1 Search in full texts

1.1 Different search options for the text search

The search field used most commonly is the search field *Text.* Here the search for words in the original texts of the patent offices is carried out. Depending on the offices searches can be performed in the titles, the abstracts, the claims and the descriptions.

The text search is available in the simple search and in the extended search / research.

In the simple search, the search is always carried out "across SIP-patent". "Across SIP-patent" means that the search is carried out across the texts of all publication levels of an office. This search is designed to find as many hits as possible and mainly corresponds to the full text searches of the cost-free web portals like Depatisnet, Espacenet or Patentscope.

The search options of the text search in the extended search vary within the different Invention Navigator versions.

Search options across the different Invention Navigator Portable versions:

Invention-Navigator-Version	Detailed	Across SIP-patent	Across Family
Standard/R&D		x	
Professional	x	x	
Premium	x	x	x

The differences of the searches and the respective result lists will be explained in detail below.

In the text search you can choose if you want to carry out your research "detailed", "across SIP-patent" or "across family". To do this, click on the respective radio button.

The input of the search string is identical for all searches and will be explained in detail on the following pages.
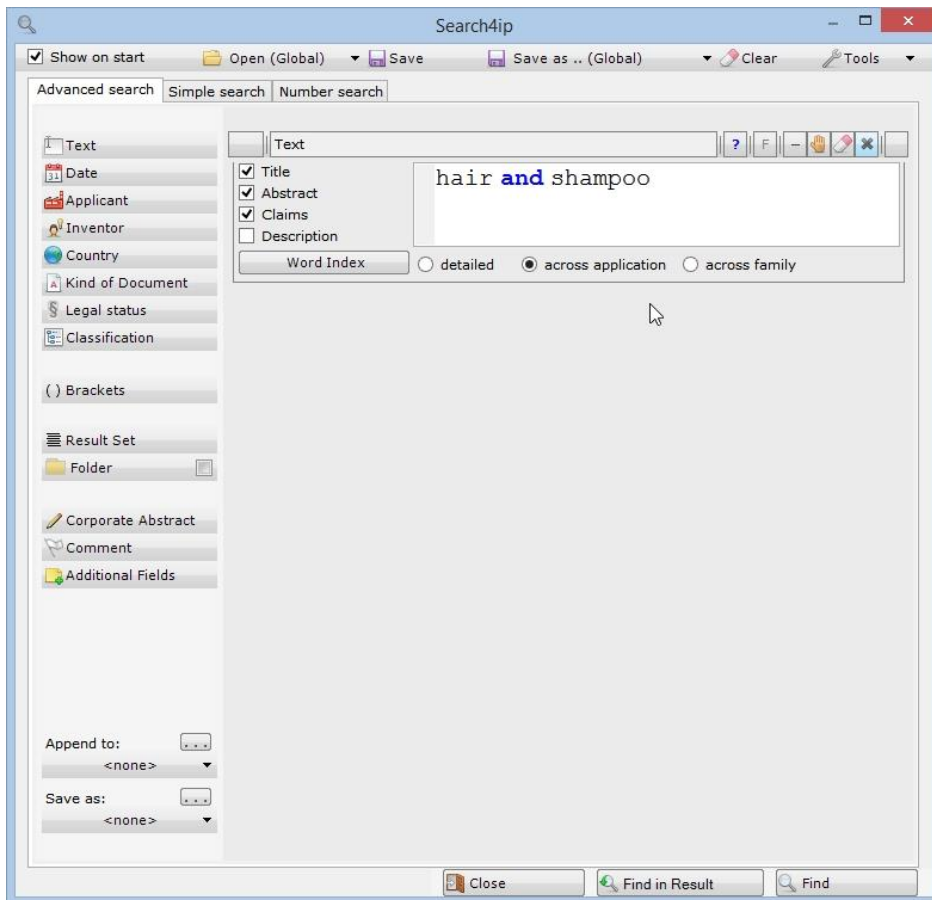
SIP Software for Intellectual Property GmbH

Gewerbering 14a   phone: +49 8024 46699-00   e-mail: service@s4ip.de
D - 83607 Holzkirchen   fax: +49 8024 46699-02   www.patentfamily.de

Image 1: Extended search in the Invention Navigator Premium

Any term entered in the text field <u>without</u> using operators like OR, AND, NOT will be automatically linked with AND.

An AND-link always has priority over an OR-link.

Example: *hair shampoo* will search for *hair* AND *shampoo.*

Search in title, abstract and claims is selected by default. By selecting the different checkboxes you can choose which areas (title, abstract, claim, descriptions) of the patent document you would like to search.

The search fields (Title, Abstract, Claim, Description) are combined with OR.

1.1.1 Selection: detailed

If you choose "detailed", you are using the search which you have already used in previous versions of the Invention Navigator solutions.

When selecting title, abstract all search terms must occur either in the abstract or all search terms must occur in the title.

(hair AND *shampoo)* in the abstract or *(hair* AND *shampoo) in the title.*

This search would not produce any results if the term *hair* would only occur in the title and the term *shampoo* would only occur in the abstract.

1.1.2 Selection: "across SIP--patent"

When selecting title, abstract, one term may be found in the title or the abstract of the application, the other term may be found in the title or the abstract of the granted patent.

(hair AND *shampoo)* in the abstract or in the title.

This search would also find hits when the term *hair* occurs in the title and the term *shampoo* occurs in the abstract of a document.

One term may also occur in the title of the application and the other term may occur in the abstract of the granted patent. This means that the search takes place in the texts of all levels of publication of the office.

Using this search will produce many more hits than the detailed search. The full text searches of most public portals perform the search across the full text of the document.

Advantages of this search: When searching in EP-patents containing different languages, additional hits can often be found if the research is carried out using multilingual synonyms. The disadvantage is that the amount of hits to be viewed is much higher.

1.1.3 Selection: "across family" (Premium version only)

When selecting title, abstract one term may occur in the title or the abstract of any family member and the other term may occur in the title or abstract of another family member.

(hair AND *shampoo)* in the abstract or the title.

This search would find hits if the term *hair* occurs in the abstract of the US-patent and in the English title of the EP-patent.

The patent family here is the extended INPADOC patent family.

This search is interesting if previous searches did not produce any usable hits, as the texts are only available in different languages with a variation in quality.

This search will produce additional hits in the research if you search using multilingual synonyms.

Which type of search you use depends on your search query. The detailed search will produce very exact hits while the search across the SIP-Patent will produce more hits. The search across family enables you comparisons with other providers who only offer family based searches. It can also produce additional relevant hits in exceptional cases. When comparing family based searches, pay attention to the different definitions of the patent family.

1.2 Display of results in the title list depending on the search

1.2.1 Amount of found hits in the title list

As of version 2.2.14 different amounts of hits will be displayed in the title list. The display of "Sip-patent", "family", "document" was hereby newly introduced. This is displayed in the bottom row of the result list.

The amount of "documents" tells you how many patent publications were found. Here every publication level of one office is counted. This means, for example, that one application EP A1 with an audit report EP A3 and a granted patent specification EP B2 leads to three hits.

The amount of "SIP-patents" tells you how many SIP-patents[1] were found. Only one hit per office is displayed. If, as in the previous example, an EP-A2, an EP-A3 and an EP-B1 were published and all three publication levels were found via a search, then not all three documents are counted, but only the Sip-patent. The hit list contains one hit.

This amount of SIP-patents corresponds to the known amount of results in the previous versions of the Invention Navigator.

The amount of families tells you in how many families the found hits occur. This information relates to the entire hit list.

If, for example, a hit is found in the DE-application, in the DE granted patent and in the EP granted patent of a family during a search, the result is displayed as one hit with the amount of families.
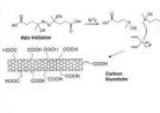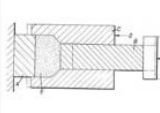
---

[1] SIP-patent is a record, which includes all publication levels of one office (country) based on one application. Example: EP A1, EPA3, EPB1

Image 2: Result list

1.2.2 The SIP-Ranking

Regardless whether a research is carried out "detailed", "across SIP-patent" or "across family", the SIP-Ranking comes into effect and the most relevant hits with a high ranking will be displayed at the beginning of the title list.

With the SIP-Ranking a patent is weighted according to the frequency of search terms in certain areas of a patent and ranked accordingly.

If, for example, both search terms occur in the title, this result will then be ranked higher than a patent where one search term occurs in the title and the other search term occurs in the descriptions. If search terms occur more frequently in the descriptions of one patent and in another patent in the claims, the patent with the search terms occurring in the claims will be ranked higher.

All publications and all publication levels of one patent family are searched. The latest publication level of the SIP-patent leading to a hit will be displayed. If, for example, the hit was found via the summary of an EP-application, the EP granted patent is displayed - even though it does not contain a summary- as the granted patent is the highest publication level based on an application.

1.2.3 Set to top in family-based view of the title list

If "one patent per family" is selected in the title list, the displayed (not the entire) result list is filtered and just one family member is displayed. It depends on your settings in Tools/Set to top which family member is displayed.

Please note that only the list displayed is filtered. Not all the hits are filtered. If, for example, you search for English keywords, then German patents (which only contain English summaries in exceptional cases) will not be found and therefore not displayed in the result. If you prioritised German patents in "Set to top" they cannot be displayed even though they are contained in the family. Of course, you can see these family members in detail view.

In the following, the possibilities of entering the search string are displayed.

1.2.4 Truncation/Wildcards

If for example the search term *motor* is entered, only the exact term *motor* will be found. In order to find the terms *motor*, *motors*, *shipmotor* etc. wildcards and truncations can be used. In the Invention Navigator you can use left, right and middle truncation as well as a combination of the three.

Wildcard Symbol		Example	Finds
*	0 or any number of characters	motor* *motor*	Motor, Motors, Motorbike, motors shipmotor, shipmotors
?	Exactly this amount of characters	motor?	motors, not: motor
%	Maximum amount of characters	motor% M%%ller	motor, motors Müller, Mueller, Muller
Combination	Combination of all wildcards	?ha*p*o%%%	channelsupport, shampooing,

There is no automatic search for the plural of a term.

1.3 Search for special characters

Umlauts or other country specific special characters can be entered and searched.

When searching for the term *Bäume* the search terms *Bäume* and *Baeume* are created and both will be searched for.

Attention with applicants' and inventors' names!

In many countries where Umlauts are not known (e.g. USA), the applicant *Müller* is entered as *Muller* in the patent document and therefore also transferred accordingly into the database. In this case a search has to be conducted using the search terms *Müller* OR *Muller*.

1.4   Phrase search

In order to search for two or more terms in their exact order, i.e. as a phrase, the search terms are entered with quotation marks. Hereby the hyphen is ignored.

Example: "omega 6 acid" finds omega-6 Acid, Omega-6-acid, etc.

Searching for "omega-6 acid" will also find all terms.

This way, whole sentences, phrases or compound expressions can be searched for.

If words like AND, OR, NOT, which also serve as operators are used as search terms, they will need to be put in quotation marks.

Attention!!

It often happens that hyphenated terms lead to false search terms when searching with an OR-link.

The following example is entered:

polypropylene* OR polyolefin* OR propylen* OR (impact AND copolymer?) OR random OR heterophasic OR homopolymer? OR "olefin polymer" OR alpha-olefin? OR a-olefin?

Now the search is converted by the search engine:polypropylene* OR polyolefin* OR propylen* OR (impact AND copolymer?) OR random OR heterophasic OR homopolymer? OR "olefin polymer" OR alpha olefin? OR a olefin? As described in the first paragraph, the input without using an operator is an AND-link.

Therefore the search is interpreted as: polypropylene* OR polyolefin* OR propylen* OR (impact AND copolymer?) OR random OR heterophasic OR homopolymer? OR "olefin polymer" OR alpha AND olefin? OR a AND olefin?

It is therefore recommended to place hyphenated terms in quotation marks or search via SPAN-Operators.

1.5 Linking search terms with Boolean Operators: AND, OR, NOT

The Boolean Operators AND, OR, NOT can be used.

The operators can be entered in lower case or uppercase letters. Usually the input with uppercase letters is clearer.

AND

If the operator "AND" is used between the search terms in the text input screen, all terms linked this way must be contained in the text.

Depending on the selection of the search type "detailed", "across SIP-Patent" or "across family" both terms must be contained either in the same area of the document (title, abstract, claim, description) respectively in the text of the SIP-patent or in the text of all family members.

Example:

Detailed search: selection of search fields: title, abstract, claims, description

Example: *Hair* AND *shampoo*

The term *hair* and the term *shampoo* have to occur in one text area of the document, e.g. in the title.

OR

If the operator OR is used between the search terms, at least one of the search terms needs to be contained in one of the selected fields.

Example: *Hair* OR *shampoo*

When using this search, the selection of the search (detailed, across SIP-patent, across family) has no impact.

NOT

If the operator NOT is used between the search terms, the selection of the type of search has a significant impact on the result.

The NOT-Operator in detailed search

Hair AND *shampoo* NOT *body* searched title, abstract, claim

This search finds all documents containing both terms *hair* AND *shampoo* in the title, abstract or claims. Documents which also contain the term *body* in either the abstract, title or claims are excluded.

Now it may happen that an EP-application whose claims and descriptions are written in German contains the term *hair* and the term *shampoo* in its English title (available in 3 languages). The granted patent contains the term *body*, as claims of EP granted patents are written in 3 languages, hence containing an English text.

The search produces one hit for the EP-application. However, the English claims contain the term *body*. The latest document of the SIP-patent is displayed, hence the B document, even though the B document contains the term *body*.

In order to rule out that no hits containing the term *body* are found, the search needs to be carried out in two steps.

Drag a text field into the search mask and search for *hair* AND *shampoo*. You will now find all hits containing both search terms in either abstract, title or claims. Then open a second text search block and enter *body*. Now combine both text fields using NOT. The second search field produces all patents containing the term *body*. Both result lists are now intersected. This way all patents also contained in the second list will be subtracted from the first list (*hair* AND *shampoo*).

The NOT-operator in the search across SIP-patent.

What happens in the same example?

Search query:

Hair AND *shampoo* NOT *body* across patent, selection title, abstract, claims

All patents, which contain the term *hair* either in the abstract, the title or the claims and contain the term *shampoo* either in the abstract, the title or the claims will be found. At the same time all patents containing the term *body* in the abstract or the title or the claims will be excluded.

This means that the EP-B will not be found, as it contains the term *hair* and the term *shampoo* while containing the term body in the claims.

As the search is now performed across the SIP-patent (all publications of an office), meaning that both publication levels are considered a unit, these hits will not be displayed.

The NOT-operator in the search across family.

Search query:

(*hair* OR Haar) AND (*shampoo* OR schampoo) NOT (*body OR körper)* across family with selection title, abstract, claims

Here, all patents containing the term *hair* or *haar* and the term *shampoo* or *schampoo* anywhere in the title, abstract or claims are found. However, if one of the family members also contains the term *body* or *körper* then all family members are excluded and will not be displayed.

Example:

The German patent document contains the term *hair* and the term *shampoo* but not the term *body* and the term *körper.* This means that the patent corresponds to the search criteria and would be displayed. If a US patent in the family exists containing the term body then the entire patent family would not be displayed, as the entire patent family is considered one unit, hence corresponding to the search query.

These examples show that the detailed search, the search across SIP-patent and the search across family can lead to very different results. It is therefore important to be aware, which results are to be expected with each search.

This is especially important when performing comparative researches with other cost-free and chargeable providers.

1.6   Simultaneous use of operators and brackets

Attention:

AND and NOT will always be processed before OR.

Example:

Search: hair AND shampoo OR body finds all documents which contain both terms *body and shampoo* or only contain the term *body*

Hair and shampoo or body matches (**hair and shampoo) or body**

When searching in full texts always use brackets or divide the searches in different search fields. The amount of brackets is not limited. This increases the clarity and the weighting of individual operators is not crucial.

1.7   Proximity operators

In most cases a limitation with Boolean Operators leads to many not relevant hits. In order to refine text searches proximity operators can be used.

Proximity operators specify exactly how many words may be present in which order between search keywords in a text.

In the Invention Navigator the operators SPAN and NEAR can be used.

SPAN

With the SPAN function you can search for keywords which occur in an exact order with a defined maximum number of terms between them.

Example:

Search: **span(**hair shampoo composition,**5)**

Finds all documents in which all 3 terms occur in this exact order with a maximum of 5 words between them.

Within a SPAN search string all wildcards / truncation characters can be used.

span(hair(shampoo OR lotion) composition,**5)** or **span(**hair (shampoo AND lotion) AND composition,**5)**

matches the query

span(hair shampoo composition,**5) OR span(**hair lotion composition,**5)**

span(hair shampoo composition, 0) matches "hair shampoo composition"

NEAR

This operator also is a proximity function. The difference from the SPAN operator consists in the fact that the terms may occur in any order.

It is also possible to use NOT in the SPAN/NEAR function.

near(hair shampoo composition NOT (body OR skin),**5)**

matches the query

near(hair shampoo composition NOT body,5) and near(hair shampoo composition NOT skin),5)

Nesting of NEAR, SPAN and phrase search is also possible.

Example for nesting multiple NEAR functions:

near(near(hair shampoo,**5) near**(composition red clay **10),50)**

Finds all documents in which

1. the terms *hair u*nd *shampoo* occur with a maximum of 5 words between them

2. the terms *composition, red a*nd *clay* occur with a maximum of 10 words between them

3. the hits from 1. and 2 occur with a maximum of 50 words between them

The proximity search always takes place in the detailed search. It is not possible to use proximity operators when searching across SIP-patent / across family.

1.8 Extend to family

Another feature, which has a big impact on the results, is the option to "extend to family". This feature is not available for Invention Navigator (R&D, Standard).

If you carry out a "detailed" or "across SIP-patent " text search with English keywords you will only find patents which have a text in English language and contain these terms. Patents that have been published, for example, in German or French, are not included in the results.

Now click on the "F" icon located above the text field. The list will now be extended to all family members. The result list will subsequently also contain patents without a text or with a non-English text, which would not have led to a hit in the text search.

The amount of found families will not change. The amount of found patents/SIP-patents will increase.

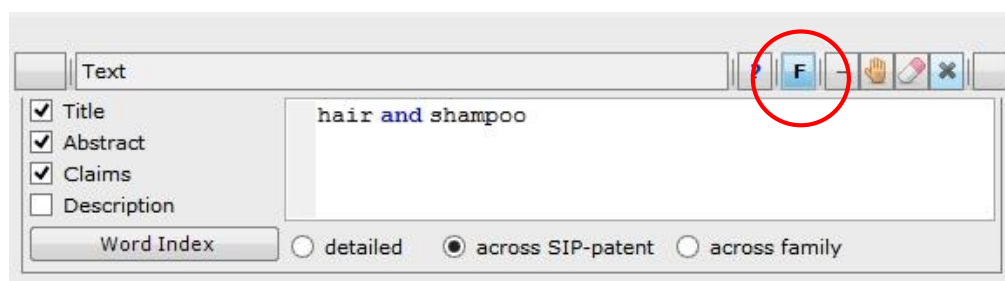The result list can also be extended to families accordingly when using the search across family.

Image 3: Search in the Invention Navigator Premium version with activated family extension

1.9   The history

From Invention Navigator Portable version 2.3.0, in addition to the display of SIP-patents, the display of the amount of documents and families is available (INPADOC family).

The history displays the found amount of SIP-patents. Additionally the amount of documents and families is displayed. This information is saved in the history. If the history is opened after a database update, the history displays the amount of hits at the time of the execution of the research. The number of documents and families derived thereof may change after the update.

If, for example, an EP-application was found prior to the database update it is counted as a SIP-patent and saved in the result list. If a database update is performed, which contains an EP granted patent belonging to the previously found EP-application, the amount of SIP-patents will not change. However, as 2 documents belong to the SIP-patent, the number of documents will increase. As families are corrected by separating false families or merging individual families, the amount of families can also change. This, however, does not occur frequently.

Summary: In the history, the amount of documents and families does not change. In the title list the number of SIP-patents remains the same but the number of documents and families will be newly calculated accordingly.

Image 4: History with the same search prior and after the database update

As you can see in image 4, 2608 documents in 1966 SIP-patents and 1400 families were found prior to the database update.

Image 5: Results in the title list after opening the result R152

If you now open the result list R152 the result shown in image 5 will be displayed.

The number of SIP-patents remains the same. The number of families as well. The number of documents increases by one hit to 2127 documents (see The history)

This means that during the update only one document was added of which one SIP-patent was already included in the result list prior to the update.

These differences don't have to be considered when carrying out researches. Combining searches and results always happens on application level. The amount of patents and families is only a further information to enable comparisons with document based or family based databases.

We hope that we have answered the most important questions about the text search with this manual. Our service team will be happy to answer your questions.

Your SIP-Team

Service@s4ip.de