



CEPIUG 10th Year Anniversary Conference



POLITECNICO
MILANO 1863

Pharma & Biotech Patent Searching in Public Scientific Databases

LUCA FALCIOLA
IP & CI MANAGER
(DROIA ONCOLOGY VENTURES, BELGIUM)

OVERVIEW

INTRODUCTION

PUBMED & PATENT-CITING SCIENTIFIC LITERATURE

SEARCHING PATENT-CITED SCIENTIFIC LITERATURE

COMPOUND CODING ACROSS PATENT & SCIENTIFIC LITERATURE

FINAL OBSERVATIONS

Introduction:

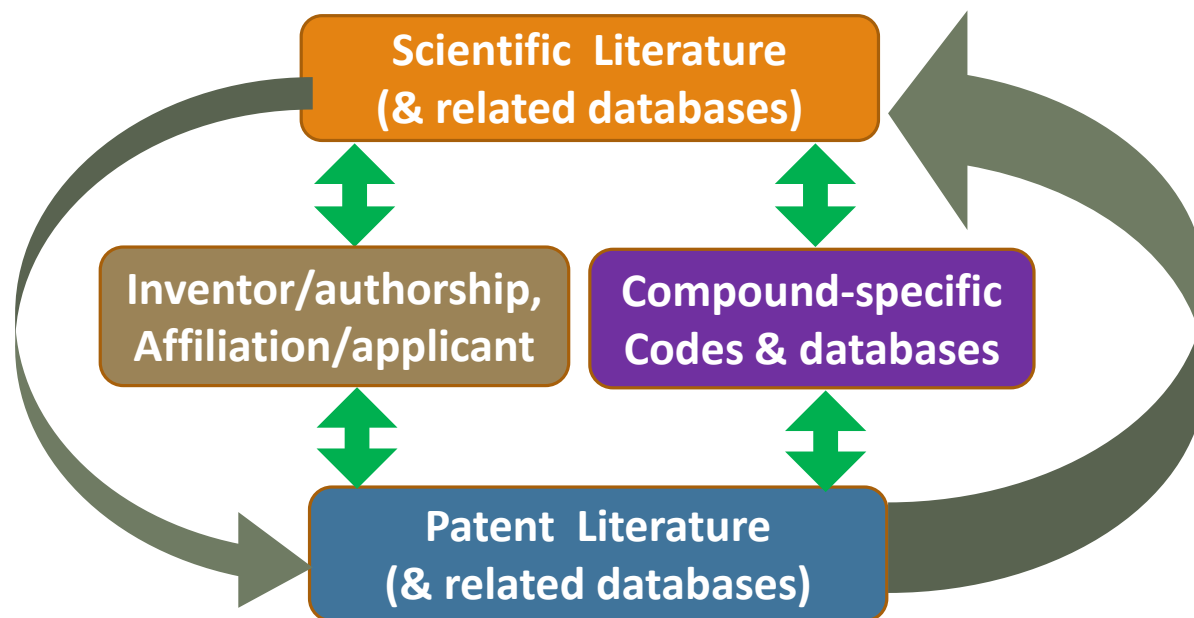
(Broken) Links Between Patent & Scientific Literature

- A complete, patent-relevant search in pharma & biotech requires checking the content of scientific literature beyond patent databases, looking for:
 - ✓ Text-/table-based based disclosures
 - ✓ Details over biological sequences, chemical structures, technologies
 - ✓ Features identified in graphs, photographs, raw datathat may be absent or hard to find in patent literature
- Non-patent scientific databases & literature rarely refer to patent-originated or patent-related information, and often in a incomplete and/or inconsistent manner due to
 - ✓ Poor understanding of patent system
 - ✓ Qualitatively/quantitatively Incomplete coverage of patent informationLimiting the possibilities to identify and consolidate the best of the two “worlds”

Introduction:

(In)directly Cross-citing Patent & Scientific Literature

- The citation flow is uneven between patent and scientific literature
 - ✓ Still a limited number of articles cite patent documentation at all or properly
 - ✓ Scope/format of scientific literature cited in patent documentation is uneven
- The information that is actually "shared" between these two types of disclosure may be based on identity of names of people, organization, and/or compounds but other means may facilitate this task or expanding the scope

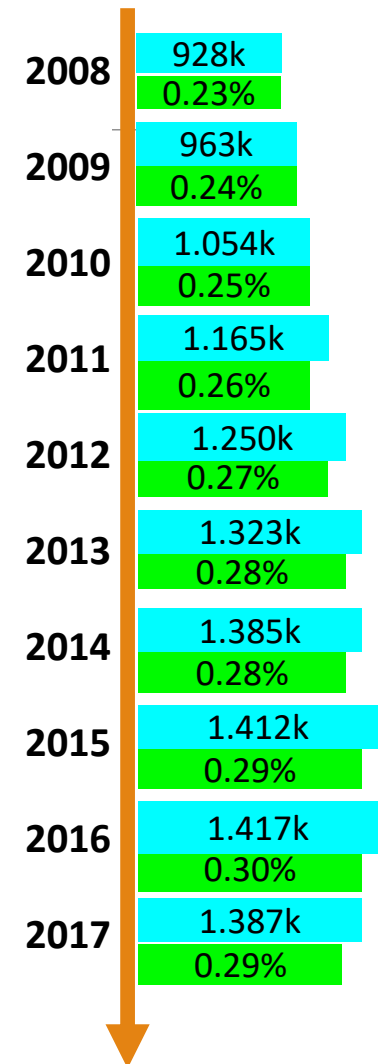


Pubmed & Patent-Citing Scientific Literature: Introduction

- Pubmed is a major entry point for searching scientific literature in life sciences, and beyond, that present some major advantages, eg:
 - ✓ High number of covered journals & books, with high frequency of update
 - ✓ Complex search & field features, MeSH indexing (Medical Subject Headings)
 - ✓ Links to full-text, similar articles, & records in other NIH databases
- At the same time, some general limitations need to be acknowledged:
 - ✓ Dependency from provider's formats & policy for adding new entries
 - ✓ Inconsistent/unclear own policy on when/how entries are updated & indexed
 - ✓ No links/search features to citations from/to a given entry ("similar" feature)
- When dealing with patent-related topics, other limitations appear in Pubmed:
 - ✓ No distinction between entries on economical/social/political matters about patent system and those about technical/legal content of patent documents
 - ✓ Confusion over the actual use and meaning of "patent" across entries

Pubmed & Patent-Citing Scientific Literature: Patent-related content & Pubmed Searching

- A quick search in Pubmed shows that, over the last 10 years:
 - ✓ Number of Pubmed entries has regularly increased
 - ✓ Percentage of entries including *patent** over this period is 0.27% but increasing somehow over the years
- By browsing these entries, a quite large variety of patent-related topics appears covered in such articles, but a substantive fraction of them appears having nothing to do with any patent matter (quite confusing/discouraging)
- Truly/falsefully patent information in Pubmed can be more directly identified by “deconstructing” how Pubmed actually present and use such entries
 - ✓ Four “search levels” can be defined
 - ✓ The outcome of those alternative search approaches can be then combined with technical keywords



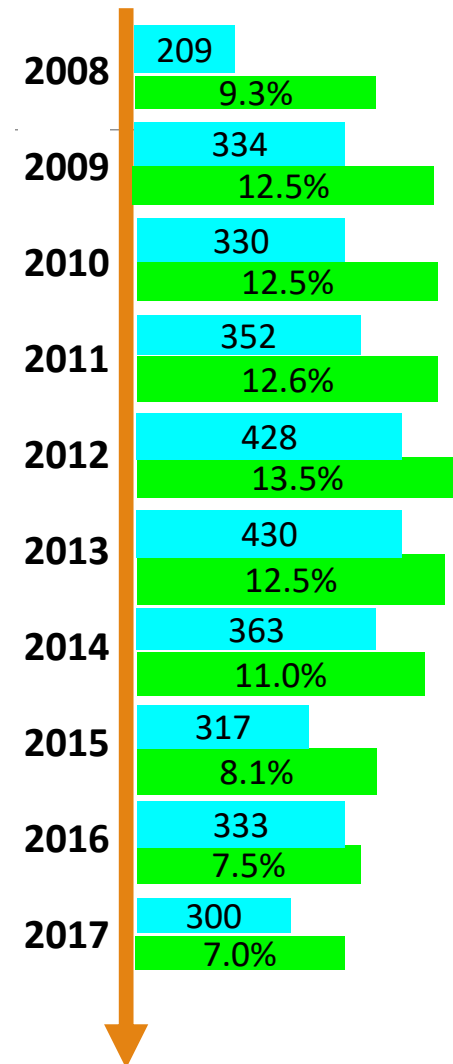
Pubmed & Patent-Citing Scientific Literature : Journals Covering Patent Literature (Level 1; I/II)

- Pubmed indexes articles a series of journals explicitly covering patent matters information is available in a variety of formats:
 - ✓ Expert opinion on therapeutic patents
 - ✓ Pharmaceutical patent analyst
 - ✓ “Recent Patents” series (14 distinct titles among the whole series, but some of them are discontinued)
 - ✓ Some, but limited, entries from legal journals
 - ✓ Other potentially relevant journals (such as World Patent Inf., J. Tech. Transf., Technovation, Scientometrics) are **not** indexed in Pubmed

- General observations:
 - ✓ Possibly the level providing the highest percentage of patent-related technical content since authors & audience are mainly people working regularly on patent matters in academia or companies
 - ✓ High article / subscription cost, inversely proportional to Impact Factor

Pubmed & Patent-Citing Scientific Literature: Journals Covering Patent Literature (Level 1; II/II)

- Pubmed search in these journals requires considering:
 - ✓ The [Journal] field need to be fully written according to Pubmed format (no truncation or limited search criteria), that is, the initial search using *patent** does **not** find their articles not including *patent** elsewhere in the entry
 - ✓ Incomplete year coverage and irregular update (eg coverage of Exp. Opin. Ther. Pat. starts from 2004,vol.14)
- However,
 - ✓ These journals provide a **number of entries** that is uneven and does not contribute to most recent increase but still substantial in terms of **percentage over *Patent** entries**
 - ✓ If a topic of a search is fully new for a searcher, and an article in these journals covers such a topic, maybe worth getting it at least as a starting point for keywords, etc.
 - ✓ Search can be pursued also at publisher's website



Pubmed & Patent-Citing Scientific Literature: Articles with *Patent** in Original Title/Abst. (Level 2; I/II)

- The “*patent**” search can be limited to Pubmed entries in which authors made use of this wording in title and/or abstract (*as patents, patenting, patentee(s), patented, “patent landscape”, “drug patent(s)”, “patent expiration”, “patent pending”, “patent review(s)”, “patent protected”, “patent application(s)”, “patent literature”, “patent survey”, “patent search”, “patent protection”, patentab*, “patent value”, “us patent”, “patent office”, “patent citations”, “patent citation”, “patent system”, etc.*) but there are two main source of false hits:

- ✓ The typo pati*e*nt / pati*e*nts is not corrected, with the result of selecting many articles of clinical research and summaries of medical case studies
- ✓ The regular use of “patent” as adjective in medical and/or older English indicating “*unobstructed, obvious, readily visible, exposed*” eg

Patent medicine	Patent foramen ovale	Patent fistula	Patently
Patent anastomosis	Patent ductus arteriosus	Patent shunt	Patent urachus

- If alternative search approaches may allow finding entries that present the desired “patent” anyway, the recall/precision of Pubmed search in title/abstract is improved by using more appropriate phrases or variants

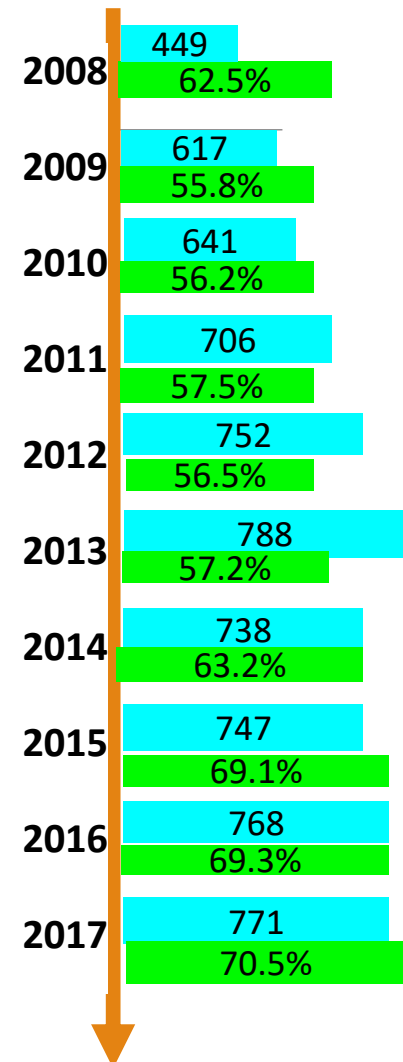
Pubmed & Patent-Citing Scientific Literature : Articles with *Patent** in Original Title/Abst.(Level 2;II/II)

- A complex but still not fully “clean” search strategy can be

(patenta*[Title/Abstract] OR patente*[Title/Abstract] OR patenti*[Title/Abstract] OR patents[Title/Abstract] OR "new patent"[Title/Abstract] OR patents[Title/Abstract] OR patenting[Title/Abstract] OR patente*[Title/Abstract] OR "patent landscape"[Title/Abstract] OR "drug patent"[Title/Abstract] OR "drug patents"[Title/Abstract] OR "patent expiration"[Title/Abstract] OR "patent pending"[Title/Abstract] OR "patent review"[Title/Abstract] OR "patent reviews"[Title/Abstract] OR "patent protected" [Title/Abstract] OR "patent application"[Title/Abstract] OR "patent applications"[Title/Abstract] OR "patent literature"[Title/Abstract] OR "patent survey"[Title/Abstract] OR "patent search"[Title/Abstract] OR "patent protection"[Title/Abstract] OR "patent value "[Title/Abstract] OR "us patent"[Title/Abstract] OR "patent office"[Title/Abstract] OR "patent citations"[Title/Abstract] OR "patent citation"[Title/Abstract] OR "patent system"[Title/Abstract])

- This search provides with a **total number of entries** that has grown and now is stabilized but **percentage of those not present in *Patents* journals** has continued increasing

- ✓ Contributing at least in part to recent growth of patent information-related technical content in Pubmed
- ✓ Showing the higher interest in publishing about patent matters and/or referring to patent documentation in generally non-patent-oriented journals



Pubmed & Patent-Citing Scientific Literature: Articles with Patent-Related MeSH (Level 3)

- MeSH indexing provides:
 - ✓ A single “patents as topic”[MeSH Terms]
 - ✓ Some “patent”[MeSH Terms] related to the medical use of adjective (eg “patent medicines”, “patent ductus”)

- However:
 - ✓ “patents as topic”[MeSH Terms] covers less than half of the hits found using journal-based/Level 1 or title&abstract-based/Level 2 strategies
 - ✓ “patents as topic”[MeSH Terms]-only hits increase the total number of patent-related hits by 10%-15% each year
 - ✓ A large majority of such “patents as topic”[MeSH Terms]-only hits refers to articles covering economical, social, and/or legal topics rather than technical contents of patent documents

- The “patents as topic”[MeSH Terms] search is worth to be used but always completed with technically focused, patent term- based search strategies

Pubmed & Patent-Citing Scientific Literature: Articles with Patent Originator Info (Level 4; I/IV)

- Once that a list of companies and/or authors is available for comparing patent publications and scientific publications, it is a common practice to use these names in Pubmed but worth reminding some search features:
 - ✓ Both [Author] and [Affiliation] fields can be searched using the truncation operator * and all other usual operators
 - ✓ However [Affiliation] field is neither standardized, nor always provided by publisher in a complete format, but more importantly, Pubmed tutorial clearly indicates “before 2014, only the **first** author's affiliation is included”
- This limitation is quite important, especially with the growing number of authors and entities involved in each publications but there is a further, subtler manner to associate Pubmed entries to (un)published patent documents

Pubmed & Patent-Citing Scientific Literature: Articles with Patent Originator Info (Level 4; II/IV)

- In recent years, some publishers and journals (eg PLOsone, PNAS) have strengthened obligations of disclosing any commitment/relationship linking an author with a company having any potential interest in the content of the article
- Pubmed has followed up by introducing a specific search field [Conflict of Interest Statements] **since 2016** that contains a free text including:
 - ✓ Authors and/or their employer being inventor or applicant in (un)published, (un)related patent application
 - ✓ Grants, shares, fees, involvement as founder/consultant/officer, or other financial interests that an author declares with respect to a company
 - ✓ Name of companies somehow involved in the data generated for the article and/or in other activities of an author

That is even retroactively added for entries prior to 2016 by some journals

- A search focused on this field, in association to technical criteria and/or specific name of authors/affiliation, may provide some unexpected insights

Pubmed & Patent-Citing Literature :

Articles with Patent & Originator Info (Level 4; III/IV)

- Some examples of relevant text on [Conflict of Interest Statements]:

*ABC and DE are also **co-inventors on patent applications** that include aspects of device that was used in the clinical trial*

*LN and RS are **co-inventors of an Australian provisional patent application** No. 2017nnnnnn [title] filed by Univ. NNNN that includes the present work*

*SH **owns the following patents**: US5nnnnnn, US7nnnnnn, US9nnnnnn*

*GA and TB have **filed patents in the field** of anti-CDnn therapy*

*AS and SR-J are shareholders of NNN Inc., and both are **inventors of patents owned by** NNN Inc.*

*The **technology has been patented** by Hosp. ABC and an **exclusive license** signed with NNN Ther.. Dr. AB has Founders shares in NNN Ther.*

- There are also quite a number of “negative”, irrelevant statements

AS is co-applicant for NIH grant No. nnnnnn

There are no patents, products in development or marketed products to declare

Pubmed & Patent-Citing Literature :

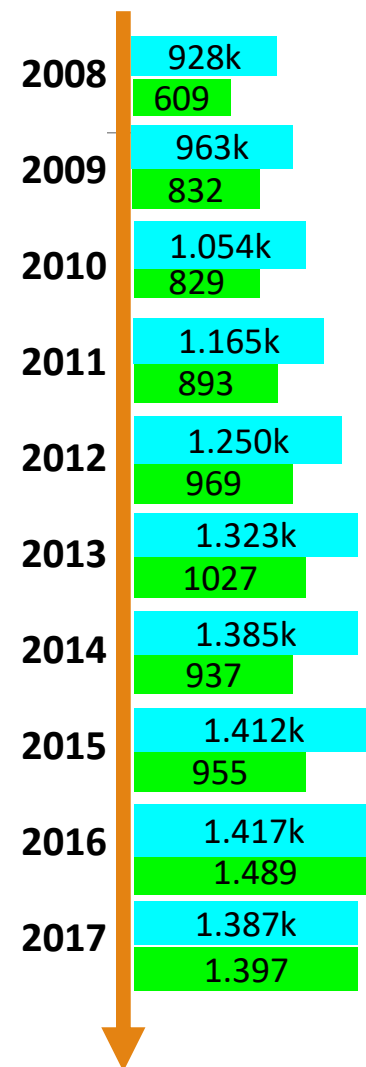
Articles with Patent & Originator Info (Level 4; IV/IV)

- The [Conflict of Interest Statements] field search by using criteria such as *invent**, *coinvent**, *coapplicant**, *applicant**, *patent**, *provisional** may be used:
 - ✓ To disambiguate the general *patent** search
 - ✓ To expand what other search levels (or even other databases) may hardly provide in a clear manner
- Indeed:
 - ✓ More than 1200 Pubmed entries in 2017 and (in Aug. 2018) more than 1100 entries in 2017 have some type of patent-relevant text in this field
 - ✓ Less than 50 of them are found by using any of previous approaches (even though a substantial percentage may present “negative statements”)
 - ✓ By combining *invent**, *coinvent**, *fili**, *file**, *licens**, *licenc**, *coapplicant**, *applicant**, with *patent**, *provisional**, approx. 50% of initial hits not found with level 1-3 searches are selected

Pubmed & Patent-Citing Literature : Take-home messages

- Pubmed may be searched for identifying patent-relevant and/or patent-originated information in scientific literature but:
 - ✓ Using the appropriate approach(es)
 - ✓ When the scope can justify spending some time and/or some names and other keywords are available for a search
- The search strategy below covering search level 1, 2, 3, or 4 is still not fully “clean” but provides a more limited, but more relevant, list of hits in Pubmed that is clearly growing over the years when compared to the total content of Pubmed

"Pharmaceutical patent analyst"[Journal] OR "Recent patents on anti-cancer drug discovery"[Journal] OR "Recent patents on anti-infective drug discovery"[Journal] OR "Recent patents on biotechnology"[Journal] OR "Recent patents on inflammation & allergy drug discovery"[Journal] OR "Recent patents on drug delivery & formulation"[Journal] OR "Recent patents on nanotechnology"[Journal] OR "Recent patents on cardiovascular drug discovery"[Journal] OR "Recent patents on food, nutrition & agriculture"[Journal] OR "Recent patents on CNS drug discovery"[Journal] OR "Recent patents on DNA & gene sequences"[Journal] OR "Recent patents on endocrine, metabolic & immune drug discovery"[Journal] OR "Recent patents on nanomedicine"[Journal] OR "Recent patents on biomarkers"[Journal] OR "Recent patents on biomedical engineering"[Journal] OR "Recent patents on regenerative medicine"[Journal] OR "Recent patents on space technology"[Journal] OR "Expert opinion on therapeutic patents"[Journal] OR ((patenta*[Title/Abstract] OR patente*[Title/Abstract] OR patenti*[Title/Abstract] OR patents[Title/Abstract] OR "new patent"[Title/Abstract] OR patents[Title/Abstract] OR patenting[Title/Abstract] OR patente*[Title/Abstract] OR "patent landscape"[Title/Abstract] OR "drug patent"[Title/Abstract] OR "drug patents"[Title/Abstract] OR "patent expiration"[Title/Abstract] OR "patent pending"[Title/Abstract] OR "patent review"[Title/Abstract] OR "patent reviews"[Title/Abstract] OR "patent protected" [Title/Abstract] OR "patent application"[Title/Abstract] OR "patent applications"[Title/Abstract] OR "patent literature"[Title/Abstract] OR "patent survey"[Title/Abstract] OR "patent search"[Title/Abstract] OR "patent protection"[Title/Abstract] OR patents as topic[MeSH Terms] OR (fili*[Conflict of Interest Statements] OR file*[Conflict of Interest Statements] OR invent*[Conflict of Interest Statements] OR coinvent*[Conflict of Interest Statements] OR coapplicant*[Conflict of Interest Statements] OR applicant*[Conflict of Interest Statements] OR licenc*[Conflict of Interest Statements] OR licens*[Conflict of Interest Statements]) AND (patent*[Conflict of Interest Statements] OR provisiona*[Conflict of Interest Statements]))) AND ("2008/01/01"[Date - Publication] : "2018/01/01"[Date - Publication])



Searching Patent-cited Scientific Literature

Introduction

- The citation of scientific literature in patent proceedings is particularly important in pharma & biotech patent proceedings across:
 - ✓ The text of the patent application
 - ✓ In official search report by patent offices
 - ✓ During examination or post-grant proceedings
- Various tools are available for identifying cross-citing scientific literature, less opportunities are available for searching scientific literature according to their use in patent literature
- These searches may be particularly relevant for competitive intelligence and post-grant proceedings but require taking into account the lack of consistency on how available tools identify and link patent and scientific literature
 - ✓ Different format of article citation by different patent attorneys/offices
 - ✓ Coverage limited to the literature in the patent application, in the search report, in the IDS, and any of their combinations



Searching Patent-cited Scientific Literature : Lens.org Contents & Searches

- Lens.org provides a quite unique double database and double search interface covering scientific literature and patent information with links to:
 - ✓ PDF version of patent documents
 - ✓ Article official indexing
 - ✓ Cross-citing frequency & other metadata
- In addition, the PatSeq section provides different access to DNA & Protein cited into patent documentation with links to:
 - ✓ Specific webpage for each sequence in the official sequence listing
 - ✓ Corresponding patent documents
- Lens.org provides users with large datasets, several layers of linked information, and different search approaches, but not clear if/how the content and links are actually validated for update, completeness, and precision, eg:
 - ✓ Incomplete coverage for older documents
 - ✓ IDS only often used for US patents

Searching Patent-cited Scientific Literature: Lens.org Patent Structured Search Interface

- The Structured Search interface for patent literature covers both DNA/protein sequence content (but only as sequence listings) and (non)patent citations

The screenshot displays the Lens.org Patent Structured Search interface. At the top, there are tabs for "Structured Search", "Search Tips", and "Support". Below the tabs, the "Configuration" section includes radio buttons for "Patents" (selected) and "Scholar". The "Predicate" section has radio buttons for "AND" (selected) and "OR". The "Field" section features a dropdown menu set to "Title" and three search criteria boxes, each containing the text "Fidget Spinner". The "Dates" section includes radio buttons for "Published" (selected) and "Filed", along with "From" and "To" date input fields. The "Jurisdictions" section has a dropdown menu set to "Defaults to All". The "Doc Type" section includes a dropdown menu set to "Defaults to All". The "Query Language" section has a dropdown menu set to "English". The "Other options" section includes a dropdown menu set to "Other Options". A blue "New Search" button is located at the bottom left. A yellow box highlights the "Non Patent Citation" and "Citation Id" options in the "Citations" section of the field list on the right.

Structured Search | Search Tips | Support

Configuration:

Patents Scholar

Predicate: AND OR

Field:

"Fidget Spinner" Title + -

"Fidget Spinner" Title + -

"Fidget Spinner" Title + -

Dates: Published Filed

From To

Jurisdictions: Defaults to All

Doc Type: Defaults to All

Query Language: English

Other options: Other Options

New Search

Looking for the old structured search?

general

- Lens Id
- Pub Key
- Kind
- Pub Num
- Pub Year
- Title
- Abstract
- Applicant
- Inventor
- Owner
- Full Text
- Claims

classification

- Classification Ipcr
- Classification Nat
- Classification Cpc

citations

- Cites Patent Pub Key
- Cites Patent Count
- Cited By Patent Count
- Citing Orcid Works
- Non Patent Citation
- Citation Id

family

- Family Of Pub Key
- Family Jurisdiction
- Family Size
- Simple Family Of Pub Key
- Simple Family Jurisdiction
- Simple Family Size

sequence

- Sequence Count
- Sequence Length
- Sequence Type

Searching Patent-cited Scientific Literature: Lens.org Scientific Literature Search Interface

- The Structured Search interface for scientific literature is even more complete but unclear if actually all indexed literature is searchable with all listed criteria

The screenshot displays the Lens.org search interface. On the left is the 'Structured Search' configuration panel, which includes tabs for 'Structured Search', 'Search Tips', and 'Support'. The 'Configuration' section has radio buttons for 'Patents' and 'Scholar' (selected). The 'Predicate' section has radio buttons for 'AND' (selected) and 'OR'. The 'Field' section has a warning icon and two dropdown menus. The 'Flags' section has checkboxes for 'Cited by Patent', 'Abstract', 'Keywords', 'Substance', 'Funding', 'Clinical Trials', 'MeSH Term', 'Affiliation', and 'Field of Study'. A 'New Search' button is at the bottom left.

In the center is a list of search criteria, with several items highlighted in yellow: 'Citation Id', 'Reference Count', and 'Referenced By Patent Count'.

On the right is a list of search results, with several items highlighted in yellow: 'Reference Id', 'referenced_by_patent', and 'Referenced By Patent Lens Id'.

Search Configuration Panel:

- Structured Search | Search Tips | Support
- Configuration:
- Patents Scholar
- Predicate: AND OR
- Field:
- Search Input 1:
- Search Input 2:
- Flags:
 - Cited by Patent Abstract Keywords Substance Funding
 - Clinical Trials MeSH Term Affiliation Field of Study
-

Search Criteria List:

- general
 - Lens Id
 - Citation Id**
 - Volume
 - Issue
 - Start Page
 - End Page
 - Year Published
 - Date Published
 - Title
 - Abstract
 - Keyword
 - Reference Count
 - Referenced By Patent Count**
 - Publication Type
- author
 - Author Last Name
 - Author First Name
 - Author Display Name
 - Author Affiliation Name
- mesh_term
 - MeSH Heading
 - MeSH UI

Search Results List:

- chemical
 - Chemical Substance Name
 - Chemical Registry Number
 - Chemical Mesh Ui
- funding
 - Funding Country
 - Funding Funding Name
 - Funding Funding Id
 - Funding Organisation
- clinical_trial
 - Clinical Trial Registry
 - Clinical Trial Trial Id
- reference
 - Reference Id
 - referenced_by_patent**
 - Referenced By Patent Lens Id**
- source
 - Source Title
 - Source Publisher
 - Source Country
 - Source Issn
 - Source Asjc Subject
- conference
 - Conference Name
 - Conference Instance
 - Conference Year
 - Conference Date
 - Conference Location

Searching Patent-cited Scientific Literature: Lens.org Patent Literature Search Language

- The Lens.org help pages provide a number of examples and details also for constructing long, complex patent searches combining criteria and operators

Index Fields

The following are all of the fields indexed on The Lens.

- > `lens_id` e.g. 186-488-232-022-055
- > `pub_key` e.g. US_2013_0227762_A1
- > `jurisdiction` e.g. US
- > `kind` e.g. A1
- > `pub_num` e.g. 2013/0227762
- > `pub_date` e.g. 20170905 - yyyymmdd
- > `pub_year` e.g. 2018
- > `filing_date` e.g. 20000519
- > `earliest_priority_date` e.g. 20000519
- > `title` e.g. "Fidget Spinner"
- > `abstract` e.g. "Super Conductor"
- > `applicant` e.g. "Smith David"
- > `inventor` e.g. Sally
- > `owner` e.g. "Sony Ltd"
- > `has_full_text` e.g. true
- > `full_text` e.g. robot
- > `claims` e.g. semiconductor
- > `classification_ipcr` e.g. "H01L21/768"
- > `classification_nat` e.g. "221/220" (US classifications)
- > `classification_cpc` e.g. H01L2924/*

- > `cites_patent_pub_key` e.g. US_7128866_B1 - docs citing US_7128866_B1
- > `cites_patent_count` e.g. 5
- > `cited_by_patent_count` e.g. 10
- > `non_patent_citation` e.g. (health OR medicine)
- > `citation_id` e.g. 10.1038/NATURE03090
- > `family_of_pub_key` e.g. US_6408520_B1
- > `family_jurisdiction` e.g. (US OR EP)
- > `family_size` e.g. [4 TO 6]
- > `simple_family_of_pub_key` e.g. US_6408520_B1
- > `simple_family_jurisdiction` e.g. US
- > `simple_family_size` e.g. 3
- > `sequence_count` e.g. [2 TO 3]
- > `sequence_length` e.g. [1 TO 100]
- > `sequence_type` e.g. N - nucleotide, P - peptide

To search for a value in an indexed field you type the name of the field followed by a colon and then the value you wish to search. For example:

```
* title:malaria
* pub_date:[20070101 TO 20070631]
```

When you want to search on multiple fields you can use boolean operators which all must be upper case. For example:

```
* rice AND pesticide
* malaria OR mosquito
* printing NOT inkjet
```

Lucene supports AND, OR, NOT as well as "must" + and "must not" - as Boolean operators

For more information on how boolean logic works, see [this tutorial \(/support/help-resources/other-articles/boolean-logic-tutorial/\)](#).

Other operators available include:

```
* term grouping: Lucene supports using parentheses ( ) to group terms into sub queries e.g. (red AND yellow) OR (blue and green)
```

```
* Field grouping: Lucene supports using parentheses to group multiple clauses to a single field e.g. title:(car OR truck)
```

* * and ? for wildcard searches. Note that wildcard search terms are not stemmed and therefore may not work as expected for searches where "Stemming" (located in "Query Tools") is turned on. This is because when stemming is turned on, search terms are matched against stemmed values in the index. For example the terms valve and valves will be stemmed to valv and both match the stemmed value valv in the index. The term valve* will not be stemmed and therefore won't match valv in the index. [Read more about stemming on wikipedia \(http://en.wikipedia.org/wiki/Stemming\)](#).

* ~ for fuzzy/proximity searches. "foo bar"~4 searches for foo and bar within 4 words from each other. Exact matches are proximity zero and word transpositions (bar foo) are proximity 2.

* TO for range searches

* ^ to boost the relevance of a value in a search, affects the result order eg. car abstract:coke^2

* \ to escape the following special query syntax characters in a search term that are not inside quotes:

```
(+ - && | ! ( ) { } [ ] ^ ~ * ? : \ /)
```

Searching Patent-cited Scientific Literature: Lens.org Scientific Literature Search Language

- The Lens.org help pages are less complete and precise about literature coverage and search criteria and operators

Scholarly Search

23/02/2018

Since 2014, the Lens has begun offering scholarly works that are disclosed in the global patent literature and in partnership with NIH PubMed, USA and Crossref, UK, the Lens team managed to resolve 47% of such non-patent literature (NPL) by open persistent identifiers, such as DOIs, PMIDs, and PMCIDs and match them to NPLs in the patent corpus.

Currently, the Lens hosts and serves scholarly records from PubMed (27,860,556), Crossref (17,889,848), and PubMed Central (4,155,640) with the following brief list of the matched and merged granular metadata:

- * citation identifiers
- * title
- * publication date
- * publication type
- * authors (first and last name, order, affiliation)
- * start end pages, volume, issue
- * journal
- * abstract
- * references (string with identifiers if available)
- * funding/grant information
- * keywords (PubMed only)
- * mesh_term (PubMed only)
- * chemicals (PubMed only)
- * clinical_trial data (PubMed only)
- * citing patents

Searching Patent-cited Scientific Literature: Lens.org Scientific Literature Overview & Metadata

- Examples and data overview clearly show the focus in life science but also on some journals

Scholar Results ▾

21,424 Results for: Arthritis

Applied Filters: Cited by Patent,

Scholarly Works Citing Patents **New**

Only Scholarly works cited by Patent Results

Patent Citations ↑

Angiogenesis in cancer, vascular, rheumatoid and other disease.

Journal: Nature medicine, Nature Publishing Group Issue: 1, Volume: 1

Authors: J Folkman

Published: Jan 1, 1995 Info: Patents Abstract Chemical MeSH Terms Affiliation

497 Patent citations 3,935 Scholarly citations (count from Crossref)

Journal: pmid: 7584949 doi: 10.1038/nm0195-27

Scholar Results

108,326,540 Results for: All Scholarly Works

Scholar Filters: Publication Type = (exclude) Unknown

Scholarly Works

Citing Patents

List

Analysis



OFF



Patent Citations ↑



Basic Local Alignment Search Tool

Stephen F Altschul, Warren Gish, Webb Miller, Eugene W Myers, David J Lipman

Journal of Molecular Biology, Issue: 3, Volume: 215, Pages: 403-410. | Oct 1, 1990

Additional Info: Patents Abstract Funding MeSH Terms Affiliation Field of Study

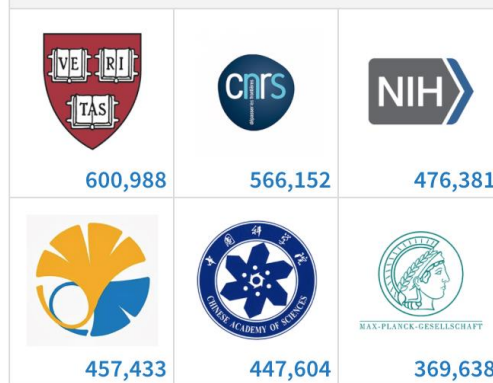
8,286 Patent Citations 28,492 Scholarly Citations

Reference Count: 23

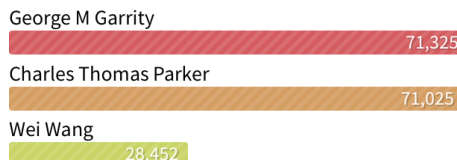
Journal 056-201-471-905-934 10.1016/s0022-2836(05)80360-2

10.1006/jmbi.1990.9999 2055043387 2231712

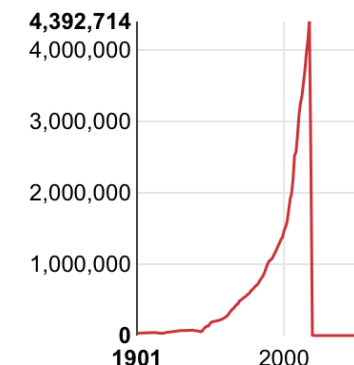
Institution



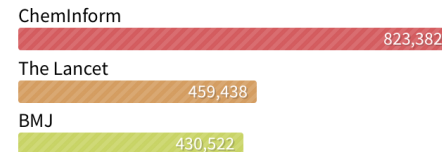
Author



Year Published



Journal



Searching Patent-cited Scientific Literature: Lens.org Scientific Article Presentation

➤ Each article is presented with links to the external sources (with own metrics...)

Article Summary

Osmotic drug delivery : a review of the patent literature

Journal of Controlled Release, Issue: 1, Volume: 35, Pages: 1-21. | Jul 1, 1995

Additional Info: [Patents](#) [Abstract](#) [Affiliation](#) [Field of Study](#)

[206 Patent Citations](#) [181 Scholarly Citations](#) [Reference Count: 1](#)

Abstract

Because patents are hard to read, to analyze, and to judge due to the peculiar style in which they are written, they constitute an under-utilized literature resource. This review analyzes U.S. Patents on osmotic drug delivery through December 1993. The 240 patents cover a period of a little over 20 years. The very first patents dealt with variations of the Rose-Nelson osmotic pump, a rather complex device that has yet to find commercial use. During the 1970s and 1980s, Felix Theeuwes and others at the Alza Corporation made a series of modifications and simplifications to this original concept, leading to the invention and development of the elementary osmotic pump. This device is now used in products with sales that total almost two billion dollars. The development of these devices is an interesting example of how true innovation is sometimes achieved by leaving things out.

Journal

021-506-736-904-214
 10.1016/0168-3659(95)00013-X
 2062374144

This item is not Open Access

Journal Article: Journal of Controlled Release, Issue: 1, Volume: 35, Page: 1-21

Published: Jul 1, 1995

Publication Info: Journal Article

Publisher: Elsevier BV

Purchase PDF
Export ▾

Journal of Controlled Release

Volume 35, Issue 1, July 1995, Pages 1-21

Review

Osmotic drug delivery: a review of the patent literature

Giancarlo Santus ^a, Richard W. Baker ^b

[Show more](#)

[https://doi.org/10.1016/0168-3659\(95\)00013-X](https://doi.org/10.1016/0168-3659(95)00013-X) [Get rights and content](#)

Recommended articles ▲

[A review on controlled porosity osmotic...](#)
 Bulletin of Faculty of Pharmacy, Cairo Uni...

[Download PDF](#) [View details ▾](#)

[Feasibility of optimizing trimetazidine di...](#)
 Journal of Advanced Research, Volume 5,...

[Download PDF](#) [View details ▾](#)

[Extrudable core system: Development ...](#)
 Journal of Controlled Release, Volume 13...

[Purchase PDF](#) [View details ▾](#)

1 2 [Next >](#)

Authors

- Giancarlo Santus
- Richard W. Baker

Institutions

- Alza

Field of Study

Osmosis

Organic chemistry

Process engineering

Drug delivery

Medicine

Dosage form

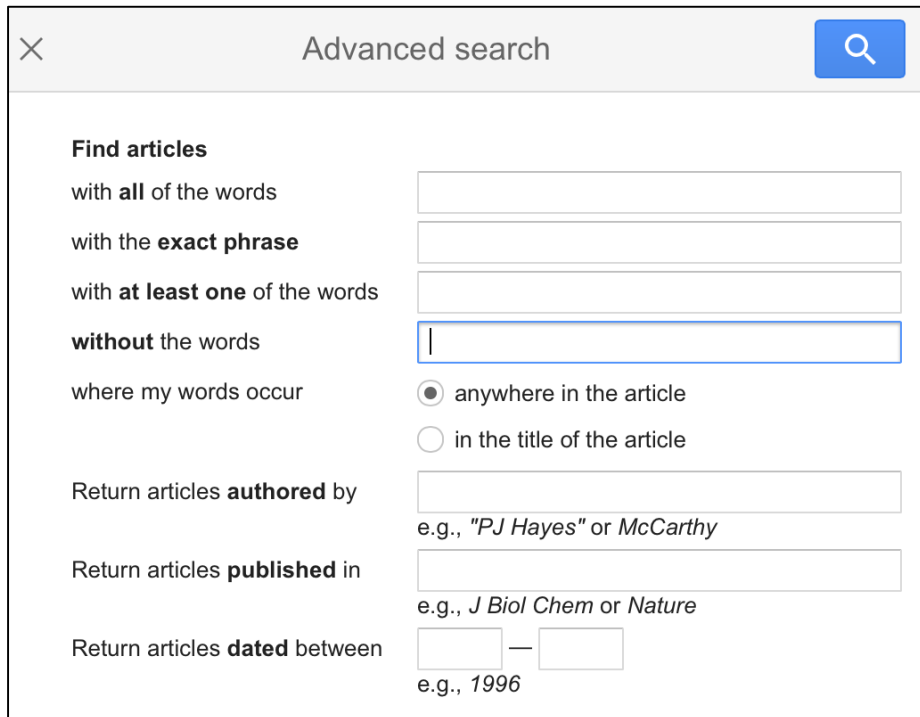
Citation Counts	
Patent Citations	206
Scholarly Citations	181
Reference Count	1

➔ **Citing articles (191)** ▾

Searching Patent-cited Scientific Literature: Google Scholar/Patents Advanced Search Interfaces

➤ Google also provides its own double database and double search interface covering scientific literature (Google Scholar, which covers also thesis and not-so-often indexed journals) and patent information (Google Patents) with links to external providers and citation metrics but seemingly:

- ✓ Less indexed / searchable fields
- ✓ Less cross-indexed information, metadata



Advanced search

Find articles

with **all** of the words

with the **exact phrase**

with **at least one** of the words

without the words

where my words occur

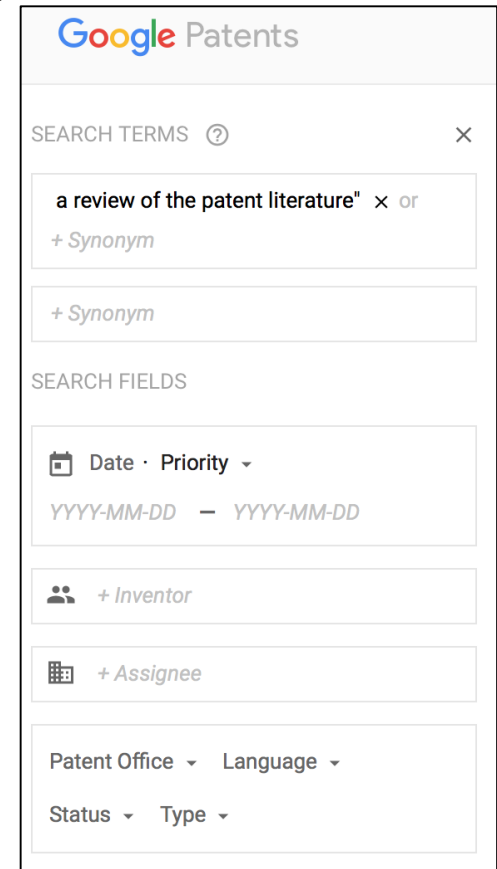
anywhere in the article

in the title of the article

Return articles **authored by**
e.g., "PJ Hayes" or McCarthy

Return articles **published in**
e.g., J Biol Chem or Nature

Return articles **dated** between —
e.g., 1996



Google Patents

SEARCH TERMS ⓘ ×

a review of the patent literature" × or
+ Synonym

+ Synonym

SEARCH FIELDS

📅 Date · Priority ▾
YYYY-MM-DD — YYYY-MM-DD

👤 + Inventor

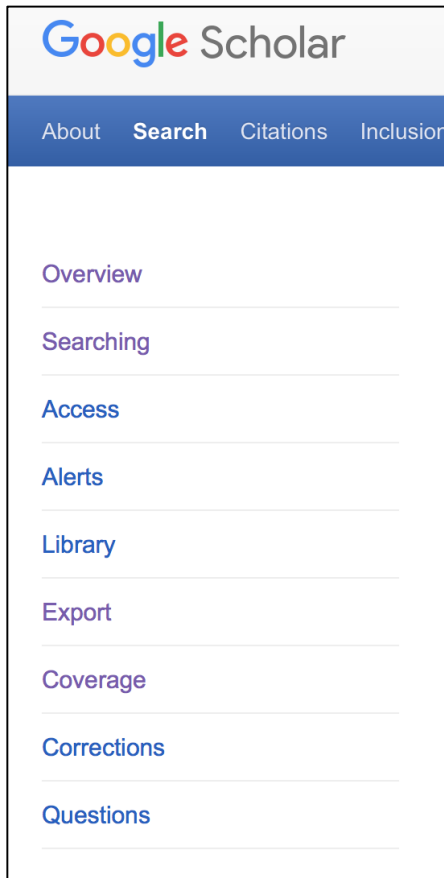
🏢 + Assignee

Patent Office ▾ Language ▾

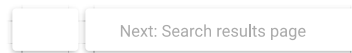
Status ▾ Type ▾

Searching Patent-cited Scientific Literature: Google Scholar /Patents Advanced Search Operators

- Google Patents Help page explains how to build quite complex searches (but not so much with Google Scholar)



About Google Patents Searching



From the homepage, you can begin your search in a few ways:

- Enter a patent publication or application number, such as [US9014905B1], [9014905], or [US 14/166, 502]

Searching by metadata (inventor, assignee, date, etc)

1. From the [homepage](#), enter one of the following field names (without the [brackets]). Press ENTER to search after typing:
 - a. Search inventors [inventor:bell] or [inventor:"Alexander Graham Bell"]
 - b. Search assignees [assignee:google] or [assignee:"Google Inc"]
 - c. Search before or after filing date (can be changed to priority or publication date after the search) [before:2010], [after:"Jan 2006"], [before:"1/2/2006"]
 - d. Search patents by country [country:US]
 - e. Search patents by status [status:grant]
 - f. Search patents by language [language:english]

Proximity: You can use proximity operators to boost the score of documents if they contain expressions near each other. *Note:* proximity operators only change ranking, not retrieval. NEAR, NEARx, NEAR/x, or /xw means matches are a maximum of x words away, in any order. WITH means 20 words away, any order, and SAME means 200 words away, any order. AJD, AJDx, ADJ/x, or +xw are the same as NEAR, but

Example: (safety ADJ/5 belt) NEAR/10 (baby OR child) SAME vehicle.

Searching in title, abstract, claims, CPC: You can search inside a specific field using field names. Use TI=(safety belt) to search in the title, AB= for the abstract and CL= for the claims. For CPCs, CPC=B60R22 will match documents with exactly this CPC, CPC=B60R22/low matches documents with this CPC or a child classification of this CPC. CPCs can also be used without a field prefix (see Searching by keywords and classifications above).

Wildcards, truncation: Although we automatically generate plurals, you can instead specify wildcard patterns of words to search. Wildcards only work on single words. The top 25 most common matches will be ORed together. *Note:* wildcards only work on English words. The wildcards are ? (zero or one character), * or \$ (zero or more characters), \$x (zero to x characters), and # (exactly one character). You can include more than one wildcard symbol per word.

Examples: *saccharide? and hydroxy*phenyl*.

Searching Patent-cited Scientific Literature: Google Scholar Actual Search Opportunities

➤ However Google Scholar can be also searched for the citations of a given article in patent literature according to “Google Patents” by searching:

- ✓ an article by the title, then
- ✓ the linked list of citing references by limiting with “Google Patents”

The image shows two screenshots of the Google Scholar interface. The top screenshot shows a search for "Osmotic drug delivery: a review of the patent literature". The search results include a list of filters on the left (Any time, Sort by relevance, include patents, include citations) and a single result on the right. An orange arrow points from the "Showing the best result for this search" text to the search bar of the bottom screenshot. The bottom screenshot shows a search for "google patents". The search results include a list of filters on the left and two results on the right, both titled "Osmotic drug delivery: a review of the patent literature".

Google Scholar "Osmotic drug delivery: a review of the patent literature"

Articles

Any time
Since 2018
Since 2017
Since 2014
Custom range...

Sort by relevance
Sort by date

include patents
 include citations

Osmotic drug delivery: a review of the patent literature
G Santus, RW Baker - Journal of Controlled Release, 1995 - Elsevier
Because patents are hard to read, to analyze, and to judge due to the peculiar style in which they are written, they constitute an under-utilized literature resource. This review analyzes US Patents on osmotic drug delivery through December 1993. The 240 patents cover a period of a little over 20 years. The very first patents dealt with variations of the Rose-Nelson osmotic pump, a rather complex device that has yet to find commercial use. During the 1970s and 1980s, Felix Theeuwes and others at the Alza Corporation made a series of ...
☆ ⓘ Cited by 339 Related articles All 3 versions

Showing the best result for this search. See all results

Google Scholar "google patents"

Articles About 44 results (0.33 sec)

Any time
Since 2018
Since 2017
Since 2014
Custom range...

Sort by relevance
Sort by date

include citations

Osmotic drug delivery: a review of the patent literature
 Search within citing articles

Controlled release hydrocodone formulations
B Oshlack, H Huang, JK Masselink... - US Patent 7,514,100, 2009 - Google Patents
Connect public, paid and private patent data with Google Patents Public Datasets
Controlled release hydrocodone formulations. Download PDF Info. Publication number US7514100B2. Authority US Grant status Grant. Patent ...
☆ ⓘ Cited by 167 Related articles All 4 versions ⓘ

Controlled release hydrocodone formulations
B Oshlack, H Huang, J Masselink, AP Tonelli - US Patent 7,943,174, 2011 - Google Patents

Searching Patent-cited Scientific Literature: Google Patents Actual Search Opportunities

- Google Patents can be then also searched with the article by title, but citation data do not correspond neither with Google Scholar nor with “pure” Google...

The screenshot shows the Google Patents search interface. The search bar contains the text "Osmotic drug delivery a review of the patent literature". Below the search bar, there are filters for "SEARCH TERMS" and "SEARCH FIELDS". The search results are displayed as a list of patents, with the first two results visible. Each result includes a title, a patent number, the inventor's name, and a brief description of the patent's content.

Google Patents

"Osmotic drug delivery a review of the patent literature"

SEARCH TERMS ⓧ × About 80 results

Sort by · Relevance ▾ Grouped by · None ▾ Results / page · 10 ▾

"Osmotic drug delivery a review of the patent literature" ×
or + *Synonym*

+ *Synonym*

SEARCH FIELDS

📅 Date · Priority ▾
YYYY-MM-DD — YYYY-MM-DD

👤 + *Inventor*

🏢 + *Assignee*

Method of inducing a CTL response
[Grant US6994851B1](#) · Thomas M. Kundig · Mannkind Corporation
Priority 1997-07-10 · Filing 1998-07-10 · Grant 2006-02-07 · Publication 2006-02-07
A method of inducing a cytotoxic T-lymphocyte (CTL) response to an antigen is disclosed, involving delivering the antigen to the lymphatic system of an animal regularly over a period of time using, e.g., an osmotic pump. The method is advantageous over prior art methods for inducing a CTL response.

Method of inducing a CTL response
[Grant US6977074B2](#) · Thomas M. Kundig · Mannkind Corporation
Priority 1997-07-10 · Filing 1998-07-10 · Grant 2006-02-07 · Publication 2006-02-07
Disclosed herein are methods for the regular delivery of the antigen to the lymphatic system of an animal delivered at a level sufficient to induce a CTL response.

The screenshot shows the Google search interface. The search bar contains the text "google patents" "Osmotic drug delivery a review of the patent literature". Below the search bar, there are filters for "All", "Images", "News", "Videos", "Shopping", and "More". The search results are displayed as a list of patents, with the first two results visible. Each result includes a title, a patent number, and a brief description of the patent's content.

Google

"google patents" "Osmotic drug delivery a review of the patent literature" 🔍

All Images News Videos Shopping More Settings Tools

About 22 results (0,41 seconds)

US20070196487A1 - Method and apparatus for drilling orifices in ...
<https://www.google.com/patents/US20070196487>
A review of such osmotic tablets is found in Santus and Baker, "Osmotic drug delivery: a review of the patent literature," Journal of Controlled Release 35 (1995) ...

US20070031496A1 - Osmotic dosage forms comprising ... - Google
<https://www.google.com/patents/US20070031496>
A review of such dosage forms is found in Santus and Baker, "Osmotic drug delivery: a review of the patent literature," Journal of Controlled Release 35 (1995) ...

Compound Coding across Patent & Scientific Literature: Nucl. Acid Res. Database of Databases (Introduction)

- The first stop for identifying scientific biotech/pharma databases that refer to and/or collect data from patent documentation is the section of *Nucleic Acid Research* journal covering the annual, free “*database issue*” published each January and searchable using the dedicated advanced search tool by using “patent”, “database” and the relevant keywords for a topic:

The screenshot displays the Nucleic Acids Research search interface. At the top, the title "Nucleic Acids Research" is in red. Below it is a dark red navigation bar with links: "Issues", "Section browse", "Advance articles", "Submit", "Purchase", "About", and "All Nucleic Acids R".

The main content area is divided into two columns. The left column, titled "Modify your search", contains a search input field, a "Filter" dropdown, and two buttons: "Add term" and "Update". Below this is a "Journal citation" section with a right-pointing triangle icon. At the bottom of the left column is a "Format" section with a radio button selected for "Journal Article (686)".

The right column shows search results. It starts with "1-20 of 686" and "Database Issue" with a close icon. Below that is a "Save search" link. A "Sort by" dropdown menu is set to "Relevance". To the right of the dropdown are page numbers "1 2 3 4 5 Next". A "Journal Article" label is positioned above the first result. The result title is "The 2018 Nucleic Acids Research database issue and the online molecular biology database collection" with an ORCID icon. The author is "Daniel J Rigden, Xosé M Fernández". The publication information is "Nucleic Acids Research, Volume 46, Issue D1, 4 January 2018, Pages D1–D7, https://doi.org/10.1093/nar/gkx1235". The publication date is "Published: 18 December 2017".

Compound Coding across Patent & Scientific Literature: Nucl. Acid Res. Database of Databases (Commentaries)

- This separate section of NAR website covers only from 2014 but prior that date such issue was simply the one of Jan. 1st every year, going back until 1996, but
 - ✓ Still more detailed for patent information on chemicals/drug than biological compounds (having a very diverse, not always consistent coding)
 - ✓ Authors of these articles are generally database originators, being mostly academics who do not provide with a fully objective / complete report
- The databases can cover very specific topics but with no guarantee about the selection criteria and database records are actually correct/updated/complete:
 - ✓ Indeed, many reported databases are “one-off” result of a project and may be not updated since their first release (or are even no more available)
 - ✓ The more reliable authors and databases are those affiliated to main bioinformatic institutions such as EBI/EMBL (Europe) and NCBI/NIH (USA)
 - ✓ Redundancy problems since the same database can be covered in more than one article, during its development over the years and/or the same codes/patent information are cross-referenced across databases

Compound Coding across Patent & Scientific Literature

NCBI Databases & Patent Information (introduction)

Search NCBI databases

patent

!

Literature

Bookshelf	3,414	Books and reports
PubMed	45,987	Scientific and medical abstracts/citations
PubMed Central	110,297	Full-text journal articles
PubMed Health	623	Clinical effectiveness, disease and drug reports

Genes

EST	1,005,851	Expressed sequence tag sequences
Gene	54	Collected information about gene loci
GEO DataSets	1,158	Functional genomics studies
GEO Profiles	7,440	Gene expression and molecular abundance profiles

Proteins

Identical Protein Groups	2,307,966	Protein sequences grouped by identity
Protein	5,631,344	Protein sequences

Chemicals

PubChem Substance	15,227,775	Deposited substance and chemical information
-----------------------------------	------------	--

➤ NCBI web site provide a global view of a number of literature-based, compound-based, and other scientific databases of interest for biotech/pharma searches

Compound Coding across Patent & Scientific Literature: NCBI Databases & Patent Information (Protein/DNA)

- However their indexing is limited as content and/or sources (mainly USPTO), with redundancy across republished WO/US/EP documents and codes poorly used elsewhere (in particular within scientific publications)

Nucleotide

Items: 1 to 20 of 7379073

- [Sequence 1 from Patent WO2013068413](#)
 - 2,011 bp linear DNA
Accession: JB393413.1 GI: 549038390
[GenBank](#) [FASTA](#) [Graphics](#)
- [Sequence 216 from Patent EP2722395](#)
 - 36 bp linear DNA
Accession: JC559089.1 GI: 684925391
[GenBank](#) [FASTA](#) [Graphics](#)
- [H.sapiens gene for B cell differentiation factor 1](#)
 - 3,230 bp linear DNA
Accession: X12706.1 GI: 29392
[GenBank](#) [FASTA](#) [Graphics](#)
- [H.sapiens mRNA for endothelial plasminogen activator inhibitor PAI](#)
 - 1,482 bp linear mRNA
Accession: X12701.1 GI: 31146
[GenBank](#) [FASTA](#) [Graphics](#)
- [H.sapiens mRNA for protein S](#)
 - 3,344 bp linear mRNA
Accession: X12892.1 GI: 35692
[GenBank](#) [FASTA](#) [Graphics](#)

Protein

- [protein histidine phosphatase interacting partner, PHPIP-180, partial \[Homo sapiens\]](#)
 - 571 aa protein
Accession: CAE11964.1 GI: 32954229
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [unnamed protein product \[Homo sapiens\]](#)
 - 362 aa protein
Accession: CAJ44921.1 GI: 83410955
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [unnamed protein product, partial \[Homo sapiens\]](#)
 - 232 aa protein
Accession: CAK95595.1 GI: 109715524
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [binding protein \[Homo sapiens\]](#)
 - 328 aa protein
Accession: CAA00862.1 GI: 412023
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [unnamed protein product \[Homo sapiens\]](#)
 - 353 aa protein
Accession: CAH69774.1 GI: 55581783
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)
- [Tpr1 \[Homo sapiens\]](#)
 - 726 aa protein
Accession: CAD86491.1 GI: 29370785
[GenPept](#) [Identical Proteins](#) [FASTA](#) [Graphics](#)

Compound Coding across Patent & Scientific Literature: NCBI Databases & Patent Information (PubChem)

- This NCBI/NIH database provides “a public repository for small-molecule and RNAi data” associating information from many sources, including own unpublished assays, with a code that is cited present in articles and links to other NCBI databases (including PubMed/MeSH) but:
 - ✓ Patent information is generally US only, with title only
 - ✓ NO specific filter/keyword for patent information
 - ✓ Confusing code (only a number)

STRUCTURE VENDORS DRUG INFO PHARMACOLOGY LITERATURE PATENTS BIOACTIVITIES

PubChem CID: 3415

Chemical Names: Foscarnet; Phosphonoformic acid; Phosphonoformate; Carboxyphosphonic acid; Foscarnet; 4428-95-9





Molecular Formula: $\text{CH}_3\text{O}_5\text{P}$

Molecular Weight: 126.004 g/mol

InChI Key: ZJAOAACCNHFJAH-UHFFFAOYSA-N

Drug Information:

1 to 5 of 6,247 ... ▾

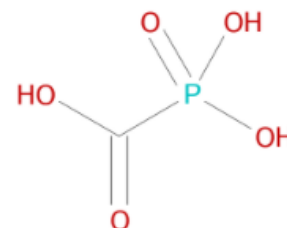
Patent ID 	Title 	Submitted Date 	Granted Date 
US8367701	Crystalline pharmaceutical and methods of preparation and use thereof	2011-11-04	2013-02-05
US7820660	N-Substituted hydroxypyrimidinone carboxamide inhibitors of HIV integrase	2008-11-06	2010-10-26

Compound Coding across Patent & Scientific Literature: EBI Databases & Patent Information (ChEmbl)

- This EBI database provides “an open large-scale bioactivity database” associating information from many sources, with a code that is cited/present in articles and links to many other databases including the parallel EBI chemical patent databases SureChEmbl (also searchable by structure, patent no., keywords, and other numerical range/similarity criteria) but still NO specific filter/keyword for patent information

Compound Name and Classification

Compound ID	CHEMBL666
Compound Name	FOSCARNET
ChEMBL Synonyms	Foscavir FOSCARNET SODIUM EHB 776 Dihydroxyphosphinecarboxylic Acid Oxide Foscarnet
Max Phase	4 (Approved)
Trade Names	FOSCARNET SODIUM FOSCAVIR
Molecular Formula	CH3O5P



PharmGKB	PA449706
PubChem	3415
PubChem: Thomson Pharma	15194456
SureChEMBL	SCHEMBL23633
ZINC	ZINC000008101109

UniChem Cross References Related structures Patent hits

UniChem cross references (Click to expand)

ChEMBL (1)	DrugBank (1)	PDBe (1)	Guide to Pharmacology (1)	KEGG Ligand (1)	ChEBI (1)
ZINC (1)	IBM Patent System (1)	FDA SRS (1)	PharmGKB (1)	Human Metabolome Database (1)	
PubChem: Thomson Pharma (1)	PubChem (1)	ACToR (1)	Nikkaji (1)	BindingDB (1)	
EPA CompTox Dashboard (1)	DrugCentral (1)	Brenda (2)			

UniChem Cross References Related structures Patent hits

Showing 1-3 of 20,946 total documents found for selected structure

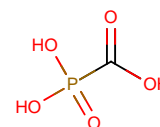
Publication Number	Publication Date	IPCR	Assignee/Applicant	Structure hits
1. US-20180220943-A1	2018-08-09	A61B 5/1486	DEXCOM, INC. ZWITTERION SURFACE MODIFICATIONS FOR CONTINUOUS SENSORS	
2. US-20180222890-A1	2018-08-09	C07D 403/14	VIIV Healthcare UK (No.5) Limited 5-(N-[6,5]-FUSED BICYCLIC ARYL TETRAHYDROISOQUINOLIN-6-YL) PYRIDIN-3-YL ACETIC ACID DERIVATIVES AS INHIBITORS OF HUMAN IMMUNODEFICIENCY VIRUS REPLICATION	

Compound Coding across Patent & Scientific Literature: Drug Databases & Patent Information (DrugBank)

- This Canadian database provides “*molecular information about drugs, their mechanisms, their interactions and their targets*” associating information from many sources in a single record for both chemical and biological drugs, with a code (general format DBnnnnn) that is cited present in many articles but:
 - ✓ Patent information is generally CA/US only, not specifying if each cited patent covers composition-of-matter, specific use, production process, etc.
 - ✓ NO specific filter/keyword for patent information (embedded into Pharmacoconomics” super-field, but not systematically completed)
 - ✓ Recent move towards a “commercial” model

Name	Foscarnet
Accession Number	DB00529 (APRD00669)
Type	Small Molecule
Groups	Approved
Description	An antiviral agent used in the treatment of cytomegalovirus retinitis. Foscarnet also shows activity against human herpesviruses and HIV. [PubChem]
Patents	Not Available

Structure



🔍 3D Download ▾ 🔗 Similar Structures

Synonyms

Carboxyphosphonic acid
Foscarnet
Phosphonoformate
Phosphonoformic acid
Phosphonomethanoic acid

Compound Coding across Patent & Scientific Literature: Drug Databases & Patent Information (IUPHAR-DB)

- This database provides “*expert-curated molecular interactions between successful and potential drugs and their targets in the human genome*” but:
 - ✓ Patent information is unevenly completed and embedded in specific fields that cannot be specifically searched using a code / field
 - ✓ Actual interest is how the information is assembled from various sources

Summary		Biological activity	Clinical data	Structure	Similar ligands
Classification ?					
Compound class		Synthetic organic			
Approved drug?		Yes (FDA (1991))			
IUPAC Name ?					
phosphonoformic acid					
International Nonproprietary Names ?					
		INN number			
4712				foscarnet sodium	
Synonyms ?					
carboxyphosphonic acid Foscavir® phosphonocarboxylic acid phosphonoformic acid					
Comments					
Foscarnet is an antiviral medication.					

Database Links ?	
CAS Registry No.	4428-95-9
ChEBI	CHEBI:127780
ChEMBL Ligand	CHEMBL666
DrugBank Ligand	DB00529
GtoPdb PubChem SID	178102144
PubChem CID	3415
RCSB PDB Ligand	PPF
Search Google for chemical match using the InChIKey	ZJAOAACCNHFJAH-UHFFFAOYSA-N
Search Google for chemicals with the same backbone	ZJAOAACCNHFJAH
Search PubMed clinical trials	foscarnet sodium
Search PubMed titles	foscarnet sodium
Search PubMed titles/abstracts	foscarnet sodium
Search UniChem for chemical match using the InChIKey	ZJAOAACCNHFJAH-UHFFFAOYSA-N
Search UniChem for chemicals with the same backbone	ZJAOAACCNHFJAH
Wikipedia	Foscarnet

Compound Coding across Patent & Scientific Literature: Enhanced Use & Access in Scientific Literature

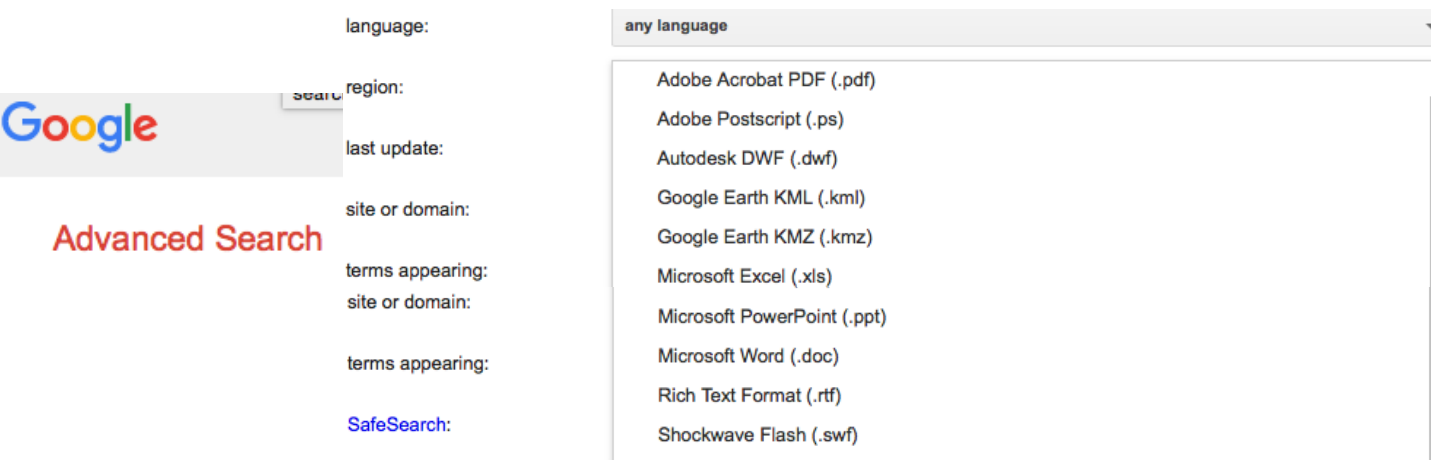
- Aside from reviews summarizing patent literature, patent references are still unevenly, rarely present in scientific literature, especially for drugs
- Compound-specific, patent-related codes from scientific databases are now more commonly used in articles, especially for chemical compounds, due to:
 - ✓ Greater, simpler experimental access to chemical libraries with patent-related codes by public institutions for large compound screening
 - ✓ Greater, simpler in silico access to chemical information through databases such as DrugBank, Pubchem, ChEMBL, SureChEMBL
 - ✓ More interest in translating the findings from pharmaceutical, commercial research in more basic, public research context with screening equipment
 - ✓ More common use of (large) supplementary information files in web version of scientific articles for providing full access to original “raw data”
- However neither scientific literature nor compound databases index which articles refer to compounds by means of corresponding patent-related codes

Compound-related Patent Information: (Advanced) Google Search & Scientific Literature

- Patent-related and other compound coding and commercial names can be made available through scientific literature (where such compounds are tested and compared) in a variety of formats summarized in two main categories:
 - ✓ Within text, figures, or tables of the PDF of the article
 - ✓ Within supplementary files of an article (available in PDF or other formats, in particular .xls, .txt, .xlsx, .doc, .tsv, .csv, or .sdf)
- Such files are indexed by Google in an uneven manner, depending from the publisher's policy, but, when they are actually available, they require:
 - ✓ Identifying the relevant database compound codes (by prior searches)
 - ✓ Making use of Advanced Google functions, in particular for file extension
- The coverage of information about a drug or other compounds, in particular for evaluating evidence with respect to specific assays/uses, can be evaluated in a more complete manner, also in view of experiments that have been performed by others than drug originators/ licensees

Compound-related Patent Information: Advanced Google Search Main Info

- Google search page (by using “filetype:” or “ext:” search operator) can be used for finding references to database and database codes in freely available files
- Regular Advanced Google search page suggests a series of ten available formats



- However, other file extensions are actually supported, as many of those used in the supplementary files associated to scientific articles available for downloading and analysis, and also to some databases and websites, including:
 - ✓ Generic formats (filetype:txt filetype:sql filetype:csv filetype:tsv)
 - ✓ Scientific format (filetype:tex filetype:sdf filetype:dat filetype:xml)

Compound-related Patent Information: Advanced Google Search Additional Info

- Using the keyword “*drug*” and no limitation, Google indexes 550 millions and 99.5 millions with filetype:pdf, but other filetype limitations can be applied

filetype:doc 440.000

filetype:xml 406.000

filetype:txt 386.000

filetype:ppt 101.000

filetype:rtf 80.300

filetype:xls 79.900

filetype:dat 27.500

filetype:tex 10.800

filetype:csv 9.960

filetype:sql 567

filetype:tsv 26

filetype:sdf 18

- Such “deep” information require some time and it may be useful for obtaining the most complete coverage possible about a compound:
 - ✓ Identifying which database code for a compound also refers to (reliable) patent references with minimal redundancy/precision over other codes
 - ✓ Testing different file extensions (possibly using other relevant keywords)
 - ✓ Testing different site limitations (using the “site:” operator as well, eg “*drug*” filetype:pdf site:.gov OR “*drug*” filetype:pdf site:.nih.gov)
 - ✓ Sometime, making use of the right software to actually read the content of file identified in this manner

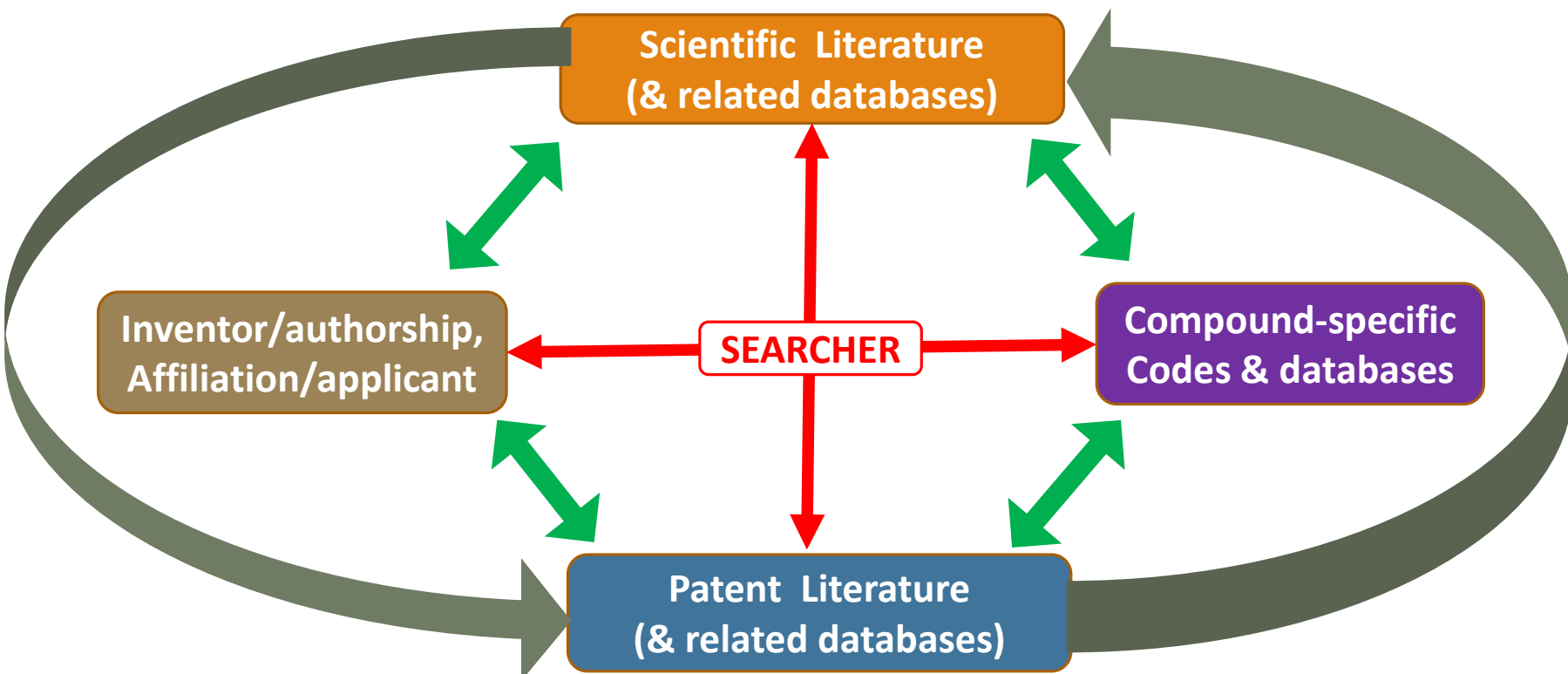
Final Observations: (Fixing) Links Between Patent & Scientific Literature

- The patent-relevant search in pharma & biotech scientific literature requires:
 - ✓ Some substantial preparation in mapping keywords and relevant resources
 - ✓ Testing the search strategies across the resources
 - ✓ Comparing names, findings, but still being aware of and making the distinction between what is experimentally confirmed and what is suggested within such disclosures
- Even more than the search in patent documentation, the preliminary analysis, the search execution, and analysis of results in scientific literature requires time and discussions with client, who may already use sources covering both patent and scientific literature but is rarely aware of actual limitations & opportunities
- Progresses are made by various information providers but still a lot relies upon the searcher's experience and choices in navigating the two oceans

Final Observations: Central Role of Searcher

➤ The citation flow is continuous between patent and scientific literature but links are deeper than expected and of major interest for the searcher, who has a central position when dealing with strategic objectives such as

- ✓ Legally relevant topics
- ✓ Competitive intelligence



References

➤ Main cited websites

[Pubmed Tutorial](#)

[Pharm Patent Analyst](#)

[J. Tech Transfer](#)

[Lens.org Support Page](#)

[NCBI Tutorial](#)

[ChEMBL FAQ](#)

[Google Advanced Search Page](#)

[Expert Opin Ther. Patents](#)

[World Patent Inf.](#)

[Google Scholar Help Page](#)

[Pubchem Help page](#)

[DrugBank Documentation](#)

[NAR database section](#)

[Recent Patents series](#)

[Technovation](#)

[Google Patents Help Page](#)

[EBI Services](#)

[Iuphar-DB](#)

➤ Some free readings

Bousfield D et al. , Patterns of database citation in articles and patents indicate long-term scientific and industry value of biological data resources. [F1000Res. 2016 Feb 11;5](#)

Ahmadpoor M, Jones BF. The dual frontier: Patented inventions and prior scientific advance. [Science. 2017 Aug 11;357\(6351\):583-587](#)

Thank you and (maybe) see you in Brussels!!
Luca.falciola@droiagroup.com



The views and the opinions expressed in this presentation are the author's personal thoughts on these subjects. They are not intended to be considered opinions and positions of Droia Oncology Ventures nor imply any commitment by Droia Oncology Ventures to any particular action.