DESIDOC Journal of Library & Information Technology, Vol. 32, No. 3, May 2012, pp. 204-213

Tools for Search, Analysis and Management of Patent Portfolios

Muqbil Burhan and Sudhir K. Jain

Department of Management Studies, Indian Institute of Technology, Hauz Khas, Delhi-110 016
E-mail: muqbil.burhan@dmsiitd.org; Skjain51@hotmail.com

ABSTRACT

Patents have been acknowledged worldwide as rich sources of information for technology forecasting, competitive analysis and management of patent portfolios. Because of the high potential of patents as an important indicator of various technology measurements and as econometric measure, patent analysis has become vital for corporate world and of interest to academic research. Retrieving relevant prior art, concerning the technology of interest, has been vital for managers and consultants dealing with intellectual property rights. Tremendous progress in the field of electronic search tools as of late has led to a specialised and less time consuming search capabilities even in the fields where search is mostly based on formulas, drawings and flowcharts. Online patent databases and various other analytical tools have given patent analysis an important edge, which otherwise required extensive and time consuming data collection and calculations. Patents provide valuable information which could be used for various purposes by industry, academia, and policy analysts. This article explores the various options and tools available for patent search, analysis and management of patent portfolios, for efficiently identifying the relevant prior art, managing their own patent clusters and/or competitive intelligence.

Keywords: Intellectual property rights, patent portfolio management, competitive intelligence, prior art

1. INTRODUCTION

In the present economy, management and protection of knowledge is of prime focus in corporate strategy. Technology monitoring systems that allow timely anticipation of technology change within the competitive environment, in a situation where a significant portion of the technical information is contained in patent documents and not published in any other form, have become the cornerstone of technology management[1]. Patents offer a wealth of information and if properly processed and analysed can yield a wealth of information on competitors' activities, R&D trends, emerging fields, collaborations, etc., apart from uncovering relevant prior art and decreasing the risk in huge R&D investments.

Public disclosure of patent application in the form of publication takes place within 18 months of its filing. In spite of a time lag of around 2 years of disclosure from the invention and widely varying filing practices, patents remain one of the most reliable and comprehensive source of information on R&D activity and together with the products in development and market data form the backbone of competitive intelligence activities in various technology sectors[2].

Patent portfolios analysis holds an important strategy for companies to give them a reasonably accurate idea of the volume of the activity in specific research areas, reveal the underlying trends, and detect emerging or hidden information or deviations from expected patterns while carrying competitive intelligence. Patent analysis is an important tool for research activities, for example on collaborations, location of research work, key inventors, licensing, etc. Irrespective of the purpose for searching a patent database, information from a patent document can be retrieved based on bibliographic data, full-text or surrogate documents. Surrogate document would mean an abstract, series of keywords, series of indexing terms, claims, title, IPC information, etc.[3]

Tremendous progress in the field of electronic search tools as of late has led to a specialised and less time consuming search capabilities even in the fields where search is mostly based on formulas, drawings and flowcharts, etc.[4] Some of the examples include Markush formulas,[5] combinatorial chemistry,[6] drawing and flowcharts in mechanics,[7] and sequences in biology.[8] Internet search has been a vital and irreplaceable tool for patent and non-patent prior art search.[9,10]

Received 09 December 2011, online published 18 May 2012

2. TYPES OF PATENT STUDIES

Most patent offices, and in particular industrialised countries, including the International searching authorities, have developed procedures for performing searches for patent applications. Patent search options are available on regional patent office sites, which offers raw patent search. With the advent of various online patent search tools and databases, the patent search and analysis has become more rigorous and convenient for purposes of industrial interest. These search tools provide various options of search, analysis, sorting and representation of valuable information present in patent data. To achieve satisfactory results, companies seek help from legal firms and/or various knowledge process outsourcing groups for a comprehensive search. These groups offer search and analysis of data for various purposes of interest to different industries. Apart from patent data, non-patent data is also of great importance depending on the types and purpose of search. Various strategies are followed to get the relevant hits (results). Huge importance associated with these searches and analysis for the companies seeking information has made patent search and analysis a lucrative and heavily paid job, charging the clients based on hourly services.

2.1 Patent Landscape Search

A Patent Landscape search is a very comprehensive patent search in a given technical discipline which offers a deeper analysis of a state of the art search. This helps in taking an informed decision based on the bigger picture offered by a systemised and analysed large pool of patent data. Patent Landscape reveals the past and present activities of various entities in a given area of technology by way of graphical representation and visuals displays of patent data and segregated based on various search fields, such as keywords, citations, inventors, assignee, classifications, etc. This helps in revealing the patenting trends, competitor's R&D over the years, monitor markets and rapid innovations in the field of interest, identify gaps in R&D, compare one's own patent portfolio to determine the valuable patents, and potential customers for licensing, etc.[1,11,12] In this kind of patent search, the definition of the technical subject matter of interest is usually relatively broad, although a combination with precise keywords of IPC sub-classes or by (range of) IPC main groups can be used to delimit the search[4].

2.2 Validity Search

Validity search involves patent and non-patent literature search based on selected claims of an issued patent and is required for several reasons. Apart from helping in the valuation of a patent where the closely related prior art may bring the value down and no such discovery will make the patent strong, validity search is brought up during litigations and more often to identify references that may challenge the presumption of validity of a subject patent.

A strong understanding of the subject and some creativity is required while identifying analogous technologies that may also fit into the claim limitations. This also requires a very broad interpretation of selected claims to successfully define the scope of a validity search. Search cut-off date or 'critical date' for search is another important factor for consideration and depends on the national laws in the issuing country from which the subject patent originates[13]. Even if such art does not seem to constitute a direct challenge to the claims, it may still form the basis for a legal argument against validity.

2.3 Prior Art, Novelty or Patentability Search

A preliminary patentability search before filing a patent application gives an idea of closest related prior art and, therefore, helps in drafting the patent claims 'around' the existing information. This helps in explaining the novelty of the technology to the examiner. Patentability searches are, basically, quick searches often carried out to quickly scan a large set of search results for prior art that appears relevant to the main idea. This also aids in drafting patent specifications and defining the appropriate breadth of claims. Novelty being the universal requirement for patentability, these types of searches are very widely carried globally in order to explain this criteria in the patent application. Other criteria like 'non-obviousness', 'usefulness or industrial application' etc. are also included in most patent laws. These criteria influence the formulation of claims[14]. Although 'Novelty' criteria is defined by the technical features explained in claims, criteria like "non-obviousness" cannot be devised without taking into account the way the prior art will be used during the substantive examination phase[4]. This influences the search strategy and therefore, one should ideally carry prior art search like the patent examiner.

2.4 Freedom to Operate or Infringement Search

A 'freedom to operate (FTO) search' examines the claims language of third-party in-force patents to identify those enforced patents or published patent applications with claims that cover the target technology, process, or product and is typically conducted as due diligence to assess the risk of potential infringement. Infringement or FTO search enable the company to determine the barriers, if any, in the form of patents in their territories of commercial interest for exercising desired industrial activities[8]. This also provides direction to product development programmes and uncovers licensing needs. These searches hold very high importance for companies, especially, before launching a new product. Failure to identify patent that may be infringed by the product can result in heavy financial losses in the form of costly legal actions with uncertain outcome[15]. The search includes active patents, published and pending patent applications, and also expired patents (as potential clearing documents). Broad full-text searches are performed on the patent databases in the countries of interest. Final results are based on a full review of the claims of the most relevant located patents. Due to this high importance associated with this type of search, companies seek stringent search strategies for precision in the results. Sometimes, other non-subject data (e.g. country codes, expired or not patents, own patents, non-patent literature) are also used to improve precision, within the given constraints[16].

3. SEARCH SYNTAX

Generally, the search engines offer their own syntax definition queries, nevertheless understanding them help in precision of search statements.

3.1 Boolean Operators

Widely used in programming and forming database queries, Boolean operators define the relationships between words or groups of words, e.g., use of 'AND' for a narrow search and retrieving of records containing all of the words it separates. Similarly, 'OR' can be used for a broad search and 'NOT' for very narrow search. Relative positions are of little relevance in operators like 'AND', 'OR' etc.

3.2 Proximity Operators/Searching

Proximity operators/searching limit the number of words between the search terms, e.g., use of no operator helps in finding the words as a phrase and within 'X' finds words in a specific radius. The proximity operators like 'NEAR', which is same as 'WITHIN10' finds words within 10 words of each other and 'BEFORE'/'AFTER' help finding words in relative order.

3.3 Wildcard Symbols

Wildcard symbols expand the scope of search. Truncation (*) expands the search term to include all forms of root word and can be limited by word length, e.g., Gen* to retrieve Gene, Genes, Genome, Genetic, etc., which can be limited to Gen*2 to limit it to Gene, Genes and exclude Genome, Genetics, etc., from the search.

3.4 Nesting

Nesting is done to expand or limit searches using various Boolean operators together by use of parentheses, e.g. (X OR Y) NEAR (Z AND S).

4. SEARCH STRATEGIES

A comprehensive search strategy involves a well planned methodology and use of many different factors and tools. In this section, the methodology and some important factors, to be taken into consideration as a part of search strategy, are discussed.

Table 1. Search syntax widely used during patent search

Search syntax	Operator	Use
Boolean	AND	Narrow search and retrieve records containing all of the words it separates
	OR	Broaden search and retrieve records containing any of the words it separates. The | can be used instead of 'or' (e.g., 'mouse | mice | rat' is equivalent to 'mouse or mice or rat').
	NOT	Narrow search and retrieve records that do *not* contain the term following it.
Proximity	No Operator	Find words as a phrase, e.g., life stage transitions retrieves records containing the three words immediately adjacent to one another and in the same order.
	WITHIN "X"	Find words within a specified radius, e.g., carbon within 3 fiber retrieves records that contain *carbon* and *fiber* in any order and within a three word radius of one other. Any number may be used to determine the proximity radius.
	NEAR	Find words within 10 words of each other, e.g., (women near violence)—retrieves records that contain *women* and *violence* in any order and within a 10 word radius of one other. Note: near is the same as within 10.
	BEFORE	Finds words in a relative order, e.g., social before security. Note: adjacency is not implied.
	AFTER	Finds words that contain words in the relative order specified with the after expression, e.g., scope after science. Note: adjacency is not implied.
Wildcard symbols	*	Truncation—This expands a search term to include all forms of a root word, e.g., patent* retrieves patent, patents, patentable, patented, etc.
	*	Multi-character wildcard for finding alternative spellings. Use to indicate an unlimited number of characters within a word, e.g., behavi*r retrieves behaviour or behavior.
	?	Single-character wildcard for finding alternative spellings—The ? represents a single character; two ?? represent two characters; three ??? represent three characters, and so on. Use within or at the end of a word, e.g., wom?n finds woman as well as women, and carbon fib?? finds carbon fiber or carbon fibre.
Nesting	( )	Group words or phrases—when combining Boolean phrases and to show the order in which relationships should be considered: e.g., '(mouse or mice) and (gene or pseudo gene)

A search process should be organised in a systematic and structured manner, and proceeded in a logical way. One of the approaches in carrying structured search strategy is to try the "most rational path"[17]. To help building a systematic search process and get relevant hits without missing any important document, relative relevance of the search keys for the corresponding searched concepts should be taken into account. Crude keyword-based search strategy could also be used to carry title-abstract-based analysis of patent document to retrieve relevant patent documents for further analysis. A focused keyword search strategy can be merged with the final set of crude keyword search results to obtain a final set of relevant documents where the probability of getting more hits is high. One can also start with a very narrow and precise set of keywords with a relevant classification, followed by a broad keyword based search. With the use of broad keywords, the precision in the hits will decrease and the number of retrieved documents will increase. This can be used to judge the quality of search strategies and better explained by the concepts of 'recall' and 'precision'[16,18].

'Recall' is the ratio of relevant retrieved documents to the number of total relevant documents, while 'precision' represents the ratio of relevant retrieved documents to the number of retrieved documents, e.g., when a search engine returns 40 pages, and only 30 of which are relevant while failing to return 50 relevant pages, its precision is 30/40 = ¾ while its recall is 30/70 = 3/7. The aim of an analyst is, therefore, to have high recall and a high precision[19] which means retrieval of all relevant documents and a low proportion of non-relevant documents. Recall and precision, however, share an inversely proportional relationship-an increase in precision will lead to decrease in recall. Therefore, an analyst needs to find a proper balance between the two.

Before beginning the search, it is necessary to understand some vital elements associated with the search. A healthy conversation with the client (R&D experts, legal firm etc.) to gain an understanding of essential points will make sure that spurious searches and hits are avoided. This also helps in generating the best possible keyword search. It is also quintessential to understand the legal motivations and ramifications for the search. The legal aspects helps apprehend not only why but also how search is to be conducted. An understanding of subject matter and technical background of the subject is always an advantage. It helps in an efficient search. However, the analysts are often made to adapt to subject areas that are only tangentially related to their areas of expertise. Therefore, a basic level of background competency should be gained before devising a search strategy and proceeding with that.

A search strategy is devised based on the search scope to select the sources in accordance with the motive behind. Therefore, an understanding of the search scope is also an important step at the outset of devising search strategy, e.g., infringement or FTO studies will include searching only active patent documents while as a patentability or prior art search will require searching any publicly available information anywhere. While proceeding with the search, one should carry some flexibility of adjusting the strategy iteratively, depending upon discovering of alternative keywords, classifications, etc. An exhaustive analysis which may include review of detailed description and claims should also be conducted for the best search strings. However, if the search strings seem reasonably relevant, investigating all the results is suggested depending on the time availability for carrying the search. It is also advised to start with a very narrow, targeted quick search at the outset followed with structured queries and broad searches.

4.1 Quick Search

Quick search involves use of those search keys/ classifications that yield the relevant results with highest probability. Quick search is generally carried out with a very narrow set of keywords in combination with relevant classification codes, using different set of Boolean operators like AND, NEAR, etc. This is done with a precise choice of keywords, avoiding the use of synonyms of more general meaning. Truncations are also avoided in this type of search. Quick searches are the first rational step carried in order to yield results which give a fair idea of further searches and help in revising the search strategy. This is also done for projects of short tenure.

4.2 Classification-based Search

Searching patents based on their classification is being done traditionally. There have been several patent classifications in place and widely used for searching patents. International patent classification (IPC), established under Strasbourg Agreement, 1971, is being used by majority of patent offices to classify their patent documents. US patent classification system, which is applicable only to US patents, also holds importance because of their economic advantages. ECLA (European Classification System), DEKLA (German classification system) and Japanese F-term classifications are extensions of IPC system which complement other systems by offering a means for searching patent documents from different viewpoints. They are being used to divide IPC classes into smaller sub-sections. Apart from these classifications, Derwent Patent Classification System (DWPI) is also being used by various commercial patent databases like Thomson Scientific to classify their prodigal patent repository. Patent classifications hold much importance in patent search and analysis. Different names for same inventions are often used and some variations may be added during translation of patent applications in different languages. Use of patent classifications is of great value in such situations to retrieve the relevant documents which may, otherwise, be missed. Moreover, the concept of classification is applicable to patent documents irrespective of their language or terminologies being used to describe the content. Therefore, the classification system would help in finding those documents using different languages or jargon that is liable to be missed in the keyword based search. Classification search is, generally, carried by identifying the relevant classes by either going through the various classes of the relevant classification or through initial search. A highly targeted keyword based search may be carried at the outset to identify few relevant documents. Considering their classification codes further classification based search can be carried out. The documents in the later search may help in finding few more classification codes for reviving the classification based search further. Boolean operators can also be used between the classification codes for narrow or broad searches.

4.3 Sequence-based Search

Nucleic acid (NA) and amino acid (AA) sequence information is another challenge faced by researchers and analysts during normal patent search. A very less percentage of sequences have been patented or under the process of patenting. Only about 5 per cent of total NA sequences and around 29 per cent of AA sequences have been patented or are under the process of being patented globally, out of which only US contributes over 70 per cent followed by UK (10 %) and Japan (6 %).[21] This sequence information is disclosed in different formats at various locations in the patent document, e.g., claims, detailed description, electronic tables, and/or drawings. Mostly, sequences are retrieved from 'sequence listing' which is in a computer readable format, having a sequence identification number (SEQ ID NO). Sequences known earlier are, generally, referenced with a public databank accession numbers[22]. Patent sequence data is available from public as well as commercial sources. GenBank/ Entrez, etc. represent the public efforts, while as CAS registry/Derwent GENSEQ is the examples of commercial entities offering patent sequence information. GenBank registers nucleotide sequences while as GENESEQ and CAS registry include corresponding translated peptide sequences also. Important sequence databanks NCBI, EMBL, and DDBJ are under the agreement known as the International Nucleotide Sequence Database Collaboration (INSDC) to exchange their sequences with each other which reach through trilateral patent offices[23] as shown in Fig. 1.

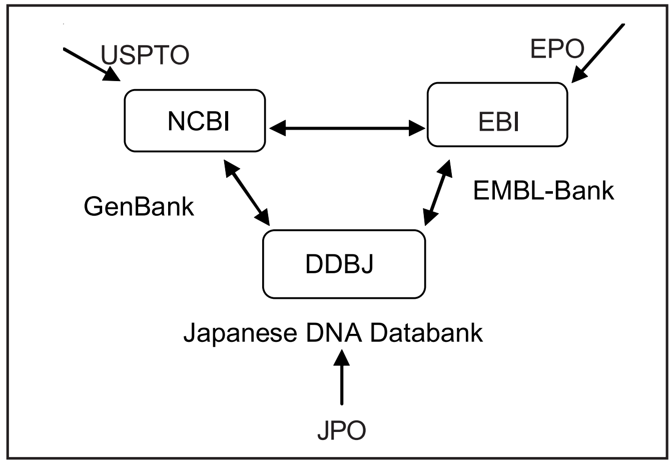


Figure 1.Trilateral exchange of sequences under INSDC.

Sequence databanks like GenBank provide a well developed similarity search interface at NCBI web system[24]. Search engines like STN provide a good platform for search on GenBank, GENESEQ and CAS registry etc. SciFinder is another good source for exploring databanks like CAS Registry for BLAST search. Similarly, Doubletwist web system can also be used for accessing GENESEQ. In absence of any international agreement for exchange of protein sequences, EBI reformats the JPO and USPTO sequence data regularly, providing those for similarity searching using the FASTA algorithm[24]. Table 3 provides a detailed list of databanks and their wide coverage along with their sources and access points.

Table 2. Master documentation databases containing family and legal status information

Database	Producer	Coverage
DOCDB	EPO	EPO master documentation database containing patent bibliographic and family data that covers patent publications from over 90 patenting authorities worldwide[20].
International Patent Documentation Centre, Patent Family Service (PFS)	Earlier produced by WIPO but later acquired by EPO in 1991	Merged into DOCDB in 2007, INPADOC PFS contained multi country bibliographic data.
International Patent Documentation Centre, Patent Register Service	Like INPADOC PFS, now part of the European Patent Office's European Patent Information and Documentation Systems Directorate (EPIDOS).	Worldwide legal status database sourced from many different patent authorities.
DWPI	Thomson Reuters	Patent bibliographic and family database with DOCDB coverage. Moreover, it has human generated indexing for searching chemical substances, structures, documents etc.

Table 3. Databases for sequence search

Databank	Producer	Description
GenBank	National Institute of Health, US	Annotated collection of all publicly available DNA sequences since 1982 from European Molecular Biology Lab, Data Bank of Japan and GenBank. Available on National Center for Biotech ftp site.
GENESEQ	Thomson Scientific	Covers all biological sequences patented since 1981 from 41 patent issuing authorities. Available through European Molecular Biology Lab like ftp site or alternatively through GenomeQuest, STN and Thomson Pharma.
Chemical Abstract Services Registry/ CAplus	CAS	Basic patents/applications published by 50 patent authorities from 1999 onwards. Accessed through STN and SciFinder.
PCTGEN	WIPO (Producer) FIZ Karlsruhe (Database supplier)	Contains nucleotide and amino acid sequences submitted as World Intellectual Property Rights Organisation/PCT applications. GETSIM (based on FASTA algorithm) and BLAST included for similarity search. Accessed through STN.
USGENE	Sequence Base Corp., US	Provides searchable access to peptide and nucleotide sequences in all published/patent applications within 3 days of their publication since 1982. Available through STN or in house customised subscription.
Genome Quest IP	Genome Quest, Inc.	Contains nucleotide and protein sequences listed in applications from USPTO (since 1980), EPO (since 1979) and WIPO (since 1980). It also includes INPADOC, GenBank, European Molecular Biology Lab, Data Bank of Japan, SIPO information which is weekly updated. Algoriths include BLAST and GenePast. Accessed easily through Genome Quest web interface.
European Molecular Biology Lab Bank	European Molecular Biology Lab - European Bioinfo. Inst	DNA and RNA sequences from EPO, JPO and USPTO since 1982. European Bioinfo. Inst. also allows similarity search against protein databases using FASTA. Access through FTP archives, European Molecular Biology Lab sequence version archive (SVA).
Data Bank of Japan	NIG, Mishima, Japan	Nucleotide sequences from EPO, JPO and USPTO since 1982. Access through Data Bank of Japan FTP server and web API.
Patent lens	Cambia, Australia	DNA, RNA and protein sequences extracted from USPTO patented/published applications. Uses NCBI's BLAST software. GenBank Id can also be used for searching sequences.

4.4 Searching Chemical Patents

Like polypeptides other chemicals also have their characteristics that make them different to search with the usual structure and/or text searching techniques. Chemical abstracts services (CAS) offers a database which includes patents published by around 26 regional patent offices, containing million of abstracted and indexed patent documents in all areas of chemistry and chemical engineering. CAS registry members are the unique identifiers assigned by CAS to compounds in their database that help in identifying them and the associated patent documents, since 1957, in the patent database.Moreover, CAS registry file includes substance information, for these indexed compounds. MARPAT is another important resource in CAT database that has an advantage over CAS Registry and BEILSTEIN files of CAS in structuring Markush Structure[25,26]. CAS REGISTRY and BEILSTEIN files allow searching only specific structures reported in chemical literature.[27]

On the other hand MARPAT file includes generic (Markush) structures representing more than one substance. These generic prophetic substances[27] represent virtually infinite set of substances by using generic terms like heterocyclic, alkyl, carbonyl, etc. The query structures on STN for MARPAT file allows use of specific atoms/symbols e.g., PH, O, N, ME, etc. and/or generic nodes, e.g., HY for hydrogen cycle.

MARPAT and CA files can be used together to obtain useful information or to transfer an answer set of one file to another for a narrow search by combination of two. This can also be done for retrieving INPADOC family information retrieval.

4.5 Text Searching

Using one or more specific keywords query indexed data, bibliographic data, title-abstract or full-text search or just claims. A researcher can search for relevant patent data which will also include patent documents which may, otherwise, be improperly classified. Most of the databases allow search based on text search, which can be combined with classification or other types of searches to carry out a narrow precise search.

A combination searching of text and classification is a powerful method for retrieving relevant data, but should be carried with passion. This may exclude some relevant data. While searching for data, 'backward citation' and 'forward citation' (references from relevant patent documents) should also be analysed. A file wrapper analysis, which includes various office actions and other important communications related to patent documents, is of particular importance for certain kinds of searches, e.g. validity search. Moreover, date range, which holds legal importance, also primarily important in case of data search analysis for FTOs, validity searches, etc.

Table 4. Databases for chemical substance search

Database	Producer	Access	Update	Coverage
CAS Registry	CAS	STN SciFinder	Daily	Substances including protein and peptide sequences searched using 1 or 3 letter abbreviations. Information about chemicals (organic, inorganic, minerals, salts, alloys, sequences, mixtures etc.) from patents, journals, catalogs, reliable web sources and GenBank seq. etc.
CAplus	CAS	STN SciFinder	Daily	Patent (61 patent authorities) and journal references since 1800 onwards. English language summaries translated from scientific literature published in >50 languages from >180 different countries. Conference proceedings, technical reports, dissertations also included.
CASREACT	CAS	STN SciFinder	Weekly	More than 37.9 million single and multi step reactions since 1840 onwards from millions of published journal articles and patent documents.
CHEMLIST	CAS	STN SciFinder	Weekly	Electronic collection of more than 2.93 lakh chemical substances that are regulated in key markets around the world.
CHEMCATS	CAS	STN SciFinder	Twice per week	Commercially available chemicals, catalog database containing information about >54 million commercially available chemicals and their suppliers.
MARPAT	CAS	STN SciFinder	Daily	Includes generic (Markush) structures (1961 onwards) representing more than one substance. Also includes citations to patents that include Markush structures.
MMS (Merged Markush Service)	Questel	Questel	-	Markush chemical substance search indexed by Thomson Reuters and INPI, French Patent & Trademark Office. Includes both generic and specific chemical structures.

5. PATENT INFORMATION SYSTEMS AND ASSOCIATED TOOLS FOR IP PORTFOLIO MANAGEMENT

Knowledge of various information systems is essentially required for a reliable patent prior art search and analysis. Various studies have recognised this fact and tried to explain comparatively a number of major patent information systems.[22,28,29,30,31] Intellogist[32] is an interesting example of a patent database assessment tool that provides a comparison of patent databases. JISC Academic Database Assessment Tool[33] is another example of comparison strategies for non-patent literature.

Apart from patent search and information retrieval, effective intellectual property portfolio management would also require sharing and collaborating information within different departments of an organisation. This is also required in case of outsourcing for various types of patent studies. ‘IP decisions are taken with contributions from different departments that include R&D, legal, finance, technology transfer, product development, competitive intelligence, etc. Therefore, a proper IP portfolio management would also require a streamline workflow processes with collaboration throughout the organisation. This helps in improving decision making within the system and reduces the time and expenses associated with IP management.

Various tools offered by different patent databases are of particular importance in this regard. Table 5, provides a concise list of some of the databases and various options offered by them that help in efficient management of IP portfolios.

5.1 Use of Macros

Macro is a collection of commands written in Microsoft Visual Basic for applications that can automate various programs such as repetitive document production tasks, etc. to save lot of time. Macros can also be used to create custom add-ins that include template from a dialogue box or to store information for repeated use. Use of macros in Microsoft Excel can be of great use in increasing the efficiency of patent portfolio analysis. Creation of macros like Family Sorter, Legal Status Extractor, Split Count, Sanitizer, etc., helps in creating a final report of relevant data in a Microsoft Excel, with current legal status and other bibliographic data extracted and sorted automatically along with INPADOC family member details without spending much of time on that. Moreover, this makes sure that redundancy of data is avoided.

A final report can have embedded search interface in Excel which helps the reviewer trace the results with ease based on various keywords like inventor, assignee, country, etc.

6. CONCLUSIONS

Patents are rich source of information and hold a high potential as an important indicator of various technology measurements. To retrieve the valuable information from the patent documents in an efficient and cost effective manner, it is essential to explore the methods of patent search strategies to devise the one suitable to one’s needs. With the advent of online patent databases and various tools that allow an effective and reliable patent search and analysis, managing patent portfolios has been

Table 5. List of some important free and commercially available databases for IP portfolio management

Producer	Database	Coverage
Thomson Reuters	Thomson Innovation	Wide patent coverage ThemeScape Maps: for identifying and relating predominant concepts. Clustering Tool – Quickly find valuable relationships through linguistic analysis of search terms. Citation Maps – Trace the history of an invention. Charting – Instantly create lists or charts that are meaningful to your search.
	Derwent Innovations Index	Includes patent coverage to 1963 and patent citations to 1973, from 47 patent-issuing authorities. Includes patent families, rewritten descriptive titles, abstracts, DWPI class codes, and cited patents.
	Aureka	Conduct competitive IP portfolio analysis and for supporting R&D project prioritisation and optimisation of R&D spending.Visualises data to reveal trends and opportunities and creates detailed reports and export data for easy analysis. ThemeScape maps to transform complex data into a visibly appealing landscape for quick review. Aureka Directory Tree and Citation Trees help in visualising the competitive activity and future trends. Aureka's Knowledge management tool creates a virtual workshop for teammates or selected individuals for collaborative contributions and sharing.
	Derwent World Patents Index	DWPI contains over 20.7 million patent families covering more than 45.2 million patent documents, with coverage from over 47 worldwide patent authorities.
	IP Payments	Helps in streamlining worldwide payment processes. IP payments calculates payment due dates using the most comprehensive IP Rules in the industry, so one can be confident he haven't missed a law change in any of the more than 400 global jurisdictions covered.
	Thomson Data Analyser	Thomson Data Analyser desktop software offers a powerful interface for managing and extracting business-critical insights from patent and scientific data within in-house or commercial databases. Thomson Data Analyser provides an easy way to analyse trends, profile competitors, avoid or uncover patent and copyright infringement, and identify strategic development opportunities in information from both in-house and commercial databases.
	Thomson IP Manager	Offers an enhanced set of tools to help organisations further protect and manage their intellectual property assets. Easily configurable software and role-based features help streamline the portfolio management process for all stakeholders, providing end-to-end management of intellectual assets and enabling collaboration between groups in the innovation and product management life cycle.
Questel	Orbit.com	Patent and industrial design search and watch covering BX, CA, CH, CN, DE, ES, EU, FR, GB, JP, KR, RU, US and WO in native languages/ machine aided/human translated English. Offers companies corporate tree and gives graphical and statistical analysis for market trends, positioning, family member relationship, potential partners or competitors, etc.
	Digipat	Provides patent copies, Us and foreign file histories. Displays 1 year of history with special features of filtering and sorting.
IP.Com	IP.com	Free database for searching over 10 million patent and patent related publications from US, Japanese, Canadian, Chinese, and European patent authority databases.
Intellectual Property EXchange Limited (IPEXL)	IPEXL Patent Search	Free databases for patent search with a focus on Asian patent offices and multilingual support. Patent search results along with bibliographic details can be exported to Microsoft Excel for analysis.
The Linux Foundation	Linux Foundation Patent Commons	Database of IT patents which the proprietors have pledged not to enforce if open-source conditions are met.
Stroke of Color, Inc.,	Patent Fetcher	Free database for US patent and published patent application PDF files
Cambia	Patent Len	Free database for full-text and status search of EP, WO, EP, and US patents
WIPO	PATENTSCOPE Search Service	Free database for full-text search of patent documents including published international patent applications (PCT).
Chemvalet Inc. and Multimus Information Technologies Inc.	PharmaValet	Database of approved US and Canadian drugs with an associated patent information like expiration, approval, extension dates, patent family and legal status. It also offers supplementary protection certificate (SPC) and patent term extension information which is particularly of importance for generic (drug) industry.

a less cumbersome job and also added much value to it. Therefore, it becomes important for IP professionals to have a better understanding of these databases and tools to be in a position to find and choose the right tool to suit his/her requirements in patent analysis and management.

REFERENCES

1. Altschul. S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W. & Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.*, 1997, **25**, 3389-402.

2. Andree, P.; Harper, M.; Nauche, S.; Poolman, R.; Shaw, J.; Swinkels, J. & Wycherley, S. A comparative study of patent sequence databases. *World Patent Inf.*, 2008, **30**(4), 300-08.

3. Archontopoulos, E. Prior art search tools on the Internet and legal status of the results: A European Patent Office perspective. *World Patent Inf.*, 2004, **26**,113-21.

4. Chakravarti, A.K. & Krishnan, A.S.A. Technology alert patent search system in electronics. *Elect. Inf. Plan.,* 1996, **23**(6), 313-14.

5. Claus, P. Patent search and information in nineties. *World Patent Inf.,* 1982, **4**(4), 149-54.

6. Community dedicated to comparing major patent search systems. *World Patent Inf.,* 2011, **33**,168-79.

7. Davis, Smith E. A comparison of some patent databases. *World Patent Inf.,* 1988, **10**(1).

8. Ebbe, T.; Sanderson, K.A. & Wilson, P.S. The chemical abstract service generic chemical (Markush) structure storage and retrieval capability. *J. Chem. Inf. Comp. Sci.*, 1991, **31**, 31-36

9. Fisanick, W. The chemical abstracts service generic chemical (markush) structure storage and retrieval capability. *J. Chem. Inf. Comp. Sci.*, 1990, **30**, 145-54.

10. Fletcher, J.M. Quality and risk assessment in patent searching and analysis. Recent advances in chemical information. *In* Proceedings of International Chemical Information Conference, edited by H. Collier, Montreux. 1992

11. Foglia, pasquale. Patentability search strategies and the reformed IPC: A patent office perspective. *World Patent Inf.*, 2007, **29**, 33-53

12. Gann Xu, G.; Webster, A. & Doran, E. Patent sequence databases. *World Patent Inf.*, 2002, **24**, 95-101.

13. Grandjea, N.; Charpiot, B.; Andres, Pena C. & Peitsch, M.C. Competitive intelligence and patent analysis in drug discovery mining the competitive knowledge bases and patents. *Drug Discovery Today: Technologies*, 2005, **2**(3), 211-15.

14. Harter, S.P. Search term combinations and retrieval overlap: A proposed methodology and case study. *J. Am. Soc. Inf. Sci.,* 1990, **41**(2), 132-46.

15. Hoetker, G. Patterns in patents: Searching the forest not the trees. *EContent*, 1999, **22**(5), 37-8, 40-2,44-5.

16. Hunt, David; Long, Nguyen & Matthew, Rodgers (Ed.). Patent searching: Tools & techniques. Wiley, Hoboken, NJ, 2007.

17. JISC-ADAT. http://www.jisc-adat.com/adat/home.pl.

18. Karasinka, E.H.; Ebbe, T. & Corning, J.F. New patent information resources from chemical abstract service. *World Patent Inf.*, 1991, **13**(3), 160-65.

19. Koschatzky, K.C. Praktische Durchfu hrung von Recher-chen in patentdatenban-ken. *Informations Technik*, 1991, **33**(5), 259-62.

20. Kristin, Whitman. Intellogist. *In* Information sources in patents, edited by Stephen R. Adams, Ed. 2. K.G. Saur, Munich, 2006. p.148.

21. Lambert, N. Internet patent information in the 21st century: A comparison of Delphion, MicroPatent, and QPat. *In* 2004 International Chemical Information Conference & Exhibition, Annecy, France, 2004.

22. Marie-Julie, J.M. Searching in flowcharts–A PD tools-doc pilot project at the borderline between text and image to access the most important features of an invention. EPO Internal Publication, 2005.

23. Michel, J. & Bettels, B. Patent citation analysis–a closer look at the basic input data from patent search reports, Dordrecht. *Scientometrics*, 2001, **51**(1), 185-201.

24. Schwander, P. An evaluation of patent searching resources: Comparing the professional and free on-line databases. *World Patent Inf.*, 2000, **22**(3), 147-65.

25. Simmons, E.S. Markush structure searching over the years. *World Patent Inf.*, 2003, **25**(3), 195-202.

26. Stevensborg, N. Searching combinatorial chemistry: From biochips to catalysts and beyond–an EPO perspective. *World Patent Inf.*, 2002, **24**, 13-23.

27. Stock, M. & Stock, W.G. Intellectual property information: A comparative analysis of main

information providers. *J. Am. Soc. Inf. Sci. Technol.,* 2006, **57**(13), 1794-803.

28. EPO. Useful tables and statistics, coverage and codes.http://www.epo.org/patents/patent-information/raw-data/useful-tables.html.

29. Van der Drift, J. Effective strategies for searching existing patent rights. *World Patent Inf.,* 1991, **13**(2), 67-71.

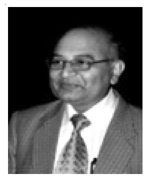
30. Wu, C. & Liu Y. Use of the IPC and various retrieval systems to research patent activities of US organizations in the People's Republic of China. *World Patent Inf.,* 2004, **26**(3), 225-33.

31. Yoo, H.; Ramanathan, C. & Barcelon-yang, C. Intellectual property management of biosequence information from a patent searching perspective. *World Patent Inf.,* 2005, **27**, 203-11.

32. Ziman, J.H. Information, communication, knowledge. *Nature,* 1969, **224**, 318-24.

About the Authors

Mr. Muqbil Burhan persuing PhD from Department of Management Studies, IIT Delhi. He is working in the field of patent portfolio management and technology transfer. He has a graduation and post-graduation in the field of Biotechnology. He has also worked as IPR specialist at Evalue Serve, Gurgaon.



Prof. Sudhir K. Jain is MHRD IPR Chair Professor and the Chairman of Economics & Entrepreneurship Group and Head, Department of Management Studies, IIT Delhi. He has published more than 100 research papers and has supervised 15 PhD theses. He has been associated with WIPO in various capacities and member of several committees set up by the such as IPR Education Policy Committee, National Entrepreneurship Policy Committee (Ministry of MSME), and MHRD Technical Committee on WTO Matters (Ministry of HRD).