Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2022 doi:10.20944/preprints202205.0054.v1

Patent prior art searches: basic principles and strategies

Massimo Barbieri, Politecnico di Milano – piazza Leonardo da Vinci, 32 – 20133 Milano, ITALY (e-mail: massimo.barbieri@polimi.it; OrcID: https://orcid.org/0000-0002-7409-5861)

Abstract

From a strategic and an economic point of view, patent prior art searches are very useful for several reasons such as avoiding the infringement of third parties' rights, reinventing something already new in the state of—the-art, speeding up the prosecution of a patent application and getting a patent granted with a lower associated cost.

A prior art search is carried out in patent databases (free of charge or fee-based) by means of keywords or classification symbols, indexing codes (IPC – *International Patent Classification* or CPC – *Cooperative Patent Classification*) or a combination of these, using Boolean or proximity operators.

The classification codes are independent from the language used and can increase the number of relevant documents retrieved.

Patent databases are all incomplete and therefore a search in multiple sources is suggested.

Introduction

Why should startups and innovative companies deal with patents, when their main competitive advantage is provided by the products and/or services they offer on the market?

The reasons are essentially two: to maintain a position of pre-eminence *vis-à-vis* direct (or future) competitors, preventing them from profiting from what has been validly protected by an industrial property right (IPR) and to avoid re-inventing something that has been already discovered or even already patented by third parties.

Creating a product/device that depends, even partially, on third parties' rights, could lead the company to have to defend itself against an infringement or to have to negotiate a license agreement in order to proceed with the product or device to market.

Dealing with patents does not only incorporate creating a portfolio of intangible assets, which has an economic value as a defensive or offensive tool, but also includes knowing thoroughly the state-of-the-art of the product sector in which the company carries out its business.

This knowledge, even if initially perceived as an additional outlay for R&D activities, brings several benefits, such as:

- Saving on R&D expenses
- The possibility of carrying out the design around of a product before entering the market
- Speed up the prosecution of a patent application (fewer office actions means lower costs for obtaining the grant of the patent)
- Limit the scope of protection (or granting) of competitors' patents by means of opposition (administrative) or court actions
- Check the freedom to operate (FTO) on its products/services [1]

Types of patent searches

Essentially five types of prior art searches can be identified [2]:

- **Informative** (or state-of-the-art): to plan R&D projects, analyze technological trends or monitor competitors/potential licensors
- **Patentability**: to ensure that an invention meets the law requirements (novelty and inventive step)
- **Validity**: to make sure that a granted patent is actually valid (a subset consists of opposition searches)
- **Freedom to Operate** (FTO): to establish the marketability of a product in one or more countries (and therefore verify that it does not depend on third parties' rights)
- **Legal status**: to prove the status of a patent (filed, granted, opposed, abandoned, withdrawn)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2022 doi:10.20944/preprints202205.0054.v1

A patentability search has a narrower scope than a state-of-the-art search and does not need to be complete. It has no limitations regarding the type of documents and jurisdiction, but the publication date of the prior documents retrieved must be earlier than the filing date of the patent application under examination.
A validity search is usually based on the retrieval of documents (not found by an examiner) relevant to undermine the inventiveness and novelty requirements.
An FTO search is focused on the claims of granted and active patents (in one or more countries), patent applications still under examination and PCT applications, which could be followed by the filing of a national or regional phase.
In this case, it is essential that the recall is as complete as possible, and the search will focus on the essential technical features of the product under examination in the claims (using precise keywords with the support of proximity operators). [3]
The legal status search is necessary to know if an application or a granted patent are still active or have been abandoned or revoked or are now available in the public domain because the terms of the patent (20 years from the date of filing of the application) have expired.

Patent databases

Patent information is made available by national (or regional) patent offices or by independent producers through databases available on the Internet.
The databases can be divided into bibliographic or full text and are both free of charge (e.g., Espacenet, Patentscope) and fee-based (see Table #1).

Patent database	Provider	Example
Free of charge	National or regional Patent Offices; independent companies	Espacenet, Google Patents
Fee based	Independent companies	Derwent Innovation, Orbit Intelligence

Table #1 – Types of patent databases

Professional databases generally offer wider and more accurate searches and data analysis.
To choose the right database, you need to know the time period each one covers, as well as the number and type of documents contained in each database.
A collection of reports on patent databases and their features is available on a platform, called WIPO INSPIRE [4].
Each patent database is incomplete and therefore a complete search is hard to achieve. To be effective, very often a search requires the use of multiple databases. [5]

Classification codes

The classification systems have been designed to organize patents according to the technical sector of the claimed invention and to simplify their retrieval.
All patent documents are assigned at least one classification symbol by the examiners in charge of assessing the patentability requirements of an application.
The two most used systems are: IPC (International Patent Classification) [6] and CPC (Cooperative Patent Classification) [7]. Both are hierarchical systems, with eight/nine sections respectively and several classes, subclasses, groups, and subgroups.
The CPC is an advanced version of the IPC, having the same structure but with a greater number of subdivisions (see Table #2).
The CPC Scheme is divided into three parts:

- The “main trunk” symbols
- The indexing codes or “2000-series”
- The “Y-section”.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2022 doi:10.20944/preprints202205.0054.v1

Classification system	No. of subgroups
IPC	80,000
CPC	250,000

Table #2 – Main classification systems

The classification symbols are assigned according to the function or application described in the claims (e.g., F61K: valves; A61F2/24: heart valves) and refer to the inventive information.
The indexing codes are used to classify aspects not covered by the classification scheme and to classify additional information only (which is useful when carrying out a search).

Search strategies

The scope of prior art searches is to find documents that claim similar (or identical) concepts to the essential technical characteristics of the invention under examination.
A search can be performed by using keywords (which is the most intuitive but subjective method), classification codes, citations (forward and/or backward) or a combination of them.
The advantage of using classification symbols and citations is their independence from the language used and this allows you to find any documents not translated into English as well as those in which the search terms were not present.
Furthermore, keywords can have different meanings and often those found in claims are used to generalize a concept (for example, screw vs. means of fastening).
There are essentially two search strategies.
The first one involves combining all the keywords (and synonyms) and classification codes corresponding to each technical feature to be searched (but assuming that the planned set of terms and codes is complete).
Imagine a planned a list of keywords (KWs) and symbols (S_n) for two technical features (F_n).
They have to be combined in the following way:

$$R_1 = KW_1F_1 \text{ OR } KW_2F_1 \text{ OR } S_1F_1 \text{ OR } S_2F_1$$
$$R_2 = KW_1F_2 \text{ OR } KW_2F_2 \text{ OR } S_1F_2$$
$$R = R_1 \text{ AND } R_2$$

The second strategy is focused on an initial quick search using precise keywords to obtain a limited number of documents to be analyzed.
If one or more relevant patents are identified in the list, the search can be repeated using the symbols with which those documents have been classified.
The fields in which it is advisable to search are title, abstract and claims (the full text could provide an output which includes too many documents, which are not all relevant).
All databases offer the possibility to use the Boolean operators (AND OR NOT) and sometimes also the proximity operators.
Some useful tips on how to set up a search [8]:

1. Avoid running search phrases that are too long and complex and which combine the various technical terms into a single solution
2. Aim to obtain a number of results not exceeding 30/40 and subsequently expand the search
3. Do not use proximity operators in the title and the abstract
4. Search for citations on a relevant document

Search example #1 – 3D bioprinting

The fist example refers to 3D bioprinting, which is a *collection of additive manufacturing (AM) technologies, whose aim is to fabricate parts imitating real tissue and organ functionalities by combining both living and non-living materials in a specific three-dimensional (3D) spatial organization structure.* [9]
The first step is to check which keywords are most commonly used by searching the scientific literature (e.g., through the Scopus database).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2022 doi:10.20944/preprints202205.0054.v1

From 2013 to today there has been an exponential growth in the number of articles on 3D bioprinting (see Figure #1), and these can provide ideas both for the execution of the patent search and for the retrieval of documents hindering the patentability of an invention (novelty is an absolute requirement and therefore does not depend on the type of document).

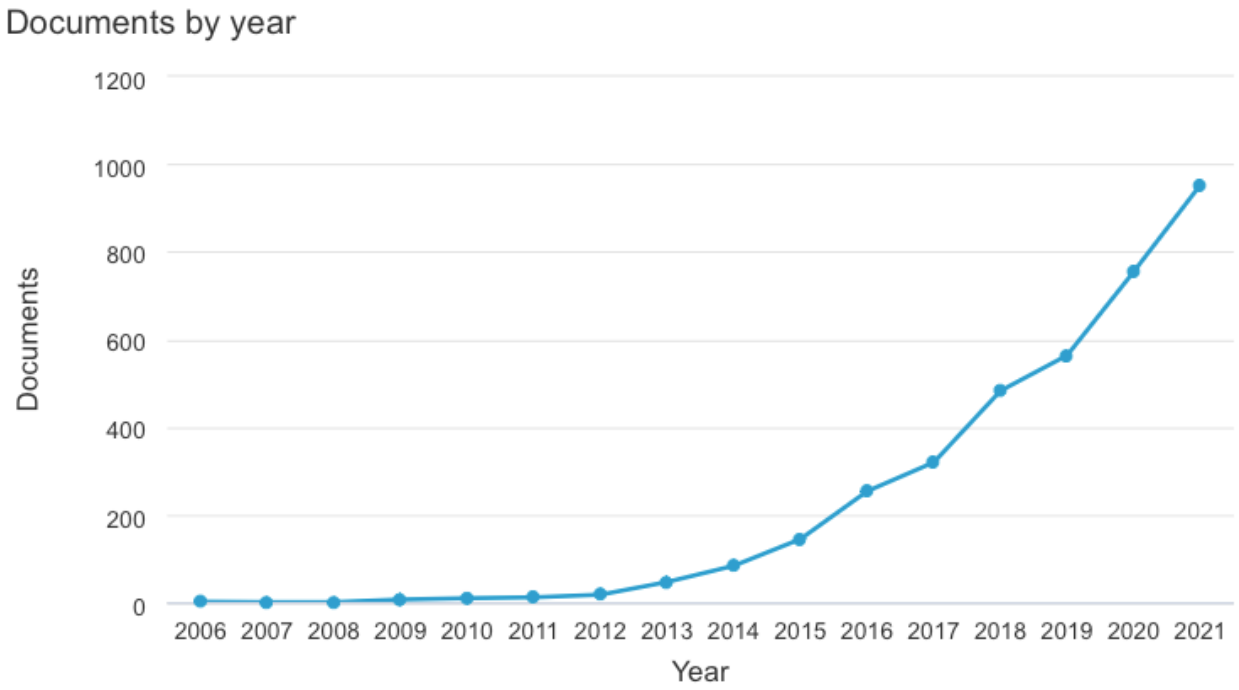


Figure #1 - Number of scientific papers on 3D bioprinting [source: Scopus, Accessed on 3/16/2022

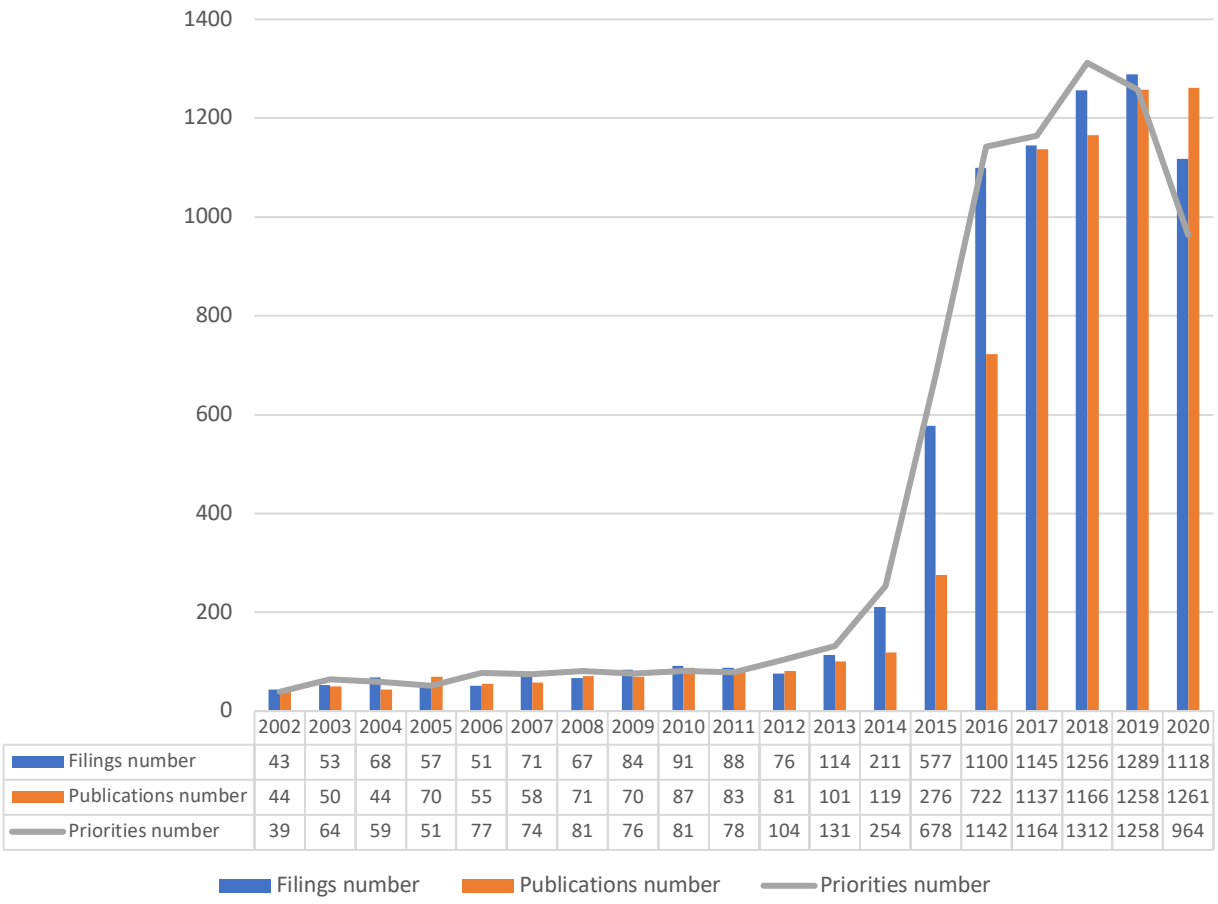
To find the classification codes, you can carry out a preliminary search on Espacenet with one (or more) precise terms, and by analyzing the data you can check the symbols that occur more frequently.
Once the relevant keywords and classification codes have been identified, a search is carried out on a professional database, for example, Orbit intelligence (https://www.orbit.com): the various search queries are listed in Table #3.

	No. of results	Search queries
1	7.927	(((A61L-027+ OR A61F-002+ OR A61L2430/00 OR C12N-005+ OR C12M-003+ OR C08L OR C09D-011+) AND (B33Y+ OR B29C-064+)))/IPC/CPC
2	993	(BIOPRINT+ OR BIOINK? OR ORGAN_ON_A_CHIP OR BIOFABRICATION OR "3D BIOPRINTING" OR "3D BIOPRINTER")/TI/AB/CLMS/ICLM
3	8.490	1 OR 2

Table #3– List of search queries used on Orbit database (search carried out on 3/16/2022)

What information can be obtained from the state-of-the-art search?
Numerical data can be obtained, for example, the total number of patent families (8,490), the percentage of applications that are no longer active (22%), and among those activities, a distinction can be made between granted patents (3,563) and applications still pending (3,055).
Furthermore, it is possible to analyze the filing trends (see Graph #1 - since 2014 there has been an exponential growth in the number of applications) and to know in which countries the most patents are filed (see Graph #2 - China and the United States are the main countries in the 3D bioprinting sector).

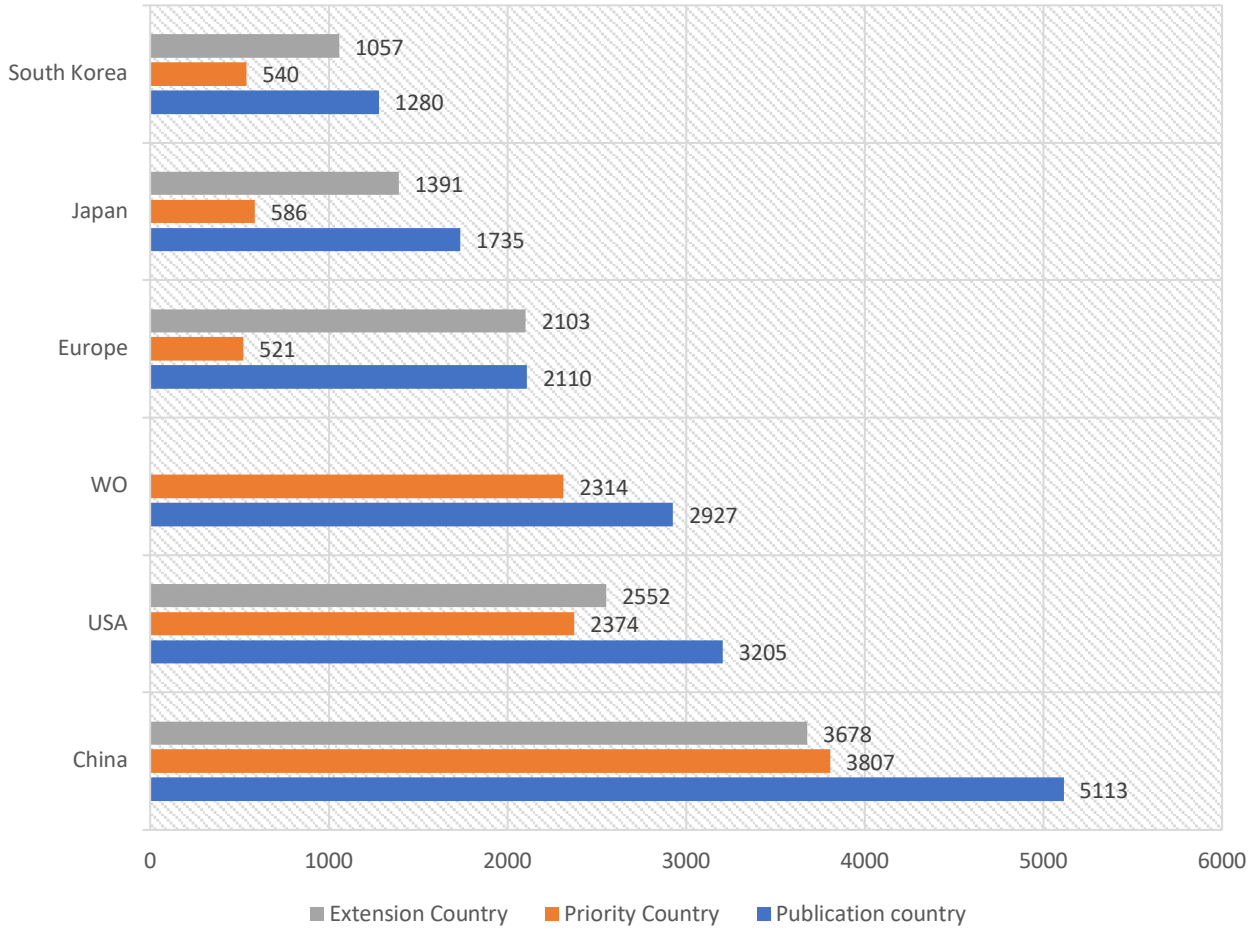
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2022 doi:10.20944/preprints202205.0054.v1



	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Filings number	43	53	68	57	51	71	67	84	91	88	76	114	211	577	1100	1145	1256	1289	1118
Publications number	44	50	44	70	55	58	71	70	87	83	81	101	119	276	722	1137	1166	1258	1261
Priorities number	39	64	59	51	77	74	81	76	81	78	104	131	254	678	1142	1164	1312	1258	964

Graph #1 – Filing, priority, and publication trends

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2022 doi:10.20944/preprints202205.0054.v1



Graph #2 – Filing trends per country

By filtering the results with the filing date and then focusing on the most recent applications, emerging developments can be studied.

Search example #2 – Lignin in tires

A second example refers to lignin used in tires as a filler instead of silica.

The first step is to combine the keyword "lignin" and the classification codes related to tires (B60C and B29D30). To improve search results, the term "lignin" can be replaced by classification symbols (C07G 1/00 – "Lignin; lignin derivatives"; C08H 6/00 – "Macromolecular compounds derived from lignin"; C08H 7/00 – "Lignin; modified lignin; high molecular weight products derived therefrom"; C08L 97/005 [CPC only] – "Composition of compounds derived from lignin").

The IPC and CPC codes used in the search query are listed in Table #4.

Classification system (version)	Classification code	Definition
IPC (2022-01)	C07G 1/00	Low molecular weight derivatives of lignin
IPC (2022.01)	C08H 7/00	Lignin; modified lignin; high molecular weight products derived therefrom
IPC (2010.01)	C08H 6/00	Macromolecular compounds derived from lignin
IPC (2006.01)	C08H 5/02	Other macromolecular compounds derived from lignin
CPC	C07G 1/00	Lignin; lignin derivatives
CPC	C08H 6/00	Macromolecular compounds derived from lignin
CPC	C08L 97/005	Composition of compounds derived from lignin

CPC	C09D 197/005	Coating composition comprising lignin-containing materials

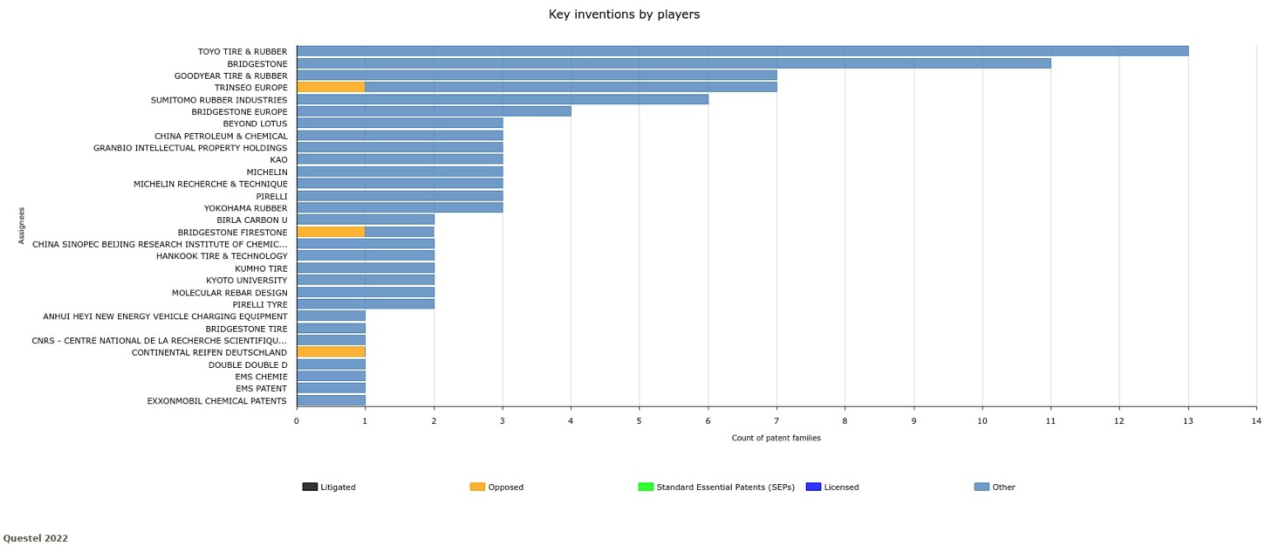
Table #4 - List of IPC and CPC classification symbols

The entire search strategy used in Orbit Intelligence is reported in Table #5

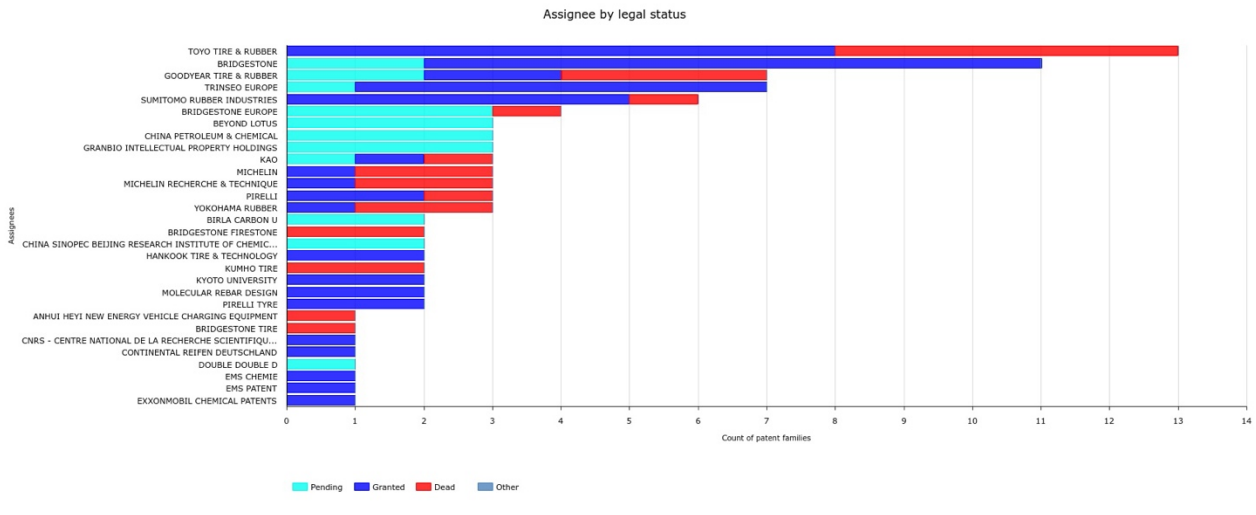
	No. of results	Search queries
1	116	(LIGNIN)/TI/AB/CLMS/ICLM AND (B60C+ OR B29D-030+)/IPC/CPC
2	196096	(B60C+ OR B29D-030+)/IPC/CPC
3	4619	(C07G-001/00 OR C08H-007/00 OR C08H-006/00 OR C08H-005/02 C08L-097/005 OR C09D-197/005)/IPC/CPC
4	13	2 AND 3
5	118	1 OR 4

Table #5 - List of search queries used on Orbit database (accessed May 3rd, 2022)

With Orbit Intelligence it is possible to view the top applicants in this technical field (see Graphs #3, #4 and #5).

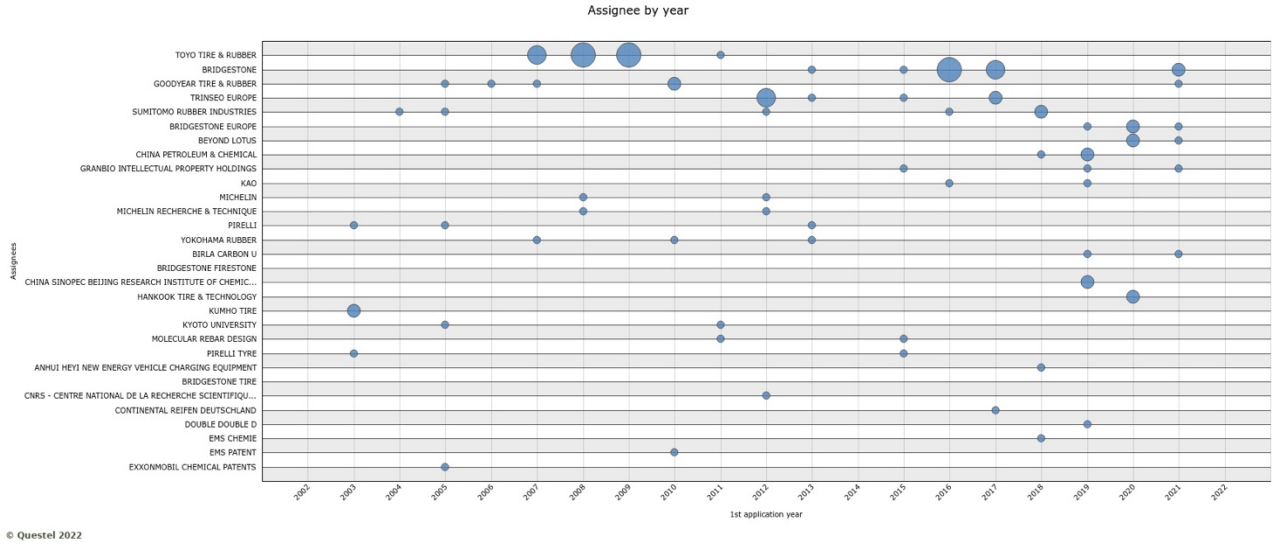


Graph #3 - Key inventions by top applicants



Graph #4 - Top assignee by legal status

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 5 May 2022 doi:10.20944/preprints202205.0054.v1



Graph #5 - Top Assignee by year

Conclusions

Prior art searches are undoubtedly useful for those involved in creating innovative products and services (entrepreneurs, startuppers, researchers) but they do require some knowledge of basic functionality in order to be properly used. The use of classification codes, for example, is what distinguishes patent searches from those commonly adopted on the internet or on databases for scientific literature. Understanding how a classification scheme is structured and how examiners attribute these codes to patent applications is of paramount importance. One limitation when carrying out prior art searches is that access to professional databases is a subscription service. Using paid service allows you to find and analyze data more effectively than just using free ones. Another barrier concerns having knowledge of the technical field of the invention, the comprehension of which is essential to obtaining good results.

Conflicts of interest: The author declares no conflict of interest.

References

[1] A. Zito, Il Diritto Industriale (2022) 1, 17.
[2] M. Barbieri, Le ricerche brevettuali, Sprint Soluzioni Editoriali (2020), p. 2-3
[3] P. Foglia, World Pat. Inf. (2007) 29, 1, p. 33-53
[4] WIPO INSPIRE: https://inspire.wipo.int/wipo-inspire
[5] J. Ohms, World Pat. Inf. (2021) 66, 102055
[6] WIPO IPC: https://ipcpub.wipo.int/
[7] CPC: https://www.cooperativepatentclassification.org/home
[8] E. Nijhof, World Pat. Inf. (2007) 29, 1, p. 20-25
[9] S. Santoni et al., Bio-Design and Manufacturing (2022) 5, p. 14-42