CHIMIA 2000, 54, No. 5

Chimia 54 (2000) 281–285

ISSN 0009–4293

Patent Information on the Internet

Thomas P. Stauffer*[a], André J. Escher[a], and Andreas Litscher[b]

Abstract: Many people believe that the entire world's information is somehow accessible on the Internet and costs virtually nothing to retrieve. Others view the Internet as a slow, overrated mechanism for obtaining information that may or may not be reliable. For the information specialist, the Internet has become another useful resource for obtaining information. This article is an attempt to provide an overview of free and commercial sites containing patent databases and to outline the possibilities as well as the limitations of these patent information sources. In addition to this written article, we have set up a Web site (*http://chemsear.ch/chimia/*) which contains all the important links as well as additional information about patents on the Internet.

Keywords: Databases · esp@cenet · Internet · Patents · Search

1. Patent Sources on the Internet

Until recently, the patent specialist had a complete monopoly on patent work. By the rapid evolution of the Internet this needs no longer be the case. A wide variety of information about patents is available on the Internet. The spectrum goes from searchable patent databases to press releases of intellectual property cases. Of main interest are the publication and distribution of patents and patent applications in patent databases, since crucial information is buried in the patents themselves. Information which helps answering questions like: Is this specific invention already protected by a patent? Do I infringe a patent by using this specific technique? Who owns the key patent in this technology?

We will show where you can retrieve this patent information and also discuss which source may be favored for a specific task. Finally, we will also discuss the limitations of these databases, since a number of questions can not be satisfyingly answered using the sources on the Internet.

*Correspondence: Dr. T.P. Stauffer
[a] Institut für Geistiges Eigentum IGE
Einsteinstrasse 2
CH–3003 Bern
Tel.: +41 31 322 48 39
Fax: +41 31 350 06 92
E-Mail: Thomas.Stauffer@ipi.ch
[b] Infolit Information Broker
Berner Technopark
CH–3018 Bern

1.1. *Patent Databases*

A lot is free on the Internet but not everything. This is also the case for the patent databases. There are searchable databases which are maintained by the commonly known commercial hosts like Dialog, Questel or STN. These hosts provide the same services and databases on the Internet as they provide online. These services are, of course, not free. On the other hand numerous sources, mainly national patent offices, put their national patents and published patent applications on the net. Normally these databases are free of charge.

Additionally to these databases maintained by the patent offices some commercial companies offer free searchable patent databases. The most famous one is the IBM Patentserver (now called Intellectual Property Network IPN). This server is developing more and more into a intellectual property portal site. But the main part is a searchable patent database which is free of charge.

1.2. *Data Sets and Retrieval of Data From Internet-based Patent Databases*

An important factor to evaluate the right database is their data sets and search tools. There is a tremendous difference in the patent data sets and search tools between the Internet-based databases. We divide the databases into two classes. The first one contains only the bibliographic data. These are, for example, the publication, the application and the priority number; the applicant's and inventor's name; the technical field (using the IPC classification system) and the text in the title. In these so-called bibliographic databases you can search only for these bibliographic items (Figure). Therefore the common search for keywords is limited to the title. On the other hand there are full text databases which contain the whole patent document including abstract, description, claims and quite often its drawings. In these databases you can search for specific words in almost every part of the document. However the drawings of chemical structures, for example, are seldom searchable. For these tasks you need specific tools which are provided by the commercial hosts of the corresponding databases.

Introductory searches in the commercial or free patent databases do not require specific skills and are therefore useful and fun for everybody. However, both systems also allow the professional information specialists to retrieve their data quickly and precisely.

The searches can be done in different ways depending on the database you are using. Quite often it is possible to type words or numbers in the appropriate text entry areas (*e.g.* the patent number in the corresponding field) and start the query. Another way is to use the Boolean operators (OR, AND, or AND NOT) which connect the text entries or simple keywords (Figure). However as we mentioned above, each database has its own search and retrieval system which is in general self-explanatory.

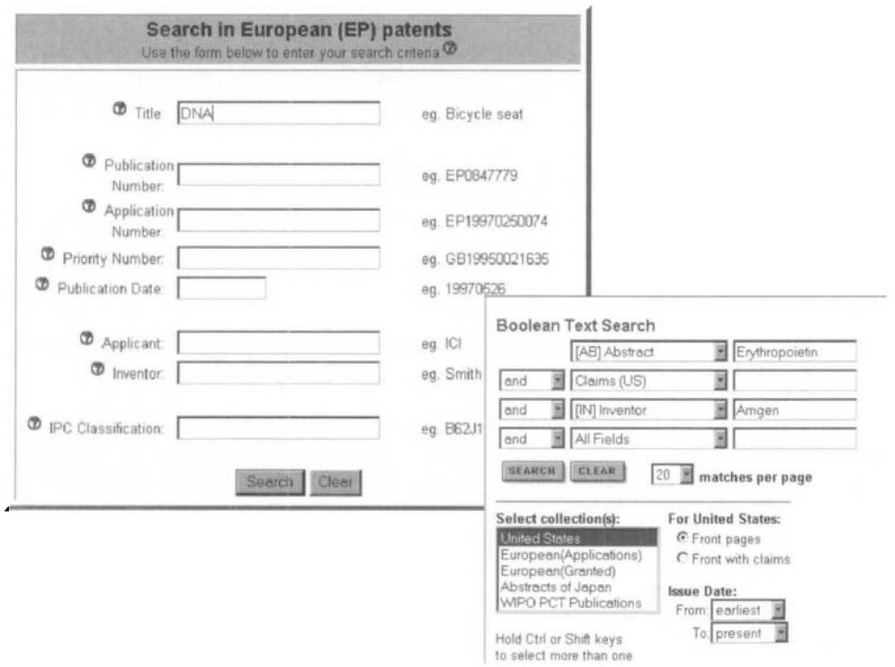

Figure: Search forms used by the esp@cenet and the Intellectual Property Servers.

2. Free Patent Databases on the Web

2.1. *Databases of National/Regional and International Patent Offices*

A number of national patent offices provide free searching of their national patents. The records often include bibliographic data and quite often the legal status of the patents or patent applications. These are perfect sources to receive detailed information about a known patent or patent application.

2.1.1. *European Patent Office*

The European Patent Office provides free information on EP patents and published patent applications through the esp@cenet: *http://ep.espacenet.com.*

Esp@cenet is a free service on the Internet provided by the European Patent Organisation through the EPO (European Patent Office) and the national offices of its member states including Switzerland. Esp@cenet is designed primarily for the general public to provide users with a readily accessible source of free patent information. It also aims to improve the kind of information that is accessible to the public. It allows searching and viewing of almost all patents published in any member state of the European Patent Organization, as well as by the European Patent Office, the World Intellectual Property Organization (WIPO) and a selected number of national patent offices. This includes the abstracts, bibliographic data and full text of patents/patent applications of the EPO, France, Germany, Switzerland, United Kingdom, United States, and WIPO. Abstracts and bibliographic data are available for China and Japan. Bibliographic data are available from more than 60 countries including Australia, Austria, Belgium, Canada, Eurasian Patent Office, India, Ireland, Israel, Italy, The Netherlands, Spain. The coverage goes back in part to 1920. However as more and more patent data become available it is necessary to check the current coverage whenever you use esp@cenet. For documents published since 1970, each patent family in the collection will have a representative document with a searchable English language title and abstract. Users will also be able to display and, where required, print out all the bibliographic data including the abstract. It is also possible to display the full-page facsimile copies of virtually any patent application available to the EPO at the touch of a button.

You can perform a quick search by searching either for a keyword, patent number or applicant's name or you do a more sophisticated search where you can search in almost every bibliographic field.

Despite the huge amount of searchable patent data in this database a couple of things are missing. For example a link to the already available register would be very helpful. Also the search and browse possibilities do not correspond to the enormous amount of data you have to search/browse through. The truncation of words, for example, is not possible. Another feature which is missing are Boolean operators. The single searchable fields are only connected with AND while the other operators are not available. Nevertheless, esp@cenet provides the public a never before seen amount of patent information and this 'for free'.

Legal Data: The Register of European Patents provides detailed information on all European applications as well as on PCT patent applications which have entered the European phase. The system provides bibliographic data such as title of the invention, classification, publication dates, name and address of the applicant, inventor, patent attorney representative, and the latest information about the status of the granting procedure of the patent application. The database is searchable with the applicant's name, the different application and publication numbers (European publication/application number; PCT publication/application number; priority number) and the International Patent Classification (IPC). The database contains no description of the invention (claims or abstracts). The files are updated daily and information is available to online users three to five days after the action date (*http://www.european-patent-office.org/epidos/epr.htm*).

2.1.2. *WIPO*

The PCT (Patent Cooperation Treaty) database contains the first page data (bibliographic data, abstract and drawing) of published patents since January 1, 1997 filed under the PCT convention. Access to the full text of PCT publications are provided by an automatic link to the European Patent Office's esp@cenet database. PCT applications published before 1997 may be consulted on the European Patent Office's esp@cenet database. All bibliographic data are searchable including publication date, international classification or application number.

Legal Data: Since the WIPO is not a patent granting authority, the PCT database contains no legal data (*http://pctgazette.wipo.int/*).

2.1.3. *Swiss Federal Institute of Intellectual Property (IGE)*

The Swiss Federal Institute of Intellectual Property provides free information on its granted patents as well as EP patents with Switzerland as designated

state by the esp@cenet. Please note that patent applications in Switzerland are published only after the patent has been granted. Therefore Swiss patent applications are not included in this database. Nevertheless the institute plans to make publicly accessible all national patents back to 1888 through esp@cenet. This database was described in more detail above (*http://www.espacenet.ch*).

Legal Data: The Swiss Federal Institute of Intellectual Property intends to publish the legal data of all national patent and patent applications by the end of June 2000.

2.1.4. *US Patent and Trademark Office*

The US Patent and Trademark Office (USPTO) now offers World Wide Web access to its bibliographic and full text patent databases. These databases are free of charge. They cover all patents starting from January 1, 1976 to the most recent weekly issued patents. The databases offer full use of all fields including bibliographic fields, for example, patent number, inventor name or US references. Additionally the full text patent database allows a search of the abstract and claim text. The user can choose between Boolean search, advanced (field) search or a search by US patent number. Similarly to Switzerland, the USPTO only publishes granted patents. Therefore no information is available on patent applications. However, recently a proposal emerged which would change the practice in such a way that the USPTO would publish patent applications 18 months after their filing date (*http://www.uspto.gov/patft/index.html*).

Legal Data: No legal data on US patents are available on the Internet.

2.1.5. *United Kingdom Patent Office*

The UK Patent Office, frequently just called the Patent Office, provides free information on its published and granted patents as well as EP patents with United Kingdom as the designated state by the esp@cenet.

Legal Data: The Patent Office provides register data on its homepage.

2.2. Intellectual Property Network (IPN) Formerly Known as IBM Patentserver

The IBM Intellectual Property Network (IPN) is a premier Web site for searching, viewing, and analyzing patent documents. The IPN provides you with free access to a wide variety of patents including the patents of the United States (from 1971), European patents and patent applications (from 1979/80), PCT applications (from 1990) as well as the patent abstracts of Japan (from 1976). Searching is fast and easy. Along with simple keyword search, IPN offers alternative searches by patent number, Boolean text, and advanced text that allows for multiple field searching. Through a review of specific classifications, you can identify topics and patents of interest. A very helpful feature is that all collections are cross-referenced and forward and backward linked to all other referencing documents. This allows immediate access to related information. An additional feature is the possibility to browse patent documents using the specific US class. All documents can be viewed and/or downloaded which will be charged separately.

What is missing? Besides its powerful performance the IPN has some substantial drawbacks. First of all the patent data provided are quite fragmentary. Specifically the European user will miss some important patent data like UK or German patents. Second the patent families are not connected. Third, the legal status of the documents are not available. However IBM plans to improve the performance of the server by adding more data and link the documents to corresponding register data. The IPN is improving and adding new features almost monthly. Therefore it is worthwhile to check on it from time to time at *www.patents.ibm.com*.

2.3. Summary: Free Patent Databases on the Internet

The Internet sites described here are useful tools in finding patent information. However, these Web sites and others may not supply 'THE' answer you could have, nor are they always the best method to retrieve the requested information. They are an excellent source to find and to look at specific patents and to check their legal status. The free patent databases are also a good starting point to get preliminary information about the patent situation in a specific field of technology. They are also excellent sources for new ideas. But they are not the source you should consult to check whether your invention is already patented or if there is the possibility of a patent infringement using a specific technique. This is mainly due to the fragmentary data set of these databases as well as the quite limited search possibilities. We will focus in the next part of the article on commercial patent sources on the Internet which overcome some of the obstacles of the free sites.

3. Commercial Sources on the Web

Convenient one-stop shopping sources for patent information are the so-called 'database hosts' like DIALOG, FIZ TECHNIK, LEXIS-NEXIS, STN INTERNATIONAL, QUESTEL-ORBIT. All of them offer their products online as well as through the Internet. The Internet services are quite similar to the online services but offer additionally the possibility of a guided search with the host-specific query language. It is important to note that searching these databases is charged in different ways, *e.g.* by connecting hours and/or by the number of citations which are checked. Using either system (guided search or retrieval language) can get quite expensive very quickly.

3.1. *Commercial Patent Databases*

We will discuss here the features of several important patent databases and focus specifically on databases with a high relevance for chemists. Since the world's principal center for chemical information including chemistry-related patent information is CAS we will discuss more in detail the advantages of the CAS databases and especially the patent searches in chemistry which are exclusively available on STN International.

Similar to the free patent databases there are bibliographic and full text databases. Currently the majority are bibliographic databases with weekly updates. Some of them even contain patent drawings. Many bibliographic databases contain not only patent data but cover also the scientific literature. These databases, like CA, CAPLUS, BIOSIS, ENERGY, FSTA, RAPRA mainly cover specific subject areas, and generally the indexing emphasizes the technical rather than the legal aspects of the original document.

3.1.1. *CAS on STN International*

The producer of the world's largest commercial databases of chemical information is Chemical Abstracts Service (CAS) in Columbus, Ohio, a division of the American Chemical Society (ACS). The principal databases of CAS, Chemical Abstracts (CA/CAPLUS file) and the REGISTRY file include 15 million abstracts of chemistry-related literature and patents and more than 22 million substance records respectively. All CAS databases are searchable on STN International (The Scientific & Technical Information Network). Chemical Abstracts Service's coverage of patents extends from the beginning of the printed CA in

1907. Enhanced coverage of patents began around 1960. There are several entries available into the CAS/STN patent databases, for beginners or experts, over the STN European Service Center at *http://www.fiz-karlsruhe.de/stn.html*. The Web-based entry point to all STN databases is called 'STN on the Web' at *http://stnweb.fiz-karlsruhe.de/*.

3.1.2. *WPINDEX*

Derwent is one of the world's largest patent services covering over 40 countries including the European Patent Office (EPO) and the Patent Cooperation Treaty (PCT) countries. The WPINDEX (World Patents Index) file covers pharmaceuticals and other categories of chemicals from 1963, agricultural chemicals from 1965, plastics and polymers from 1966 and all other chemistry from 1970. In addition Derwent's Biotechnology Abstracts (1982–) has extensive patent coverage of this relatively new subject area. The Derwent Patents Citation Index (DPCI file) covers US Examiner's citations from 1984 and EPO and PCT Examiner's citations from 1978.

3.1.3. *INPADOC*

INPADOC, the International Patent Documentation Center covers over 60 patent offices. It is the major source of patent family data and legal data. It forms the basis for the Chemical Abstracts printed patent indexes. The different IFI/Plenum files (IFIPAT, IFICDB) cover US chemical patents from 1950 onwards. More information is available at *http://chemsear.ch/chimia/*

3.1.4. *USPATFULL*

The USPATFULL file on STN has a thesaurus on the US and International Patent Classifications. USPATFULL contains the full text of all patents issued by the USPTO from 1974 onwards. The chemical patents in USPATFULL, including the CAS Registry Numbers, are indexed by CAS.

3.1.5. *PATDPA*

The German patent database PATDPA contains citations on scientific and technological patents, patent applications, and utility models filed in Germany since 1968 as well as SPCs (supplementary protection certificates) valid for Germany. Additionally you can retrieve all kinds of patent documents of the European Patent Office and the WIPO in which the Federal Republic of Germany is designated, and much more since the recent reload in 1999.

3.2. *Searching Substances, Reactions, Preparative Information*

3.2.1. *Substance Specific Searches*

The most general approach to finding patents on substances involves first finding the record for the substance in a substance-based database like the CAS REGISTRY file. Substances can be found in a number of ways. Nearly all searches for substances are based on their identifiers: Names, structures, and molecular formulas. Sometimes substances with certain properties (*e.g.* specific boiling points) are required, and properties may be used as search terms to find these substances and probably the desired patents.

Most countries allow the inventor to define the legal limits of the patent claim in both generic terms and specific terms. For chemical patents, generic inventions usually take the form of a Markush structure that contains one or more structural variables based on a list of stated alternatives. Each compound that could be constructed from the list is covered by the claim. On STN, MARPAT is the document-based file that contains the Markush structures found in patents cited in the CA file since 1988. It offers structure-based queries to hypothesized and generic substances found in chemical patent literature. MARPAT searchers use their query structures to retrieve patent citations containing Markush structures, which represent both finite sets of specific substances and infinite sets where portions of the structure have been defined using generic terms, for example: 'alkyl' and 'heterocycle'.

A MARPAT search combined with one in the CAS REGISTRY file (for exactly defined substances) offers very comprehensive coverage of chemical patents published worldwide from 1988 onwards. MARPAT already offers structure access to more than 70 million specific substance combinations. A query structure can be created to search not only in MARPAT but also in REGISTRY and BEILSTEIN. Complete bibliographic data, including the CA abstracts, are also available for direct online display or offline print for each retrieved citation. Other STN files, important to patent searchers, can also be conveniently searched and processed with MARPAT results. Structure searches are also possible using the Web version of STN but you need to download and install the structure drawing plug-in.

3.2.2. *Manufacturing Information*

Reaction databases focus on chemical reaction participants and associated information such as the yield. The database can be searched with structures or fragments of structures for reactants, products, or both reactants and products. For greater precision, reacting bonds may be specified and/or atoms may be mapped from reactant to product. In general, answers from a reaction database are chemical reactions and information about the reference from which the reaction was extracted.

3.2.3. *Reaction Databases*

STN offers four reaction databases, CASREACT, CHEMINFORMRX, CHEMREACT and DJSMONLINE. All can be searched with specific structures or substructures for the reactants, the reagents, and the products (Table). Three of the STN reaction files have patent information which differ in years covered, the sources of reaction information and the criteria for including a reaction in the database. The CAPLUS and BEILSTEIN files also contain preparative information but cannot be searched with A→B type structure reaction queries. Please check for an example at *http://chemsear.ch/chimia/*.

3.3. *Summary: Commercial Databases*

The databases on commercial hosts offer a wide spectrum of patent information. These databases cover almost all needs you may have for patent information. The Web-based hosts systems make these data now accessible to a broader audience. Together with guided search systems they want to attract non-professional users, since it looks as if the user does not need the specific skills anymore. Indeed you may want to search on these Web-based databases and you will find a number of very helpful information. There are two points you have to be aware of, because they affect the amount of money you spend for this information. The first point is that these guided search systems do not allow a very specific search, which means you will end up with quite a number of citations you have to check. Second, if you are not used to searching these databases you will stay for a long time 'online' or have to do a number of search queries. Both of it will cost you money. Summarizing these facts shows, that you still need a certain training and expertise to search these commercial databases.

4. Final Conclusions

In this article we characterized the different patent databases available on

Table: Summary of the different chemical reaction databases.

	CASREACT	CHEMREACT	DJSMONLINE *
Approx. no. of reactions	> 1.5 million (single step) > 2.1 million (multi-step)	> 390 000 (single step)	> 54 000 (single step) > 6500 (multi-step)
Sources of the reactions	Mainly the organic sections of CA: Journals 1985- Patents 1991-	Subset of the reactions from the journal and patent literature covered in the VINITI/ZIC database	Derwent Journal of Synthetic Methods (journals and patent literature)
Reactive functionality terms	Functional groups		Reacting and product keywords
CAS RNs	All reaction participants		
Yield **	✔	✔	
Number of steps	✔		✔ ***
Years covered	1985–present	1975–1995	1975–present
Update frequency	Weekly		Monthly
Producer	CAS	InfoChem	Derwent

* DJSMONLINE is available to non-subscribers. DJSMDS is available to subscribers.
** Not every reaction in the database has a reported yield.
*** Number of reaction steps is not reported for every reaction.

the Internet. Both sources, *i.e.* free patent databases and databases provided by the commercial hosts, are useful tools to get information about patents. However, the two sources have a different focus.

The free databases are designed primarily for the general public and to provide users with a readily accessible source of free patent information. They also aim at improving patent information in particular among small and medium-sized enterprises. This information is far from complete and it is hard to retrieve the complete information, since the search capabilities are quite limited. Even though these databases were created for non-professional users, a prerequisite to use these databases with all their advantages and to get the most out of them demands professional knowledge. Therefore non-professional searches will always be preliminary. On the other hand, these sources are excellent tools to catch a first glimpse of the patent situation in a specific field or to read and print out an already known patent. The fact that the different national and regional patent offices make the legal status of their patents and patent applications publicly available is a very exciting development.

The commercial sources provide similar information on the Internet as they do online. The main differences is that the commands are based on web-browser technology. Still, conventional searches with the corresponding retrieval language are possible in each host. These sources are designed for information professionals and advanced end-users, since they contain the complete content and functionality of the hosts. A novice searcher may use these systems but has to be very careful not to spend too much money for fragmentary information.

This brings us to some final remarks. Keeping in mind that the free and the commercial patent databases on the Internet are created for different customers and different purposes, both of them are perfect sources for patent information. However, the two require a certain expertise to retrieve all the relevant information and even more is needed to further analyze this information. Since the patent data are often only useful in their context, an analysis and/or interpretation of them are quite mandatory.

The Internet provides indeed a number of superb sources for patent information. To get the optimal benefit of this information you need to have the know-how to retrieve it, as well as the knowledge how to bring this information into the right context.

Received: March 7, 2000