



**DESCRIPTIONS  
OF THE MOST  
FREQUENTLY  
USED  
DATABASES**



# Descriptions of the most frequently used databases

Nordic Patent Institute utilizes the examiners of the Norwegian and Danish Patent Offices who both are members of the European Patent Organization. The examiners not only have access to the unique databases and tools most frequently used by major patent offices. We are always looking for alternative tools and database to optimize the result of their work. The examiners therefore have access to tools and databases not normally used within patent information searching which has shown to make a difference e.g. in invalidity searches.

Our unique search tools secure fast and efficient handling of searches, not only securing a very high quality but also cost efficiency. In this leaflet you will find examples of the databases most frequently used by our experienced PCT examiners.

## Patent databases

### **EPODOC**

The EPODOC database contains classified patent applications/patents from 71 patent organisations. The database covers EP, PCT, OAPI and ARIPO completely, USA from 1832, England, France and Germany from approx. 1920. Japan and other European countries are covered from 1971/74 to date. EPODOC contains the database PAJ (Japio) from the Japanese Patent Organisation. The database is updated weekly and the search is performed in the abstract of a patent/patent application filed by the applicant. A citation search in EPODOC covers those documents cited in the subject patent application, as well as the documents cited by the technical examiner. The search also covers patent literature in which the subject document or its family is cited.

### **Full text search in PCT and national patent documents**

We have access to full text search in databases that covers patent documents from US, EP, WO, GB, AU, DE, AT, CH, FR, BE, CA, RU, CN, KR, JP and more countries through several different providers. Searches are performed in original language files and/or in machine translations.

### **WPI (Derwent World Patents Index®)**

The WPI database is made by Clarivate Analytics and covers published patent applications and granted patents from 61 global sources, including 59 patent-issuing authorities and 2 literature sources.

The database is updated weekly, with editorially curated titles and abstracts, special subject matter classification coding and indexing, and proprietary patent invention-centric family building.

### **PatSeer®**

PatSeer has access to over 157+ million patent publications from more than 100 authorities. The database includes a comprehensive list of full text records from over 70 authorities, as well as machine translated applications/patents and scientific literature. The search engine supports both simple quick searches and advanced command line searches, as well as semantic searches using free text. PatSeer features a highly customizable patent data analysis tool, allowing both internally generated and imported datasets to be visualized graphically with ease.

### **PatBase®**

PatBase covers 100+ million patent and related documents from over 100 countries and it is updated weekly. The database has dynamic indexing systems. The search can be performed in the full text of a document in the original language or/and in machine translations. PatBase also offers semantic searching, citation analysis, and graphical analysis tools.

### **IPRally**

IPRally is an AI-based search tool covering patents. IPRally searches and classifies patents with knowledge graphs combined with supervised deep learning AI. The invention is defined in a graph where only the technical core of the invention is parsed to the graph.



## Asian full text patent databases

### **CNFULL (Chinese patents full text database)**

The CNFULL database via STN is produced by Questel and covers the full text of patent applications, granted patents and utility models published in the People's Republic of China. Records are available about a week after publication date with the complete content. The abstracts are initially machine translated and about three months later replaced by a human translated text. Descriptions and claims are machine translated.

### **JPFULL (Japanese patents full text database)**

The JPFULL database via STN contains the full text of patent applications, granted patents and utility models published in Japan. Records are available about ten days after publication date with the complete content. Abstracts are either machine translated or taken from equivalent documents if available. Machine translated abstracts and titles of patent applications are replaced by human translated text about three months later. Descriptions and claims are always machine translated.

### **KRFULL (Korean patents full text database)**

The KRFULL database via STN is produced by Questel and covers the full text of patent applications, granted patents and utility models published in the Republic of Korea. Records are available about four weeks after publication date with the complete content. Titles and abstracts are initially machine translated and replaced by human translated text about five months later. If an original text is not available, the English abstract is taken from equivalent documents. Descriptions and claims are machine translated.

# Descriptions of the most frequently used databases

## Non-patent literature databases Electro and mechanical engineering

### AIP

In the AIP database, full text articles from the American Institute of Physics are searchable back to 1995.

### COMPENDEX

The Computerized Engineering Index and Ei Engineering Meetings database (Ei COMPENDEX) contains citations from worldwide engineering and technology. COMPENDEX covers all engineering disciplines including chemical, computer, electrical, civil and mechanical engineering. Sources include journals, books, conference contributions, reports, and non-conventional literature. Bibliographic information and abstracts are searchable. File data from 1960 to present and the database is updated monthly.

### IEEE

The IEEE database contains documents produced by IEEE (Institution of Electrical and Electronic Engineers) in the fields of modern electronics including electronic science and engineering, telecommunications, optoelectronics and optical communication.

### INSPEC

The Information Service for Physics, Electronics and Computing database contains citations with abstracts to world physics, electronics and electrical engineering, computers and computing, and control theory and technology literature. INSPEC corresponds to physics abstracts, electrical & electronics abstracts, computer and control abstracts, and business automation. Sources for INSPEC include primarily journals, conference proceedings, books, dissertations and reports. Bibliographic information and abstracts are searchable. File data from 1898 and the database are updated weekly.

### RDISCLOSURE

The Research Disclosure database contains the full text including images of technical disclosures of inventions published as an alternative to the patent system. Each record contains the title, a detailed description and illustrations of the invention, as well as a statement as to whether the invention is disclosed anonymously or by company name. The database covers all areas of science and technology. The source is the monthly Journal Research Disclosure. File data from 1960 to the present and is updated monthly.

### 3GPP™

The 3rd Generation Partnership Project (3GPP) unites seven telecommunications standard development organizations from Europe, Asia, and North America (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC), and covers cellular telecommunications network technologies, including radio access, the core transport network, and service capabilities - including work on codecs, security, quality of service - and thus provides complete system specifications. The specifications also provide hooks for non-radio access to the core network, and for interworking with Wi-Fi networks. Specifications, draft specifications, reports, working documents, and change requests are all searchable.

### ETSI

The European Telecommunication Standards Institute produces globally-applicable standards for Information and Communications Technologies (ICT), including fixed, mobile, radio, converged, broadcast and Internet technologies. Standards and draft standards are accessible. Further relevant databases within the electronic and mechanical area are available via ProQuest Dialog, STN or the internet.



## Chemical engineering

### BIOSIS

The BIOSIS database is a bibliographic database covering worldwide research on all biological and biomedical topics. The sources are journals, U.S. patents (1986-1989, 1995 to the present), reports, meetings (abstracts and papers), reviews and books. File data from 1969 to the present. The database is updated four times per month. The search is performed on abstracts.

### Chemical Abstracts

The CAplus file is a chemistry bibliographic database available from CAS. CAplus covers international journals, patents, technical disclosures, technical reports, books, conference proceedings, dissertations, electronic-only journals and web preprints from all areas of chemistry, biochemistry, chemical engineering and related sciences from 1907 to the present. The search is performed on abstracts.

### Embase®

The Embase database is a bibliographic database that covers the worldwide biomedical and pharmaceutical literature. It is produced by Elsevier B.V. The sources are journals from about 70 countries, books, conference proceedings and reports. File data from 1974 to the present. The database is updated weekly. The search is performed on abstracts.

### FSTA® (Food Science and Technology Abstracts)

The FSTA database is a bibliographic database providing coverage of scientific and technological aspects of the processing and manufacturing of human food products. The sources are journals, books, conference proceedings, reports, patents, pamphlets, legislation and dissertations. File data from 1969 to the present and the database is updated weekly. The search is performed on abstracts.

### KOSMET (Cosmetic & Perfume Science & Technology)

The KOSMET database is a bibliographic database containing citations from the worldwide literature on cosmetics and perfumes, with an emphasis on scientific and technical research and studies. The sources are journals, conferences, letters, patents, reports and reviews. File data from 1968 to the present and the database is updated monthly. The search is performed on abstracts.

### MEDLINE®

The MEDLINE database is a bibliographic database produced by the U.S. National Library of Medicine (NLM). The database covers worldwide biomedical literature. File data from 1950 to the present and the database is updated 5 times a week. The search is performed on abstracts. Records before 1975 do not have abstracts.

### PQSCITECH

PQSCITECH (ProQuest Science and Technology) covers all areas of science and technology from engineering to lifescience back to 1962. The file is a merge of 25 STN databases formerly known as CSA databases (Cambridge Scientific Abstracts): AEROSPACE, ALUMINIUM, ANTE, AQUALINE, AQUASCI, BIOENG, CERAB, CIVILENG, COMPUAB, CONFSCI, COPPERLIT, CORROSION, ELCOM, EMA, ENVIROENG, HEALSAFE, LIFESCI, LISA, MATBUS, MECHENG, METADEX, OCEAN, POLLUAB, SOLIDSTATE, and WATER. AEROSPACE and METADEX. PQSCITECH Sources are journals, patents, books, reports, and conference proceedings. The records contain bibliographic and indexing information, and abstracts. The database is updated monthly.

### RAPRA

RAPRA covers plastics, rubber, adhesives, and polymeric composites. Sources are journals, conference papers, books, trade literature, annual reports, and patent documents related to the rubber and plastics industries and published from 1972-2021. The search is performed in abstracts.

# Descriptions of the most frequently used databases

## Structure and sequence databases

### **CAS REGISTRY(SM)**

The CAS REGISTRY(SM) File is a substance database including substances indexed in CAPlus(SM), CA(SM), and CAOLD(SM) files. Records from the CAS Registry System, which identifies substances described in journal articles, patents, conference proceedings and substances on regulatory lists, are available. Other sources are e.g. GenBank. Sequences can be searched from 1957 to present with FSTA or BLAST-algorithm. Chemical structures can be searched from 1957 to present with additional substances going back to the early 1900s. The database is updated daily.

### **CASREACT**

CASREACT is a chemical reaction database with reaction information derived from journal and patent documents. Journals covered in CA from 1985, patents covered in CA from 1991, INPI data from 1840. The database is updated weekly.

### **The Derwent GENESEQ database (formerly Dgene)**

The database contains information on nucleic acid and protein sequences extracted from Derwent World Patents Index basic patent documents. Coverage includes nucleotide sequences of 10 or more bases, all amino acid sequences of 4 or more residues and nucleic probes and primers of any length. The sources are patents from the 41 patent issuing authorities covered by the Derwent World Patents Index. File data from 1981 to the present. It is updated every two weeks.

### **DERWENT MARKUSH RESOURCE (DWPIM)**

The database DWPIM is a structure searchable database covering Markush structures from Derwent World Patents Index. DWPI Markush include indexed structures from claims, examples and disclosure. The database covers from 1961 to date and is updated weekly.

### **GenBank®**

The database GenBank® is a nucleic acid database produced by the National Institute of Health. Records in GenBank® contain sequences and data such as the GenBank® Locus Number, sequence description, source organism, sequence length and references. In addition, the file contains records with Contiguous Sequences (CONTIG) data consisting of a set of overlapping clones or sequences from which a sequence can be obtained. The data are compiled primarily from journal literature and direct author submissions from otherwise unpublished sources. File data from 1982 to the present and the database is updated daily.

### **GenomeQuest**

GenomeQuest is a web based sequence searching tool. They offer keyword search as well as four distinct search algorithms: Percent identity (GenePast), BLAST, Fragment Search and Motif search. It covers information from public nucleotide and protein sequence databases. For patent sequence searches, GenomeQuest includes the sequence patent database GQ-PAT. GQ-PAT delivers information including: Seq ID, links to bibliographic information, patent family, legal status, patent sequence location, and more.



#### **MARPAT (CAS Markush Search Service)**

The MARPAT database contains the Markush structure found in the claims for patents searchable in CPlusSM with the patent publication year of 1961 to the present. Russian patents only those published after January 10, 2000. Records from 1961-1987 are retrieved from INPI data. The structures are searchable. The sources are patents from 50 active patent-issuing organizations. File data from 1988 to the present. The database is updated weekly. The database does not include polymers, alloys, inorganic salts, intermetallics or metal oxides.

#### **ORBIT BIOSEQUENCE**

The Orbit biosequence database covers DNA/RNA and amino acid sequences extracted from patents. The database comprises over 460 million sequences extracted from 1.1 M patent publications. The sequence searching can be combined with keyword searches in patents. Blast is the default search tool available but a motif search algorithm is also available. Orbit biosequence has specific algorithms for short sequences such as CDRs.

#### **PCTGENE (World Patent Application Biosequences)**

The PCTGENE database covers nucleotide and amino acid sequence information submitted electronically to WIPO by patent applicants. File data from August 2001 to present and the database is updated weekly.

#### **Reaxys®**

Reaxys is a structure and factual database in organic chemistry. The organic substance records contain documents from Handbook of Organic Chemistry published by Friedrich Beilstein and data from 176 journals within organic chemistry. File data from 1771 to the present and the database is updated twice a week.

#### **USGENE® (The USPTO Genetic Sequence Database)**

The USGENE database covers all available peptide and nucleic acid sequences from published applications and issued patents of the USPTO. File data from 1982 to present and the database is updated weekly.

Further relevant databases within the chemical and pharmaceutical area are available via ProQuestDialog, STN or the internet.



