



ontochem

SciWalker - a Novel Comprehensive Semantic Chemistry Search Engine for Heterogeneous Documents and Databases

NIH Ultralarge DB's. December 3, 2020

Literature and Databases

There is much more research content available, let's use it!



Semantic Search

Apply semantic intelligence to

- deep data mining
- simple to use
- normalizing content
- bulk data downloads

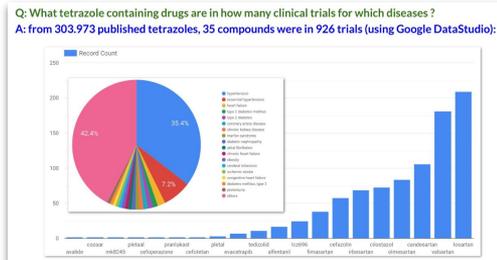


SciWalker

Analysis & Presentation

Online dashboards provide

- easy access & downloads
- automated updates
- large scale data aggregation



Research Literature



U.S. National Library of Medicine
ClinicalTrials.gov

There is much more research content available, let's use it!

256.972.611 documents are indexed, daily updates:

Database	Count
CLINICAL TRIALS (AACT)	353 572
CLINICAL TRIALS (EUDRACT)	92 747
CLINICAL TRIALS (ICTRP)	654 021
CLINICAL TRIALS (TRIALTROVE)	345 019
CORE	63 574 181
DRUG LABELS	302 896
GRANTS EU	131 320
GRANTS NIH	2 536 089
JOURNALS	1 642 633
MEDLINE	31 779 436
PATENTS	24 754 735
PATENTS DOCDB	127 451 609
PUBMED CENTRAL	3 354 353

SciWalker integrates with

Publicly available data: PubMed, clinical trials, patents, drug labels, grants, PB data in open BigQuery tables

Commercial Databases: Dimensions: 110 million publications thereof 77m full text

SciWalker semantic intelligence:

aiming to support pharmaceutical, chemistry, biology and life and material sciences

federated search in heterogeneous document resources and databases:

- Which **tetrazole** containing **drugs** were used in clinical trials for **hypertension** ?
- Which **tetrazoles** are **claimed** in patents ?
- Which **companies** produce **polymers** ?
- What are known biological **activities** of **natural products** found in **green tee** ?
- What is known about **corticosteroids** ?

... and many more

Semantic search:
simple & fast

SciWalker

cortico-steroids

SciWalker “knows”
35 million chemical concepts!

compare to the same query using
standard search technologies:

11-Dehydrocorticosterone OR
11-Deoxycorticosterone OR
11-Deoxycortisol 11-Ketoprogesterone OR
11 β -Hydroxypregnenolone OR
11 β -Hydroxyprogesterone
11 β ,17 α ,21-Trihydroxypregnenolone OR
17 α ,21-Dihydroxypregnenolone OR
17 α -Hydroxypregnenolone OR
17 α -Hydroxyprogesterone OR
18-Hydroxy-11-deoxycorticosterone OR
18-Hydroxycorticosterone
18-Hydroxyprogesterone OR 21-Deoxycortisol
21-Deoxycortisone OR 21-Hydroxypregnenolone
Aldosterone OR Corticosterone OR Cortisol OR
Cortisone OR Pregnenolone OR Progesterone OR ...

SciWalker
👤 Iweber

Dashboard Document Search Chemistry Finder Analytics My Research
Administration Help

Simple search Advanced search Expert search

cortico-steroids X
Type query (e.g. liver disease)

*OntoChem is capable of fully integrating with **Dimensions from Digital Science**. Developed using more than 110 million scientific publications, 70% of which are in full text, and more than 40 million patents, clinical trials and grants, Dimensions offers a broad database. This can be technically integrated in a number of ways, including by accessing the existing Dimensions metadata collection in Google BigQuery. According to information from Digital Science, UberResearch/Dimensions is already supporting NCI in different projects.*

CLINICAL TRIALS (AACT) 36 726	CLINICAL TRIALS (EUDRACT) 26 086	CLINICAL TRIALS (ICTRP) 72 221	CLINICAL TRIALS (TRIALTROVE) 44 262	CORE 267 160	DRUG LABELS 47 667
GRANTS EU 60	GRANTS NIH 26 199	JOURNALS 8 140	MEDLINE 309 086	PATENTS 285 820	PATENTS DOCDB 22 837
PUBMED CENTRAL 176 401					

36 726 hits in Clinical Trials (AACT)

Author

Chemistry (domain)

Companies (domain)

Drugs (domain)

Species (domain)

Substances (domain)

Sort column: Relevance descending Sort

36 726 documents found
Export XLSX

	Details	Read	Title	Trial ID	Start date	Phases	Sponsors	Source	Preview
<input type="checkbox"/>	>		Dexamethasone Versus Prednisone for Asthma Treatment in the Pediatric Inpatient Population; a Feasibility Study	NCT03133897	2018-03-05	Phase III	Children'S Hospital Of Eastern Ontario	Link	Abstract

SciWalker
👤 lweber

Dashboard Document Search Chemistry Finder Analytics My Research
Administration Help

Highlight: Matches Annotations Matches (72): << >>

Dexamethasone ↗ Versus Prednisone ↗ for Asthma Treatment in the Pediatric Inpatient Population; a Feasibility Study

Dexamethasone ↗ Versus Prednisone ↗ for Asthma Treatment in the Pediatric Inpatient Population; a Feasibility Study

Basic Data

Trial: [NCT03133897](#)
 Status: Terminated
 Start date: 2018-03-05
 End date: 2020-03-11
 Study type: Interventional
 Clinical phase: Phase III
 Conditions or Disease: Asthma
 Drugs: Dexamethasone ↗; Prednisone ↗
 Sponsors: Children's Hospital of Eastern Ontario

Abstract

Current corticosteroid regimens for children hospitalized with **asthma** typically include a 5-day course of prednisone ↗ or prednisolone ↗. However, these medications **last** poorly and are associated with **vomiting** and poor **compliance**. Outpatient evidence suggests that a 2-day course of dexamethasone ↗ is as effective as a longer course of prednisone ↗ and prednisolone ↗, and better tolerated. Studies in hospitalized patients are lacking. The investigators' primary objective is to determine the feasibility of a non-inferiority trial, comparing 2 days of dexamethasone ↗ to 4 days of inpatient prednisone ↗/prednisolone ↗ for inpatient **asthma** treatment. The investigators also wish to determine the feasibility of 1) enrolling patients upon admission to hospital, 2) asking patients and/or caregivers to complete a symptom weekly for 4 weeks, 3) reassessing patients post hospital **discharge**, 4) successfully completing **phone follow-up** 4 weeks post hospital **discharge**, and 5) collecting health utilization data post hospital **discharge**. This study will inform a future multi-site trial comparing prednisone ↗/prednisolone ↗ to dexamethasone ↗ in inpatient **asthma** treatment. It has the **potential** of improving the delivery of care in **asthma**, by improving **compliance** with a mainstay of treatment. It will also enhance collaboration within Ontario pediatric hospitals, facilitating knowledge translation and standardization of care across institutions.

Population and Eligibility

Whole Enrollment: Actual, planned

Gender: All, Minimal **age**: 18 Months, Maximal **age**: 17 Years, Healthy volunteers: Accepts Healthy Volunteers, Description: Inclusion Criteria: 1. Children admitted during the study **period** with a diagnosis of asthma exacerbation, 2. Children aged 18 months to 17 years of **age** 3. Children who received **oral Corticosteroids** in the Emergency Department prior to admission under our hospital Emergency Department Nursing Medical Directive and/or Pre-Printed Order Form Exclusion Criteria: 1. Children who received IV **Corticosteroid (CS)** in the Emergency Department (ED) 2. Children who received **oral CS** in the ED prescribed in any other way than through the Children's Hospital of Eastern Ontario ED Nursing Medical Directive and/or Pre-Printed Order Form as doses may not be standardized 3. Children who have received more than one **dose of oral CS** prior to enrollment 4. Children who have received their first **dose of oral CS** greater than 12 hours prior to enrollment 5. Children who received **oral or IV CS** in the previous 4 weeks 6. Children with any of the following: unrepaired **congenital heart disease**, **chronic lung diseases** other than **asthma**, severe **neurological impairment** and other significant co-morbid disorders as they are not typically treated on our hospital's **asthma pathway** 7. Children whose caregivers do not understand English or French

Description

My research

Bibliographic data

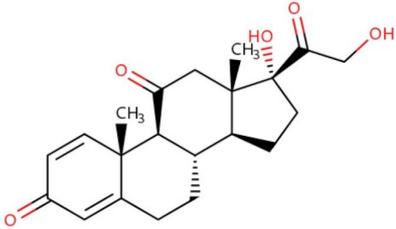
Images (0)

Chemical structures (4)

Export structures: 📄 SMI file

Prednisone

Relevance: 1.2925363 26x << >>



Semantic analysis of millions of scientific documents

Automated chemical structure identification

indexing uses open access ontologies on BigQuery:

35+ ontologies

e.g. chemistry ontology:

35 million concepts

700 million relationships

The screenshot shows the Google Cloud Platform BigQuery interface. The top navigation bar includes 'Google Cloud Platform' and 'SciWalker Open Data'. The main header shows 'BigQuery' with links for 'FEATURES & INFO' and 'SHORTCUT'. On the left, a sidebar contains a 'Resources' section with a search bar and a list of tables: 'chem_ancestors', 'chem_prelabel' (highlighted), 'chemClasses_ancestors', 'chemClasses_prelabel', 'chemGroup_ancestors', and 'chemGroup_prelabel'. The main area is the 'Query editor' for the table 'chem_prelabel'. It includes a 'Query history' sidebar, a 'Query editor' with a 'COMPOSE NEW QUERY' button, and a 'Query editor' with a 'HIDE EDITOR' button and a 'FULL SCREEN' button. Below the editor, there are buttons for 'Run', 'Save query', 'Save view', 'Schedule query', and 'More'. The table name 'chem_prelabel' is displayed with icons for search, share, and copy, and buttons for 'DELETE TABLE' and 'EXPORT'. A message states 'This is a partitioned table. Learn more' with a 'Dismiss' button. The 'Schema' tab is active, showing a table with columns 'Row', 'ocid', and 'name'. The table contains 5 rows of data. At the bottom, there are controls for 'Rows per page' (set to 100) and pagination (1 - 100 of 33030106, with 'First page', '<', '>', and '>| Last page' buttons).

Row	ocid	name
1	150000000341	2,3-xylenols
2	150000000711	D-alloheptuloses
3	150000000798	D-idopyranose derivatives
4	150000000812	D-mannofuranose derivatives
5	150000000813	D-mannofuranoses

Search for concepts

explore very large domain hierarchies, relations

SciWalker iweber

Dashboard Document Search Chemistry Finder Analytics My Research Administration Help

Domain Explorer

diethylene glycol butyl ether

Reset Domains: All (22) Filter

Tree

- Chemistry
 - structural classification
 - compounds
 - acyclic compounds
 - 2-(2-butoxyethoxy)ethanol
 - compounds by functional group
 - hydroxy compounds
 - alcohols
 - primary alcohols
 - 2-(2-butoxyethoxy)ethanol
 - ethers
 - alkyl ethers
 - dialkylethers
 - 2-(2-butoxyethoxy)ethanol
 - compounds by element
 - carbon group compounds
 - carbon compounds
 - organic compounds
 - aliphatic compounds
 - 2-(2-butoxyethoxy)ethanol
 - organic oxygen compounds
 - hydroxy compounds
 - alcohols
 - primary alcohols
 - 2-(2-butoxyethoxy)ethanol
 - ethers
 - alkyl ethers
 - dialkylethers
 - 2-(2-butoxyethoxy)ethanol

Result list

Select: Chemistry

- 2-(2-butoxyethoxy)ethanol

Concept details Deselect concept

Chemistry

2-(2-(2-butoxyethoxy)ethanol)

OCID: 190000916921

| | |
|------------|---|
| InChi | InChi=1S/C8H18O3/c1-2-3-5-10-7-8-11-6-4-9/h9H,2-8H2,1H3 |
| SMILES | CCCCOCCOCCO |
| InChIKey | OAYXUHPQHDHDDZ-UHFFFAOYSA-N |
| Formula | C8H18O3 |
| Mol weight | 162.229 |

Structure

$\text{H}_3\text{C}-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{O}-\text{CH}_2-\text{CH}_2-\text{O}-\text{CH}_2-\text{CH}_2-\text{OH}$

1 concept selected Clear Search in documents

Search for chemical structures and reactions

In publications, patents, clinical trials and grants data and BigQuery tables using a single query

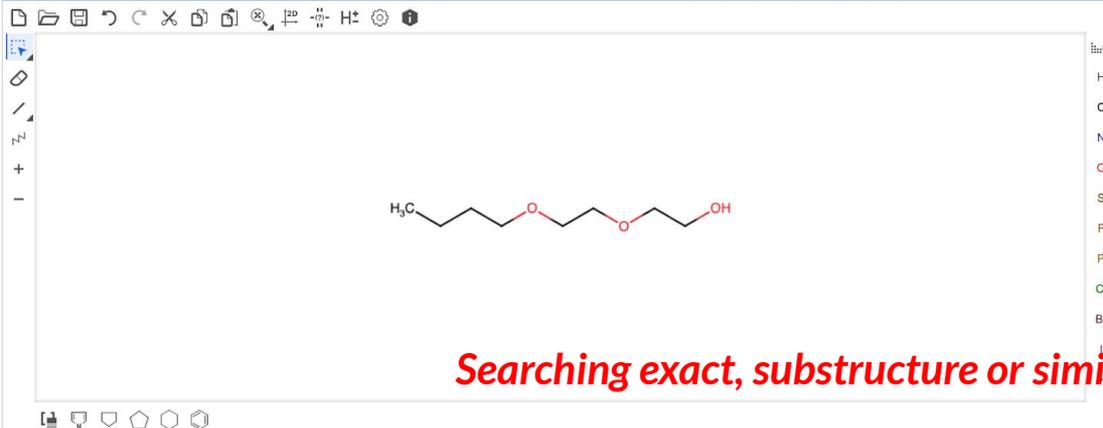
The screenshot shows the SciWalker web interface. At the top, there are navigation tabs: Dashboard, Document Search, Chemistry Finder, Analytics, and My Research. Below this, there are search options: Search compounds in BigQuery, Search reactions in BigQuery, and Search reaction products in SAVI. The main search area is titled "Search by structure" and contains a chemical structure editor with a benzene ring and an oxygen atom. To the right, there is a "Search by attribute (full match)" section with input fields for Name / CAS Registry Number / OCID, SMILES, InChI, and InChI Key. Below the search area, there is a "Sort By" dropdown and a "1 entries" indicator. The results section shows "Reaction 9390" with a chemical reaction diagram: a benzene ring with a hydroxyl group reacts with another benzene ring to form a biphenyl ether. Below the reaction, there is a "Details" section with the following information:

| | | | |
|----------------|---|------------|---|
| Product name | 1E9EBA457D98469A_BA1F62357C9A2B22_7022_UN | RinChi Key | Web-RinChiKey=LYAFAXOQYACMOVJBZ_NUHFFADPSCJSA |
| Product SMILES | C1(=CC=CC=C1)OC2=CC=CC=C2 | RSMI | [C:4]1(=[CH:5][CH:6]=[CH:7][CH:8]=[CH:9])B(O)O.[OH:17][C:18]2=[CH:19][CH:20]=[CH:21][CH:22]=[CH:23]2>=[C:4]1(=[CH:5][CH:6]=[CH:7][CH:8]=[CH:9])O.[17][C:18]2=[CH:19][CH:20]=[CH:21][CH:22]=[CH:23]2 |
| Product OCID | 180212834025 | | |

Additional details include "Chan-Lam coupling" and "Reagents". A red text overlay on the right side of the interface reads: "Searching exact, substructure or similarity on >1 billion SAVI reactions in a BigQuery table: sciwalker-open-data:chemistry_compounds:oc_savi2".

[Search compounds](#)
[Search reactions](#)

Search by structure



Searching exact, substructure or similarity in reactions extracted from US patents

Search type: [duplicate](#)Max. results: [1,000](#)

Search

Search by attribute (full match)

Name / CAS Registry Number / OCID:

SMILES:

InChI:

InChI Key:

Search

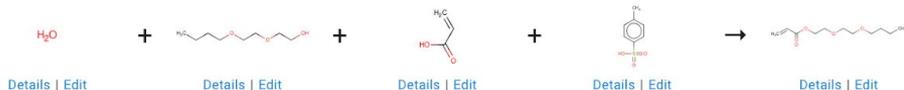
Sort By

31 entries [1](#) [2](#) [3](#) [4](#) [>](#) [>](#)

Reaction 242917

Source: [US-05053554-A](#) | Section: description

Reagents



Details

Source: [US-05053554-A](#) |

Section: description

300.1 grams Dowanol DB (diethylene glycol monobutyl ether, Dow Chemical Company) were azeotropically esterified at 95°-98° C. using nitrogen sparge and vacuum with 158.7 grams acrylic acid in 93.3 grams toluene in the presence of 9 grams p-toluenesulfonic acid and inhibitors until no more water was collected. On completion, the product

RInChI

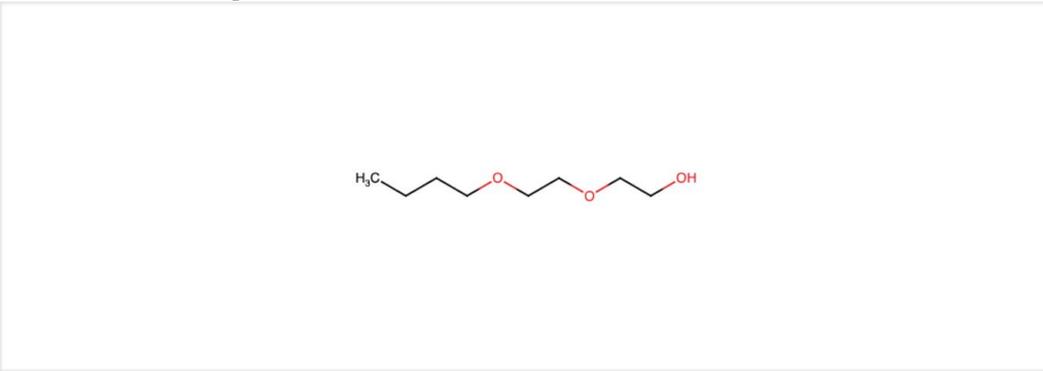
RInChI=1.00.1S/C11H20O4/c1-3-5-6-13-7-8-14-9-10-15-11(12)4-2/h4H,2-3,5-10H2,1H3<<C3H4O2/c1-2-3(4)5/h2H,1H2,(H,4,5)C7H8O3S/c1-6-2-4-7(5-3-6)11(8,9)10/h2-5H,1H3,(H,8,9,10)C8H18O3/c1-2-3-5-10-7-8-11-6-4-9/h9H,2-8H2,1H3H2O/h1H2<>C7H8/c1-7-5-3-2-4-6-7/h2-6H,1H3/d-

RInChI Key

Web-RInChIKey=CUDYCFMPKQBQFJXCX-NUHFFADPSCSJSA

Search compounds Search reactions

Search by structure



CCCCOCCOCCO

Search type: duplicate Max. results: 1,000 Search

Search by attribute (full match)

Name / CAS Registry Number / OCID:

SMILES:

InChI:

InChI Key:

Search

Export SMI

1 entries < 1 >

| Structure | Names | Mol. weight | Formula | External sources | Google Patents+ |
|--|---|-------------|---------|--|--|
| <input type="checkbox"/>  | <ul style="list-style-type: none"> 2-(2-butoxyethoxy)ethanol 112-34-5 2,2'-oxybis-Ethanol monobutyl ether 2-(2-butoxyethoxy)-Ethanol 2-(2-butoxyethoxy)ethan-1-ol ... | 162.229 | C8H18O3 | <ul style="list-style-type: none"> PubChem: 8177 ChEMBL: CHEMBL1904721 EPA DSSTox: DTXCID001519 ZINC15: ZINC01600070 ChEMBL: CHEMBL1904721 ... | <ul style="list-style-type: none"> Exact search: Patents Non-Patent Literature Both Substructure search: Patents Non-Patent Literature Both Similarity search: Patents Non-Patent Literature Both |

[Details](#) | [Edit](#)

1 entries < 1 >

Highlight: Matches Annotations Matches (3): << >>

GENTLE THERAPY TREATMENT- panthenol liquid
SOLEO

Disclaimer: This drug has not been found by FDA to be safe and effective, and this labeling has not been approved by FDA. For further information about unapproved drugs, click here.

Gentle Therapy Treatment**Active ingredients**

Panthenol (0.5%)

Purpose

Hair treatment

Keep out of reach of children

If swallowed, get medical help or contact a Poison Control Center (1-800-222-1222) right away.

Warnings**For external use only**

Do not use when scalp is red, inflamed, irritated, or painful

When using this product

- do not apply on other parts of the body
- avoid contact with eyes. If contact occurs, rinse eyes thoroughly with water
- do not apply directly to wound or open cut

Stop use and ask a doctor if

- rash or irritation on skin develops and lasts
- condition does not improve or worsen after regular use of this product as directed

Directions

Wet hair. Apply the adequate amount on the center of the damaged hair. Rinse with warm water after 3-5 minutes.

Inactive ingredients

Water, Stearyl Alcohol, Cetyl Alcohol, Behentrimonium Chloride, Propylene Glycol, Amodimethicone, Steartrimonium Chloride, Caprylyl Glycol, 1,2-Hexanediol, Glycerin, Betaine, Chlorphenesin, Fragrance, Trideceth-6, Octoxynol-40, Illicium Verum (Anise) Fruit Extract, Caprylic, Capric Triglyceride, Disodium EDTA, Butoxydiglycol, Trideceth-3, Dimer Dilinoleyl Dimer Dilinoleate, Myristyl Alcohol, Arachidyl Alcohol, Salvia Officinalis (Sage) Extract, Simmondsia Chinensis (Jojoba) Seed Oil, Macadamia Integrifolia Seed Oil, Olea Europaea (Olive) Fruit Oil, Lavandula Angustifolia (Lavender) Extract, Chamomilla Recutita (Matricaria) Flower Extract, Rosmarinus Officinalis (Rosemary) Extract, Cymbopogon Schoenanthus Extract, Citric acid

Uses

Helps provide nutrition and moisturizes dry hair.

Gentle Therapy Treatment

> My research

> Bibliographic data

> Images (0)

Chemical structures (31)

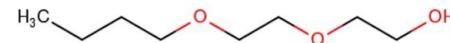
Export structures:

SMI file

2-(2-butoxyethoxy)ethanol

Relevance: 0.39737958

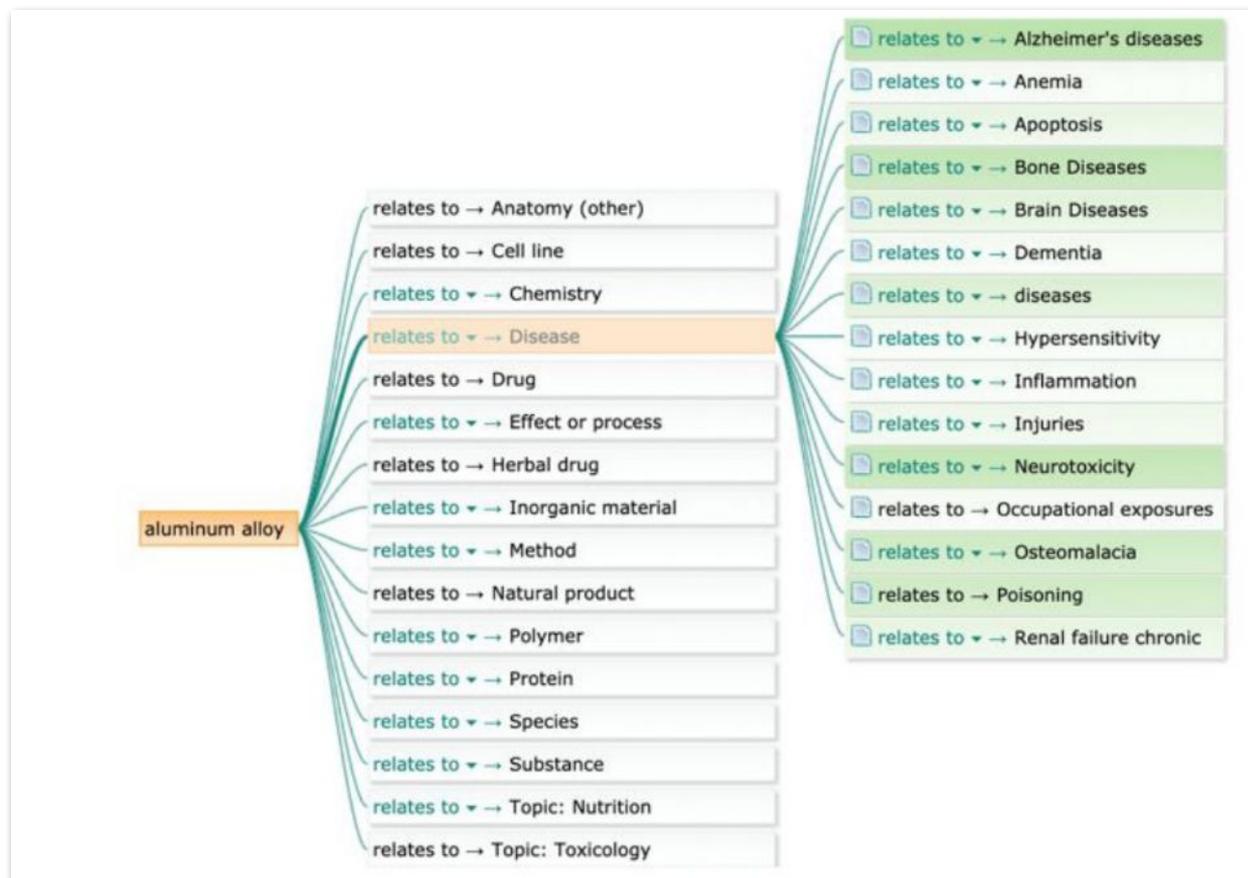
3x << >>

**Arachidyl alcohol**

Relevance: 0.25856942

3x << >>

Identify relationships between concepts such as aluminum alloy and related diseases



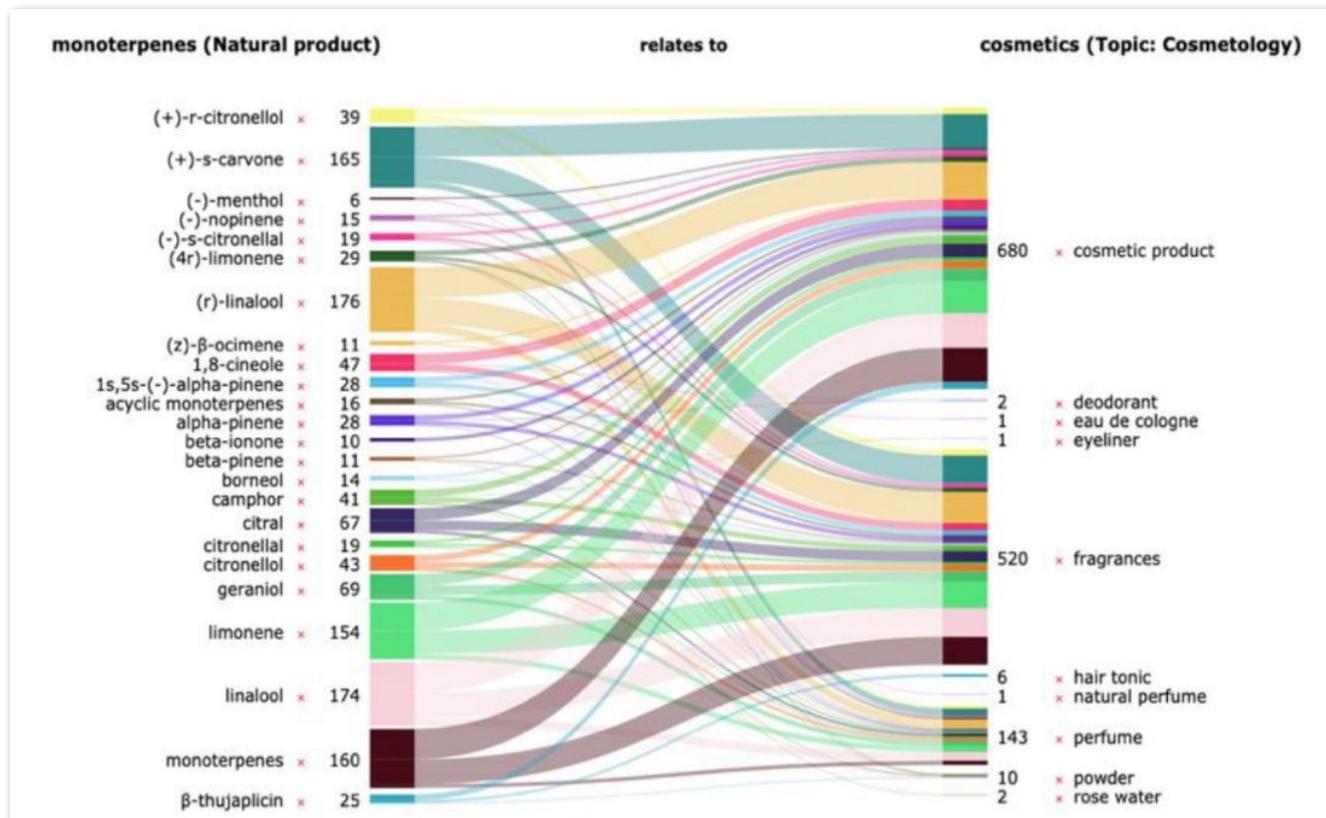
Evidence is key!

The system delivers sentences extracted from documents describing the relation and links to the source documents

The screenshot displays the SciWalker interface. On the left, a search box contains the text "aluminum alloy". To its right, a list of related terms is shown, including "Protein", "Species", "Substance", "Topic: Nutrition", and "Topic: Toxicology". The main window, titled "Sentences", shows a search result for "aluminum alloy relates to Neurotoxicity". Below this, a template "[compound] induces [disease]" is displayed. The main content area contains several extracted sentences, such as "The findings could possibly lead to studies explaining the biochemical basis of aluminium induced neurotoxicity as well as the side effects associated with the long term medication of neuropsychotropic drugs." and "Therefore, the oral ingestion of aluminum induced neurotoxicity in mice which may be seen only at an early age, but injection of aluminum can cause neurotoxicity at any age." On the right side of the interface, a vertical list of related diseases is shown, including "Alzheimer's diseases", "Anemia", "Apoptosis", "Bone Diseases", "Brain Diseases", "Dementia", "diseases", "Hypersensitivity", "Inflammation", "Injuries", "Neurotoxicity", "occupational exposures", "Osteomalacia", "poisoning", and "Renal failure chronic".

Clever visualizations make complex things easy!

Substantial relationships between monoterpenes and cosmetics.



SciWalker

lweber

Dashboard Document Search Chemistry Finder Analytics My Research Administration Help

Compound Exporter

SciWalker uses sophisticated semantic technologies to detect both known and novel chemical compounds in document texts and also in images. You can use [My Library](#) to create a collection of documents that you want to export all chemical compounds from. Just add documents to an existing collection or create a new one. You can choose the desired output format.

| Collection name | Description | Documents | Created | Last modified | Sharing |
|--|------------------------------------|-----------|-----------------------|-----------------------|---------|
| <input type="text"/> | <input type="text"/> | | | | |
| <input checked="" type="radio"/> dowtest | dow test | 7 | 2020-11-09 10:20:00.0 | 2020-11-09 10:20:05.0 | |
| <input type="radio"/> Lutz's documents | Lutz's default document collection | 2 | 2020-07-03 08:33:12.0 | 2020-07-06 07:44:31.0 | |

2 collections | < 1 > | 10

Export format: XLSX SMI

powered by [BLAST](#) | [CDK](#) | [ChemAxon](#) | [OpenChemLib](#) | [OPSIN](#) | [OSRA](#) Bildschirmfoto [Legal notice](#) © 2020

Automation of customer specific research processes:

Example: continuous biomarker screening for Boehringer Ingelheim

The screenshot displays the Mart Biomarker search interface. At the top, there are logos for Mart Biomarker, Boehringer Ingelheim, S.i.C. Scientific Information Center, and ontochem. The search bar contains the query "EGFR T790M" and "Try EGFR". Below the search bar, there are filters for Biomarker, Biomarker target, Drug, Gene variant, Population, and Type. The results section shows a table with 28 records found, sorted by Biomarker in ascending order. The table columns include Details, Biomarker, Gene variant, Target, Drugs, Type, Anatomy, Population, Kit, Sentence, Repository, Source, Source date, and Document view. The first four rows of the table are visible, showing results for EGFR T790M and EGFR.

| Details | Biomarker | Gene variant | Target | Drugs | Type | Anatomy | Population | Kit | Sentence | Repository | Source | Source date | Document view |
|---------|-----------|--------------|------------|-------|---------------------------|---------|------------|-----|---|----------------|-----------------------------------|-------------|---------------|
| > | | | EGFR T790M | | Rate of response | | patients | | These studies reported >50 % ... | PubMed Central | PMC4488453 | 2015-06-30 | Show |
| > | | | EGFR T790M | | Progression-free survival | | | | Almost all EGFR-positive NSCLC patients ... | PubMed Central | PMC6200532 | 2018-10-05 | Show |
| > | EGFR | EGFR T790M | | | Biomarker | | patients | | This has particular importance given ... | Patents | WO-2014201092; A1 | 20141218 | Show |
| > | EGFR | EGFR T790M | | | Biomarker | | | | Examples of such successes include ... | PubMed Central | PMC4574608 | 2015-09-17 | Show |

 **BoN labdocs** iweber

Dashboard Document Finder Chemistry Finder Analytics My Research Admin User Help

Simple search Advanced search Expert search Simple Search Tutorial

Save query Send to advanced search Search

LABDOCS
14 554

14 554 hits

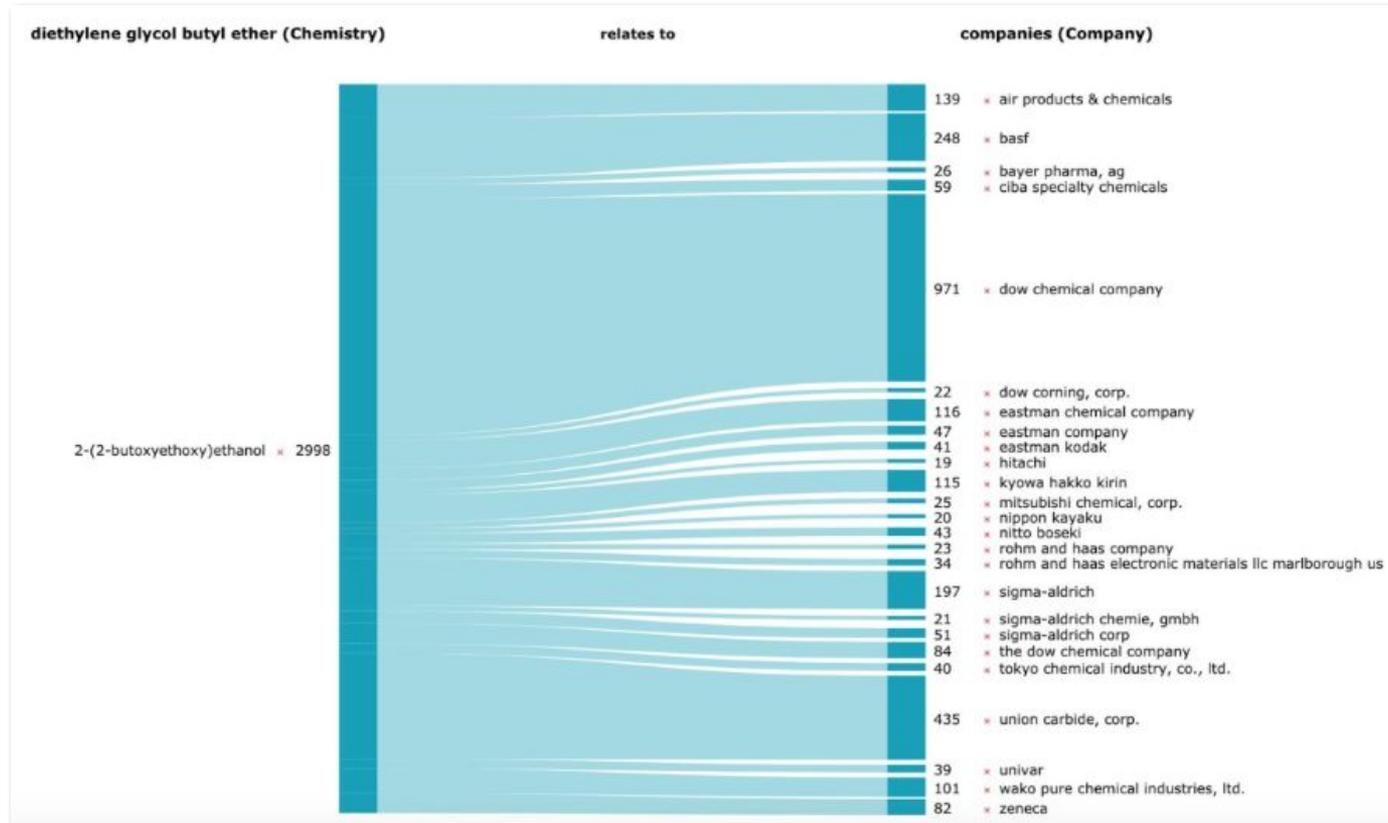
| Author | Chemistry (domain) | Companies (domain) | Drugs (domain) | Species (domain) | Substances (domain) |
|--------|--------------------|--------------------|----------------|------------------|---------------------|
|--------|--------------------|--------------------|----------------|------------------|---------------------|

Select columns to display 14 554 documents found Export XLSX

| | Details | Read | Title |
|--------------------------|---------|------|----------------------------------|
| <input type="checkbox"/> | > | | KAU037997.10.esp |
| <input type="checkbox"/> | > | | KAU037594.10.jdx |
| <input type="checkbox"/> | > | | KAU037579.10.jdx |

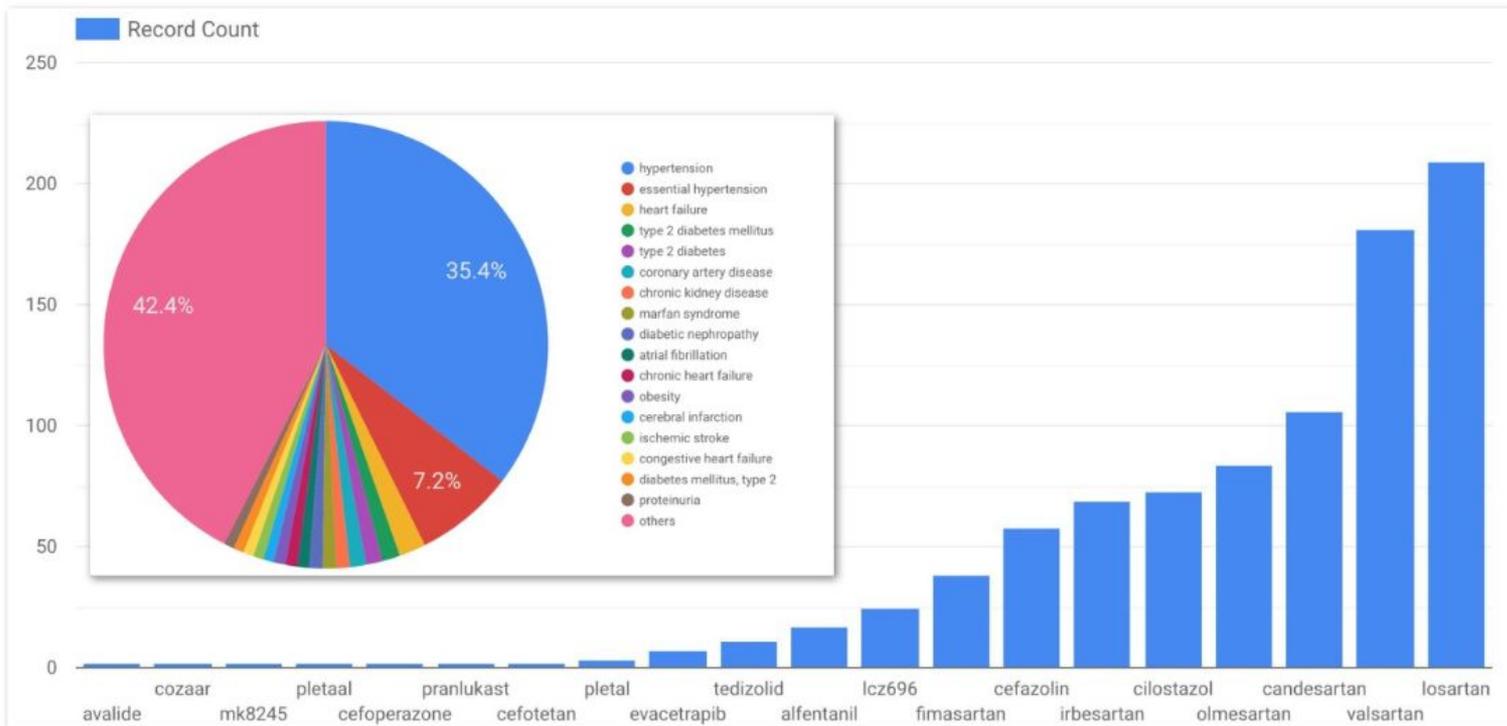
Relationships identified in patents and publications

between a chemical compound and the high-level concept “company”



Q: What tetrazole containing drugs are in how many clinical trials for which diseases ?

A: from 303.973 published tetrazoles, 35 compounds were in 926 trials (using Google DataStudio):



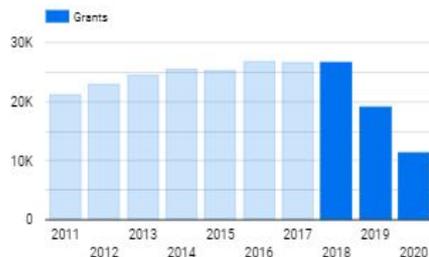
Complex technologies applied in the background

Resulting in a simple dashboard

Still enabling the interested user to interactively drill down and explore deeper insights.

Which biological targets are in the focus of current research grant applications?

Grants by start year



| | Protein name | Grants | Funding total (USD) |
|-----|--------------|--------|---------------------|
| 1. | PD1L1 family | 747 | \$193,275,507 |
| 2. | MAPK2 family | 630 | \$114,068,130 |
| 3. | MAPK family | 612 | \$108,490,877 |
| 4. | SPK1 family | 611 | \$108,369,006 |
| 5. | ERKA family | 609 | \$108,288,822 |
| 6. | NLRP3 family | 602 | \$94,401,194 |
| 7. | NF2L2 family | 571 | \$105,760,382 |
| 8. | GPBAR family | 494 | \$182,109,110 |
| 9. | OXER1 family | 458 | \$173,224,501 |
| 10. | AKT1 family | 457 | \$69,765,075 |
| 11. | GCR2 family | 453 | \$171,704,959 |
| 12. | TLR4 family | 435 | \$105,726,983 |
| 13. | AURKB family | 429 | \$252,306,051 |
| 14. | S45A2 family | 413 | \$248,404,771 |
| 15. | MBD2 family | 331 | \$88,717,985 |
| 16. | MYC family | 329 | \$88,001,168 |
| 17. | ALBU family | 308 | \$50,790,689 |
| 18. | HIF1A family | 307 | \$36,588,830 |
| 19. | APOE family | 290 | \$271,569,831 |
| 20. | IFNG family | 286 | \$63,006,600 |



Customizing SciWalker

SciWalker is a highly modular - create your own personalized applications !

Modular Functions - Personalized Applications - "iMarts"

Organization details

Organization name: Status: ● Active

Application name: Max. users:

Repositories
Text or database sources available

- All
- CORE (OC, oacore)
- ChEMBL (OC, chembl)
- Clinical Trials (AACT) (OC, ctaact)
- Clinical Trials (EudraCT) (OC, cteuct)
- Clinical Trials (ICTRP) (OC, ctctrp)
- Clinical Trials (Trialtrove) (OC, ctoteline)
- Clinical trials (OC, aact)
- Drug labels (OC, druglabels)
- Grants EU (OC, eugrants)
- Grants NIH (OC, nihgrants)
- Journals (OC, journals)
- Medline (OC, med)
- Patents (OC, patents)
- Patents DOCDB (OC, patdocdb)
- PubMed Central (OC, pmc)
- Web (OC, web)
- Web Regenold (OC, webregenold)

Features
GUI features and apps available

- All
- GENERAL**
 - Changelog
 - Dashboard
 - Document Search (Landing page)
 - Info Page
 - Manual
 - User Management
- CHEMISTRY**
 - Chemistry Search
 - Compound Search
 - Semantic Export
 - Sequence Search
- ANALYTICS**
 - Domain Explorer
 - Path Search
- USER RESEARCH**

Domains highlighting
Domains highlighted in annotated text

- All
- Anatomy
- Biomarker
- Cell lines
- Chemical Classes
- Chemical Compounds
- Chemical Groups
- Chemistry
- Clinical study dictionary
- Companies
- Cosmetology
- Countries
- Devices
- Diabetes health
- Diseases
- Drugs
- Effects, processes and functions
- Eye health

Searchable domains
Domains available in autocomplete

- All
- Anatomy
- Biomarker
- Cell lines
- Chemical Classes
- Chemical Compounds
- Chemical Groups
- Chemistry
- Clinical study dictionary
- Companies
- Cosmetology
- Countries
- Devices
- Diabetes health
- Diseases
- Drugs
- Effects, processes and functions
- Eye health

Delete organization Cancel Save

SciWalker
👤 Iweber

🏠 Dashboard 🔍 Document Search 🧪 Chemistry Finder 📊 Analytics ★ My Research
🏠 Administration 🔗 Help

My library

[Document search](#)
[Compound exporter](#)

Document collections
Documents

Filter collections

| Alert | Collection name | Description | Documents | Created | Last modified | Is shared | Shared for | Editable for shared users | Modify |
|----------------------------|-------------------------------------|------------------------------------|-----------|---------------------|---------------------|-----------|------------|---------------------------|--------|
| <input type="checkbox"/> | <input type="text"/> | <input type="text"/> | | | | | | | |
| <input type="checkbox"/> 🔔 | dowtest | dow test | 6 | 9.11.2020, 11:17:39 | 9.11.2020, 11:18:07 | 👤 🔒 | - | no | ✏️ |
| <input type="checkbox"/> 🔔 | test | test | 6 | 3.11.2020, 10:26:53 | 3.11.2020, 10:26:58 | 👤 🔒 | - | no | ✏️ |
| <input type="checkbox"/> 🔔 | monoterpenes+virus | monoterpenes+virus | 15 | 2.11.2020, 15:26:36 | 2.11.2020, 15:26:42 | 👤 🔒 | - | no | ✏️ |
| <input type="checkbox"/> 🔔 | niclosamide | niclosamide | 0 | 26.8.2020, 16:38:33 | 26.8.2020, 16:38:33 | 👤 🔒 | - | no | ✏️ |
| <input type="checkbox"/> 🔔 | PDquery6 | | 11 | 20.7.2020, 19:46:31 | 20.7.2020, 19:46:36 | 👤 🔒 | - | no | ✏️ |
| <input type="checkbox"/> 🔔 | ProteinDegradation1 | | 20 | 20.7.2020, 19:35:49 | 20.7.2020, 19:40:58 | 👤 🔒 | - | no | ✏️ |
| <input type="checkbox"/> 🔔 | Lutz's documents | Lutz's default document collection | 8 | 13.5.2020, 11:19:57 | 16.6.2020, 20:05:08 | 👤 🔒 | - | no | ✏️ |

7 collections
⏪ 1 ⏩
10

+ Create new collection

powered by [BLAST](#) | [CDK](#) | [ChemAxon](#) | [OpenChemLib](#) | [OPSIN](#) | [OSRA](#)

[Legal notice](#)
© 2020

Watchlist alerts Library alerts

| Status | Name | Description | Type | Interval | Next sending date | Created | Remove |
|--------|------|-------------|-------|----------|-------------------|-----------------------|--------|
| | dow1 | | query | WEEKLY | 2020-11-16 | 2020-11-09 10:27:52.0 | |

1 alert < 1 > > 10

SciWalker

lweber

Dashboard | Document Search | Chemistry Finder | Analytics | My Research | Administration | Help

Dashboard

New documents since: 2020-11-02

dow 574 new

| Category | Count |
|----------|-------|
| dow1 | ~40 |
| MOF | ~120 |
| aspirin | ~310 |

Bortezomib 123 new

| Category | Count |
|------------|-------|
| Category 1 | ~10 |
| Category 2 | ~60 |
| Category 3 | ~50 |

targeted protein degradation 8 new

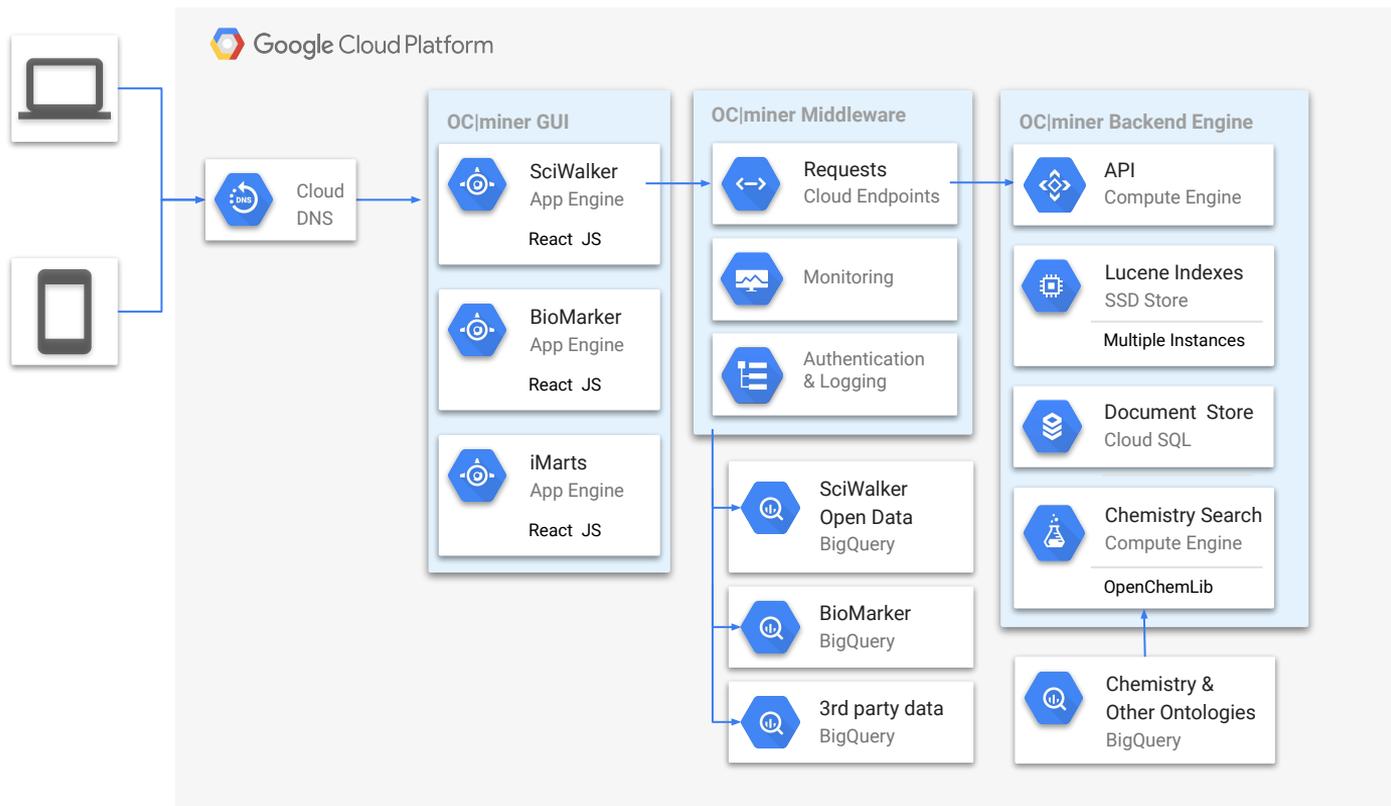
| Category | Count |
|------------|-------|
| Category 1 | ~1 |
| Category 2 | ~3 |
| Category 3 | ~4 |

powered by [BLAST](#) | [CDK](#) | [ChemAxon](#) | [OpenChemLib](#) | [OPSIN](#) | [OSRA](#)

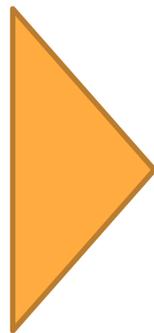
[Legal notice](#) | © 2020

SciWalker is a toolkit:

- RESTful WebAPI, programmed in Java, based on non-toxic open access software
- cloud based installations, e.g. via docker images: GCP, AWS, Azure
- local installations, e.g. on laptops - Unix, Mac, PC
- using encrypted, authenticated communications with RESTful WebAPI
- React Javascript GUI and Java 8 middleware
- Chemistry webservice: OpenChemLib, CDK or ChemAxon, different editors
- MySQL as general document database
- connectors to other databases, and 3rd party WebAPI's, e.g. Google BigQuery
- hardware: typically 32GB RAM, 8+ CPU, 12 TB HDD or SSD for large indexes



New technologies
can be successful
only if
they are easy to use



OntoChem is using the
Google BigQuery & GCP
platform

to present complex research
data & results in an easy way

providing outstanding
functionality



Thank you

contact:

www.ontochem.com

phone: +49 345 4780470

e-mail: info@ontochem.com