ChemAnalyser Mission:

provide added value to customers in chemistry patent searching:

- Cognitive Search®
- background knowledge ontologies
- Big Data analytical tools

supporting decision making:

*Freedom-to-operate opinion*
*more comprehensive through full text annotations*
*compare with non-patent literature*
*ontology based classifications and analytics*
SciFinder & CAS – advancing human annotations...

...problem:
Introduction

Solution:

human & automated annotations
Exact Search 1: More Comprehensive

Searching “pinocarvone” yields same results:

Drawing structure:

or using name

powered by ChemAxon
Searching “pinocarvone” in patent and non-patent literature:

303 patent hits in total (compared to 53); 20 patent hits in 2009
Searching “pinocarvone” in SciFinder:

Only 1 patent document hit in 2009, compared to 20 from in ChemAnalyser!
Switching to patent family view in sem-ip.com
Searching “pinocarvone” in ChemAnalyser from 2009:

export 20 patent document hits into Excel, from 8 patent families:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>COMPOSITIONS AND METHODS FOR TREATING HOOF DISEASES</td>
<td>2009-11-25</td>
<td>2010-06-03</td>
<td>US2010137451A1</td>
<td>A61K31/11; A61P31/00</td>
</tr>
<tr>
<td>4</td>
<td>Hard Surface Cleaner Compositions of Sulfonated Estolides and Other Derivatives</td>
<td>2009-01-21</td>
<td>2010-07-22</td>
<td>US2010184855A1</td>
<td>A01N41/04; C11D1/12; C11D1/37; C11D1/40</td>
</tr>
<tr>
<td>5</td>
<td>PVC FOAM PANEL AND CABINET FOR BATHROOM</td>
<td>2009-12-17</td>
<td>2011-06-23</td>
<td>KR20110069411A</td>
<td>A47B67/02; B29C67/20</td>
</tr>
<tr>
<td>6</td>
<td>PVC foam panel and cabinet for bathroom</td>
<td>2009-12-17</td>
<td>2011-12-27</td>
<td>KR101099440B1</td>
<td>A47B67/02; B29C67/20</td>
</tr>
<tr>
<td>9</td>
<td>Walnut shell alcohol preparation and applications thereof for preparing medicine</td>
<td>2009-11-06</td>
<td>2010-05-12</td>
<td>CN101703569A</td>
<td>A61K131/00; A61K31/045; A61K36/52</td>
</tr>
<tr>
<td>10</td>
<td>DISINFECTANT CLEANING COMPOSITION AND METHOD</td>
<td>2009-10-02</td>
<td>2010-07-08</td>
<td>US2010173011A1</td>
<td>A01N31/16; A01N31/12; A01P1/00; A61K31/11</td>
</tr>
<tr>
<td>11</td>
<td>DISINFECTING AGENT COMPRISING EUGENOL, TERPINEOL AND THYMOL</td>
<td>2009-09-24</td>
<td>2012-11-01</td>
<td>US2012276022A1</td>
<td>A01N31/16; A01N31/12; A01P1/00; A61K31/11</td>
</tr>
<tr>
<td>12</td>
<td>Disinfecting agent comprising eugenol, terpineol and thymol</td>
<td>2009-09-24</td>
<td>2012-04-05</td>
<td>AU2010297406A1</td>
<td>A01N25/02; A01N25/34; A01N31/04; A61K31/11</td>
</tr>
<tr>
<td>13</td>
<td>DISINFECTING AGENT COMPRISING EUGENOL, TERPINEOL AND THYMOL</td>
<td>2009-09-24</td>
<td>2011-03-31</td>
<td>CA2801143A1</td>
<td>A01N25/30; A01N31/04; A01N31/08; A61K31/11</td>
</tr>
<tr>
<td>14</td>
<td>Disinfecting agent comprising eugenol, terpineol and thymol</td>
<td>2009-09-24</td>
<td>2015-09-15</td>
<td>US9132103B2</td>
<td>A01N31/04; A01N31/06; A01N31/08; A61K31/11</td>
</tr>
<tr>
<td>15</td>
<td>DISINFECTING AGENT COMPRISING EUGENOL, TERPINEOL AND THYMOL</td>
<td>2009-09-24</td>
<td>2011-03-31</td>
<td>WO2011036048A1</td>
<td>A01N25/02; A01N25/34; A01N31/04; A61K31/11</td>
</tr>
<tr>
<td>16</td>
<td>Novel biopesticide compositions and method for isolation and characterization</td>
<td>2009-07-28</td>
<td>2012-02-23</td>
<td>AU2009350416A1</td>
<td>A01N65/00; A01N65/28; A01P7/00</td>
</tr>
<tr>
<td>17</td>
<td>NOVEL BIOPESTICIDE COMPOSITIONS AND METHOD FOR ISOLATION AND CHARACTERIZATION</td>
<td>2009-07-28</td>
<td>2011-02-03</td>
<td>WO2011013332A2</td>
<td>A01N65/00; A01N65/28; A01P7/00</td>
</tr>
<tr>
<td>18</td>
<td>Novel Biopesticide Compositions And Method For Isolation And Characterization</td>
<td>2009-07-28</td>
<td>2012-05-24</td>
<td>US2012128648A1</td>
<td>A01N59/00; A01N59/20; A01N59/26; A61K31/11</td>
</tr>
</tbody>
</table>
ChemAnalyser structure search: 1 structure hit, 7 patent hit documents
SciFinder structure search in registry: no hit
SciFinder via Markush structure search: 3 patent hit documents
Results:

SciFinder

- Triazine, pyrimidine and pyridine analogs and their use as therapeutic agents and diagnostic probes and their preparation; WO 2010052569 A2 (changed title from original !)
- Preparation process for triazine, pyrimidine and pyridine derivatives via Suzuki coupling; WO 2015162084 A1
  Does not contain the specified compound as an example but only in Markush
- Preparation of substituted pyrimidines as PI3K kinase inhibitors useful in the treatment of cancer; CN 103483345 A
  Does not contain the specified compound as an example but only in Markush

ChemAnalyser

all contain the specified compound:

- TRIAZINE, PYRIMIDINE AND PYRIDINE ANALOGS AND THEIR USE AS THERAPEUTIC AGENTS AND DIAGNOSTIC PROBES; WO 2010052569 A2
- TRIAZINE, PYRIMIDINE AND PYRIDINE ANALOGS AND THEIR USE AS THERAPEUTIC AGENTS AND DIAGNOSTIC PROBES US 2011275762 A1
- Triazine, pyrimidine and pyridine analogs and their use as therapeutic agents and diagnostic probes; US 8921361 2
- Triazine, pyrimidine and pyridine analogs and their use as therapeutic agents and diagnostic probes; NZ 592617 A
- Triazine, pyrimidine and pyridine analogs and their use as therapeutic agents and diagnostic probes; AU 2009312464 A1
- TRIAZINE, PYRIMIDINE AND PYRIDINE ANALOGS AND THEIR USE AS THERAPEUTIC AGENTS AND DIAGNOSTIC PROBES; CA 2741990 A1
- Triazine, pyrimidine and pyridine analogues and their use in therapy; GB 2465405 A
ChemAnalyzer via Markush type structure search: 216 patent hit documents
TRIAZINE, PYRIMIDINE AND PYRIDINE ANALOGS AND THEIR USE AS THERAPEUTIC AGENTS AND DIAGNOSTIC PROBES

US2011275762A1

9. A compound of claim 1, wherein R³, R⁴, R³, R³, R³ are independently selected from the structures:

10. A compound of claim 1 selected from:

[Chemical structures and formulas are shown in the image, including various combinations of atoms and substituents.]

Concept highlighting priority:

- **Query term**
- **Anatomy (other)**
- **Anatomy (plants/fungi)**
- **Cell line**
- **Chemistry: class**
- **Chemistry: compound**
- **Chemistry: group**
- **Cosmetology**
- **Disease**
- **Effect**
- **Material properties**
- **Methods**
- **Nutrition**
- **Physiology**
- **Polymer**

Export document: Export TOP 15 concepts per annotation type:
US2011275762A1: TRIAZINE, PYRIMIDINE AND PYRIDINE ANALOGS AND THEIR USE AS THERAPEUTIC AGENTS AND DIAGNOSTIC PROBES

Searching 216 patent hit documents via Markush type structure search.

US2011275762A1: TRIAZINE, PYRIMIDINE AND PYRIDINE ANALOGS AND THEIR USE AS THERAPEUTIC AGENTS AND DIAGNOSTIC PROBES

[0281] The Suzuki-type coupling reaction is useful to attach a cyclic heteroaryl at the 6-position of the triazine or pyridazine ring (see Schemes 7 and 27), or on another ring of the pyrimidine ring (see Schemes 46 and 51). Generally, 4,4′-6-chloro-1,3,6-triazine-2,4-diyl)dimorpholine [77] is used to react with 4-ethylmorpholine [79]. A catalyst amount, or more, of a palladium reagent such as bis(tri(1-phenylphosphine)pyridine)nickel(0) [80] is added and the high pressure glass vessel containing the mixture was bubbled with argon gas and sealed. A variety of boronic acid or boronic ester can be used in place of the boronic ester indicated. Also, the nitrogens of the pyrimidine-2-amines may be protected, for example with a tert-butyldimethylsilyl group. The reaction mixture was then heated at 90 °C for 15 hours, cooled down and diluted with water. The organic layer was washed with mixture of water and Na2CO3 saturated with t-butyl alcohol and NaHCO3 saturated with water. The organic layer was dried and filtered to yield pure crude Intermediate. This Intermediate was purified by column chromatography or preparative TLC to yield product 85 or 87.

General Procedure C Amide Coupling:

[0284] 7-morpholino-5-(pyrimidin-5-yl)thiazolo[4,5-d]pyrimidine-2-carboxylic acid [94] or 4-pyridylfuranylimidazole [96] is treated with 1.5 eq. HATU, 1 eq. of an amine (RNH2) and 3 eq. of DIEA in DMSO to approximately 0.1 M concentration. The reaction mixture was stirred until complete and extracted in ethyl acetate with saturated sodium bicarbonate solution. The organics were separated, washed with water, and dried with MgSO4. The residue was purified by column chromatography.

General Procedure C-1 Amide Coupling:

[0286] 4-2-piperazin-1-ylmethyl-5-(pyrimidin-5-yl)thiazolo[4,5-d]pyrimidine-7-ylmorpholine [80] or pyridylfuranylimidazole [96] is treated with 1.5 eq. HATU, 3 eq. of DIEA in DMF to approximately 0.1 M concentration. The reaction mixture was stirred until complete and extracted in ethyl acetate with saturated sodium bicarbonate solution. The organics were separated, washed with water, and dried with MgSO4. The residue was purified by preparative TLC.

General Procedure D Reductive Amination:

[0287] 7-morpholino-5-(pyrimidin-5-yl)thiazolo[4,5-d]pyrimidine-2-carboxaldehyde [92] or 4-pyridylfuranylimidazole [96] was dissolved in a 0.2 M concentration in dichloromethane. To this solution was added 1.5 eq. of 1,10-phenanthroline (RTH), 10 eq. of diethylthiocarbamate, and 1 eq. of nitric acid. The mixture was stirred for 2-6 hours prior to adding 1.5 eq. equivalents of sodium borohydride. Following 12 to 18 hours of stirring the reaction was poured into saturated sodium bicarbonate and extracted several times with ethyl acetate. This intermediate was either purified on silica gel or used crude in the next reaction.

Concept highlighting priority:

- Query term
- Anatomy (other)
- Anatomy (plants/fungi)
- Cell line
- Chemistry: class
- Chemistry: compound
- Chemistry: group
- Cosmetology
- Disease
- Effect
- Material properties
- Methods
- Nutrition
- Physiology
- Polymer
- Scaffold
- Small molecule
Assignment of compounds in structure files, databases, documents

SMARTS based

- e.g. alkylbromides
- \([C;\!\!\!$(C=[O,S,N,P])]\)Br
- D. Weininger

AND, OR, NOT logic

- e.g. AND NOT for alkenes:
  - \(C=C\)
  - !c:c

OR

http://www.ontochem.de/images/downloads/sodiac-academic-2.0.7.zip powered by ChemAxon
Understanding Concepts

Context sensitive annotation

homonym resolution: disambiguate based on environment – material properties:

gray terms: only annotate in desired context:

+ 15 wt.% Fe + 25 wt.% Fe -3 -2 -10 1 Applied field Polarisation \( \mathbf{J} \) (T) Heat flow exothermic (a.u.) 23 \( \mu \text{H} \) (T) 4 5 6 Fig. 3.
### Cognitive Search

**Search term “tomatidine”** (a class of natural products for food):
- ChemAnalyser: 664 patent hit documents
- SciFinder: 57 patent hit documents

**Search term “sequiterpenes”** (a class of natural products for food):
- ChemAnalyser: 318,707 patent hit documents
- SciFinder: 1,837 patent hit documents

**Search terms “food additives” + “natural products” + “triterpenes”**
- ChemAnalyser: 362,406 patent hit documents
- SciFinder: 33 all hit documents
### Searching Materials

<table>
<thead>
<tr>
<th>Search term</th>
<th>ChemAnalyser</th>
<th>SciFinder</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;neodymium iron boron magnet&quot;</td>
<td>25,616 patent hit documents</td>
<td>0 patent hit documents</td>
</tr>
<tr>
<td>&quot;NdFeB&quot;</td>
<td>25,616 patent hit documents</td>
<td>6 patent hit documents</td>
</tr>
<tr>
<td>&quot;304 stainless steel&quot;</td>
<td>14,578 patent hit documents</td>
<td>2,487 patent hit documents</td>
</tr>
<tr>
<td>&quot;cosmetic formulation&quot; + &quot;glyzyrrhizin&quot;</td>
<td>51,024 patent hit documents</td>
<td>No References were found containing all of the concepts &quot;cosmetic&quot;, &quot;formulation&quot; and &quot;glyzyrrhizin&quot;</td>
</tr>
</tbody>
</table>
### Statistics

**Infoapps:** >90 million full text documents from 110 patent offices  
only 4.718.519 million chemistry composition of matter patents

<table>
<thead>
<tr>
<th>Domain</th>
<th>annotation count</th>
<th>sum unique concepts deduplicated per document</th>
<th>deduplicated unique concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>chemistry compounds</td>
<td>1,465,510,682</td>
<td>360,913,205</td>
<td>294,771,572</td>
</tr>
<tr>
<td>Proteins</td>
<td>204,902,329</td>
<td>30,167,344</td>
<td>67,993</td>
</tr>
<tr>
<td>Anatomy non-plants</td>
<td>126,856,048</td>
<td>21,192,154</td>
<td>2,378</td>
</tr>
<tr>
<td>Methods</td>
<td>112,230,880</td>
<td>21,725,977</td>
<td>1,959</td>
</tr>
<tr>
<td>Species</td>
<td>105,618,715</td>
<td>25,901,359</td>
<td>81,036</td>
</tr>
<tr>
<td>Diseases</td>
<td>82,857,385</td>
<td>24,592,233</td>
<td>21,367</td>
</tr>
<tr>
<td>Physiology</td>
<td>68,504,035</td>
<td>12,703,542</td>
<td>497</td>
</tr>
<tr>
<td>Nutrition</td>
<td>59,367,731</td>
<td>12,839,777</td>
<td>3,861</td>
</tr>
<tr>
<td>Cosmetology</td>
<td>23,465,151</td>
<td>4,883,741</td>
<td>920</td>
</tr>
<tr>
<td>Anatomy plants, fungi</td>
<td>22,326,124</td>
<td>4,212,548</td>
<td>802</td>
</tr>
<tr>
<td>Cell lines</td>
<td>9,857,621</td>
<td>2,325,743</td>
<td>2,079</td>
</tr>
<tr>
<td>Toxicity</td>
<td>7,986,832</td>
<td>2,858,977</td>
<td>423</td>
</tr>
<tr>
<td>Species plants, fungi</td>
<td>7,444,143</td>
<td>2,345,605</td>
<td>7,347</td>
</tr>
<tr>
<td>Regions</td>
<td>6,974,421</td>
<td>2,781,913</td>
<td>1,040</td>
</tr>
<tr>
<td>Herbal drugs</td>
<td>162,729</td>
<td>46,830</td>
<td>131</td>
</tr>
</tbody>
</table>
Thanks!

please contact us for

• working on your cognitive Big Data computing project
• test account to ChemAnalyser®
• supplying you with custom ontologies and tools

visit www.chemanalyser.com