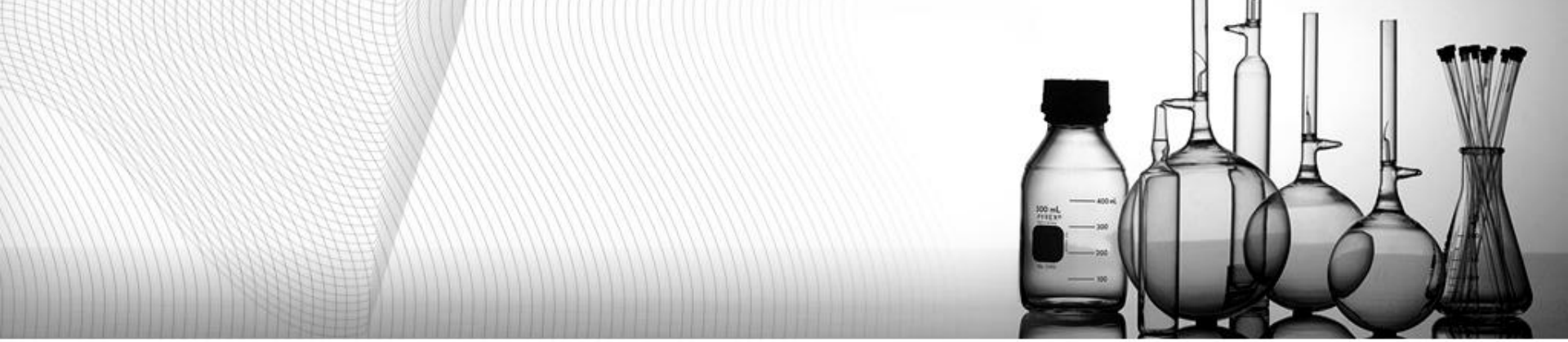


Open PHACTS
Open Pharmacological Space

A Data Platform for Drug Discovery

Paul Groth (@pgroth)
http://www.few.vu.nl/~pgroth

VU VRIJE UNIVERSITEIT AMSTERDAM

Open PHACTS
Open Pharmacological Space

1. WHY
2. THE PLATFORM
3. APPS
4. THE FUTURE

Open PHACTS

Open Pharmacological Space

Pfizer

universität wien

RSC Advancing the Chemical Sciences

130 jaar VU Vrije Universiteit Amsterdam

AstraZeneca

gsk GlaxoSmithKline

DTU

UH Universität Hamburg

cnio

MANCHESTER 1824

ESTEVE closer to you

NOVARTIS

Bio Solve IT

Parc de Salut MAR

Maastricht University

AQnowledge Semantics for Science

MERCK

Lundbeck

USC

universitätbonn

Lilly

LUMC

janssen

EMBL-EBI

connecteddiscovery

nbic

OPENLINK SOFTWARE Making Technology Work For You®

SIB Swiss Institute of Bioinformatics

The Open PHACTS Foundation

imi

2 billion Euro

1 billion Euro

Public

Partnership

1 billion Euro

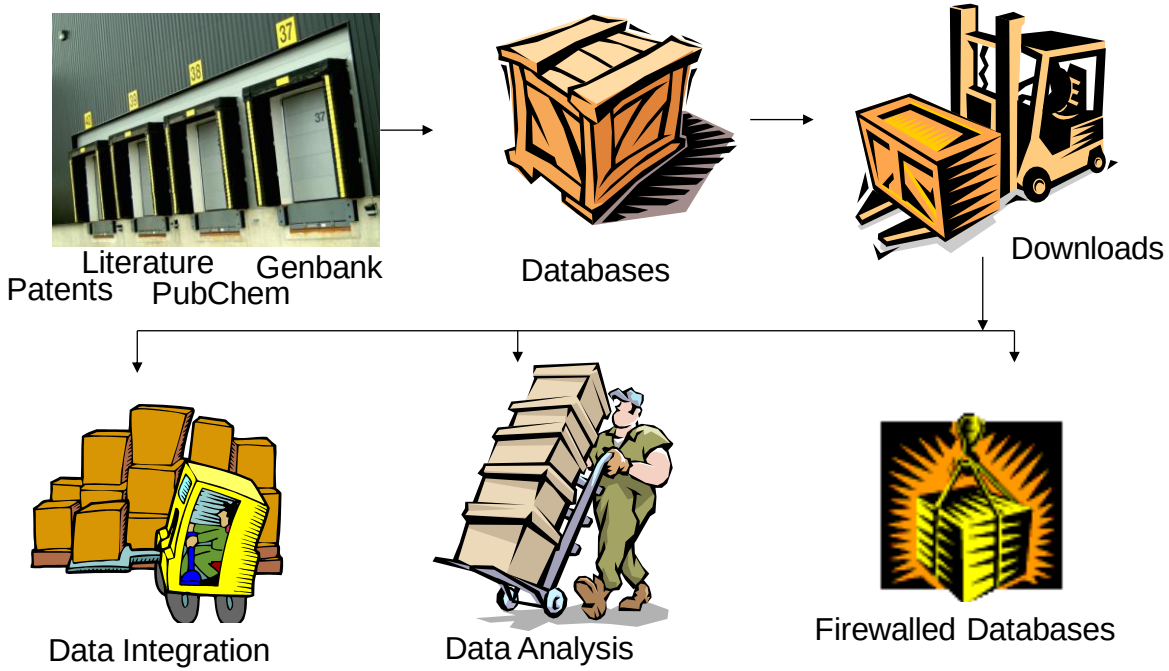
efpia

Private

Open PHACTS
Open Pharmacological Space

Pre-competitive Informatics:

Pharma are all accessing, processing, storing & re-processing external research data



X Repeat @ each company

Lowering industry firewalls: pre-competitive informatics in drug discovery
Nature Reviews Drug Discovery (2009) 8, 701-708 doi:10.1038/nrd2944

Open PHACTS
Open Pharmacological Space

Business Question Driven Approach

Number	sum	Nr of 1	Question
15	12	9	All oxidoreductase inhibitors active <100nM in both human and mouse
18	14	8	Given compound X, what is its predicted secondary pharmacology? What are the on and off,target safety concerns for a compound? What is the evidence and how reliable is that evidence (journal impact factor, KOL) for findings associated with a compound?
24	13	8	Given a target find me all actives against that target. Find/predict polypharmacology of actives. Determine ADMET profile of actives.
32	13	8	For a given interaction profile, give me compounds similar to it.
37	13	8	The current Factor Xa lead series is characterised by substructure X. Retrieve all bioactivity data in serine protease assays for molecules that contain substructure X.
38	13	8	Retrieve all experimental and cli structure (with options to match
41	13	8	A project is considering Protein compounds known to modulate the target directly? i.e. return all level of the target family (i.e. PK
44	13	8	Give me all active compounds o
46	13	8	Give me the compound(s) which (disease)
59	14	8	Identify all known protein-protei

ELSEVIER

Drug Discovery Today

Volume 18, Issues 17–18, September 2013, Pages 843–852

Review

Scientific competency questions as the basis for semantically enriched open pharmacological space development

Kamal Azzaoui[1], Edgar Jacoby[14], Stefan Senger[2], Emiliano Cuadrado Rodríguez[3], Mabel Loza[3], Barbara Zdrazil[4], Marta Pinto[4], Antony J. Williams[5], Victor de la Torre[6], Jordi Mestres[7], Manuel Pastor[7], Olivier Taboureau[8], Matthias Rarey[9], Christine Chichester[10], Steve Pettifer[11], Niklas Blomberg[12, a], Lee Harland[13], Bryn Williams-Jones[13], Gerhard F. Ecker[4]

http://www.sciencedirect.com/science/article/pii/S1359644613001542

Open PHACTS

Open Pharmacological Space

"What is the selectivity profile of known p38 inhibitors?"

"Let me compare MW, logP and PSA for known oxidoreductase inhibitors"

"Find me compounds that inhibit targets in NFkB pathway assayed in only functional assays with a potency <1 μM"

ChEMBL

DrugBank

Gene Ontology

Wikipathways

GeneGo

ChEBI

UniProt

UMLS

GVKBio

ConceptWiki

ChemSpider

TrialTrove

TR Integrity

Open PHACTS

Open Pharmacological Space

- Project Start
- Tech Team Kickoff
- Drive Team In-Place
- 1st Hackathon
- GUI Hackathon
- Hackathon 2
- Hackathon 3
- Focused User Feedback
- Alpha for public release
- Platform On Hosting Provider
- Revised Platform Implementation
- Hosting Selected
- Hosting Partner On-board
- Prototype
- Linked Data API
- Usathon
- Public Release

OPENLINK SOFTWARE
Making Technology Work For You!®

Open Phacts

												2012												2013				
b	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May

Open PHACTS
Open Pharmacological Space

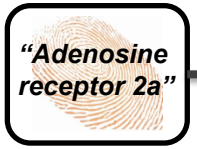
THE OPEN PHACTS DISCOVERY PLATFORM

Open PHACTS
Open Pharmacological Space

Apps

Core Platform

Identity Resolution Service



Identifier Management Service

P12374
EC2.43.4
CS4532

Linked Data API (RDF/XML, TTL, JSON)

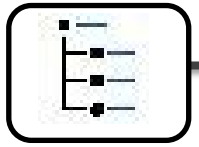
Semantic Workflow Engine

Data Cache
(Virtuoso Triple Store)

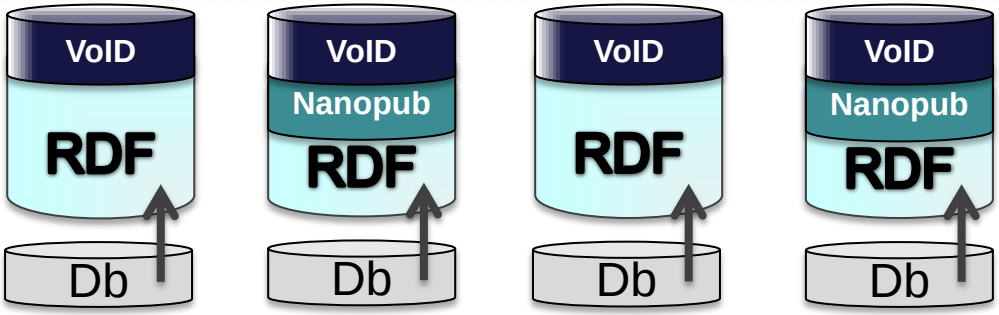
Domain Specific Services

Chemistry
Registration
Normalisation
& Q/C

Indexing



Public Ontologies



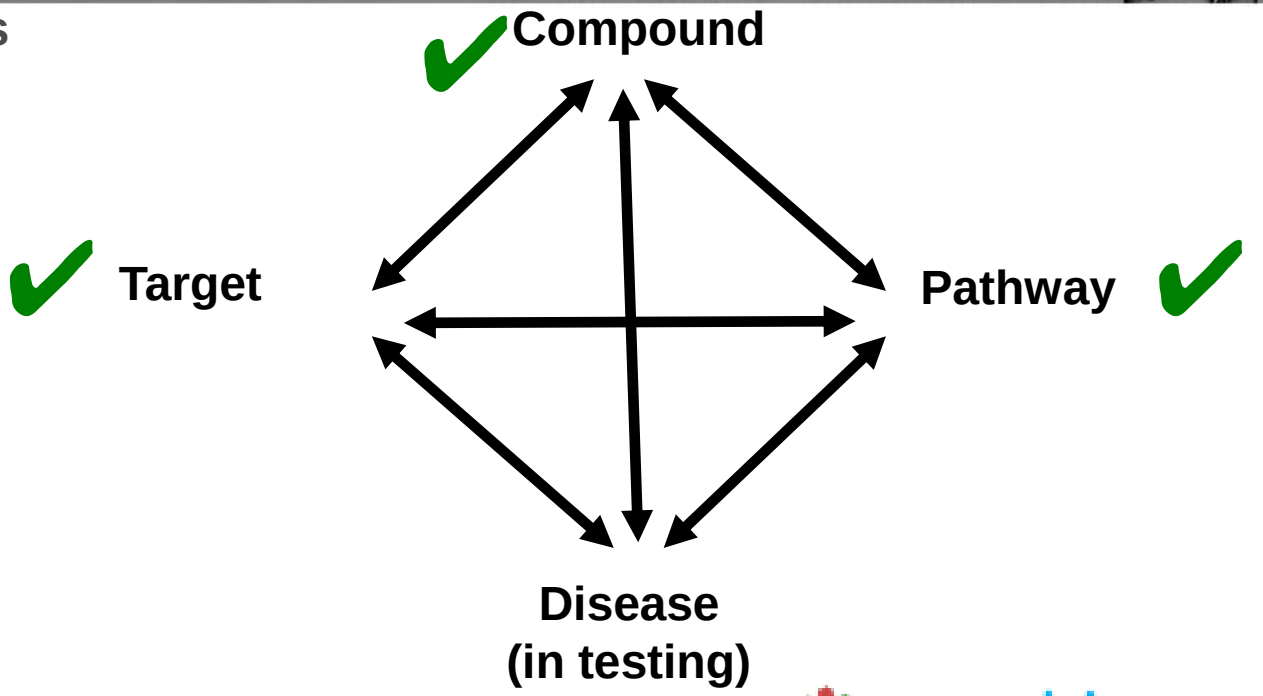
Public Content

Commercial

User Annotations

Open PHACTS
Open Pharmacological Space

Data Sources



Concept Wiki
ChEBI
ChEMBL
ChemSpider
The free chemical database
swissprot
DrugBank
Open Data Drug & Drug Target Database
UniProt
WikiPathways
Pathways for the People
GENEONTOLOGY
Unifying Biology

Play!

https://dev.openphacts.org/

OpenPHACTS API

Chemical Structure Exact Search	/structure/exact	GET
InchiKey to URL	/structure	GET
Inchi to URL	/structure	GET
Chemical Structure Similarity Search	/structure/similarity	GET
SMILES to URL	/structure	GET
Chemical Structure Substructure Search	/structure/substructure	GET
Get concept description	/getConceptDescription	GET
Map free text to a concept URL based on semantic tag	/search/byTag	GET
Map URL	/mapURL	GET
Map free text to a concept URL	/search/freetext	GET
Get ChEBI Ontology Class Members	/compound/chebi/members	GET
Get ChEBI Ontology Root Classes	/compound/chebi/root	GET
Get ChEBI Ontology Class	/compound/chebi/node	GET
ChEBI Class Pharmacology Count	/compound/chebi/pharmacology/count	GET

PARAMETER	VALUE	DESCRIPTION
app_id		Your access application id
app_key		Your access application key
searchOptions.Molecule	(required)	A SMILES string. E.g. CC(=O)Oc1ccccc1C(=O)O
searchOptions.SimilarityType		0: Tanimoto ; 1: Tversky ; 2: Euclidian
searchOptions.Threshold		Double <= 1.0
commonOptions.Complexity		(Not supported at the moment) 0: Any ; 1: Single ; 2: Multi
commonOptions.Isotopic		(Not supported at the moment) 0: Any ; 1: Labeled ; 2: NotLabeled
commonOptions.HasSpectra		(Not supported at the moment) Boolean
commonOptions.HasPatents		(Not supported at the moment) Boolean
resultOptions.Limit		Integer. Search limit. Specefy how many results return back during the search. Default value: -1 .
resultOptions.Start		Integer. Return results starting the index. Default value: 0
resultOptions.Length		Integer. How many results should be returned starting from Start index. Default value: -1.

Open PHACTS
Open Pharmacological Space

Secure Cloud Hosted + Virtualized

Triple Store
- Virtuoso 7 column store
- Scale to > 100 billion triples

Network
- AMX-IS
- Extensive memcache
- Monitored

Hardware (development)
- 2 x Intel Xeon E5-2640 - 384 GB DDR3 1333MHz RAM - 1.5 TB SSD - 3TB 7200rpm

OPENLINK SOFTWARE
Making Technology Work For You®

Open PHACTS
Open Pharmacological Space

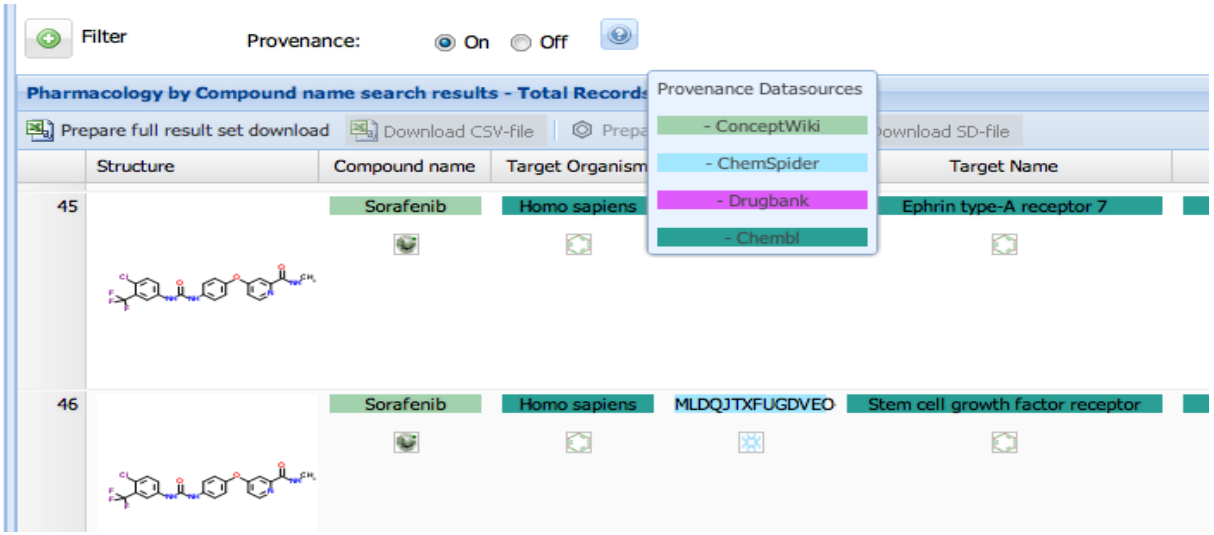
Dealing With The *Really* Tough Parts

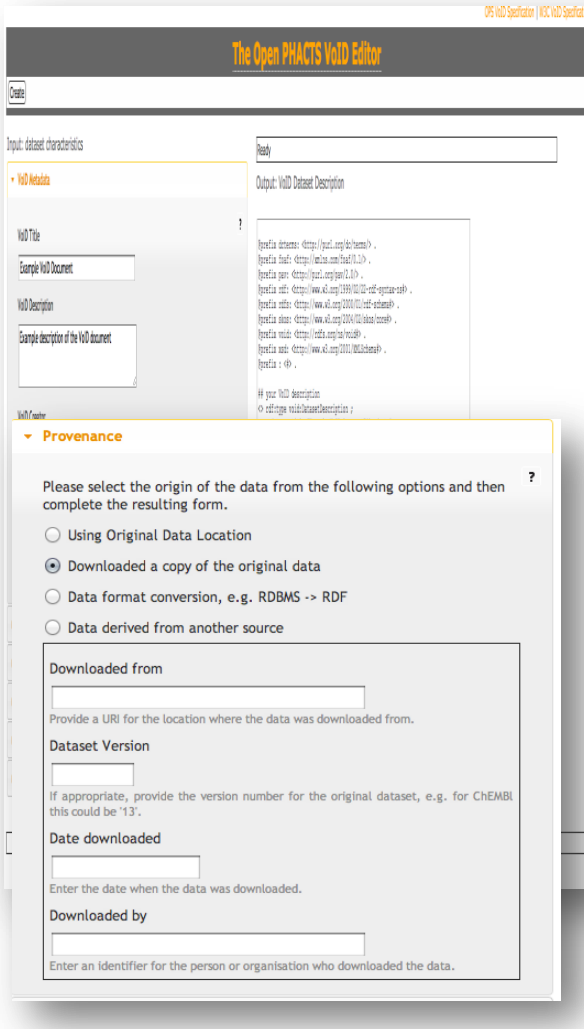
Data Licensing

John
Wilbanks
http://del-fi.org/

Compatibility chart		Terms that may be used for a derivative work or adaptation						
		BY	BY-NC	BY-NC-ND	BY-NC-SA	BY-ND	BY-SA	PD
Status of original work	PD	■	■	■	■	■	■	■
	BY	■	■	■	■	■	■	
	BY-NC		■	■	■			
	BY-NC-ND							
	BY-NC-SA				■			
	BY-ND							
	BY-SA						■	

Open PHACTS
Open Pharmacological Space

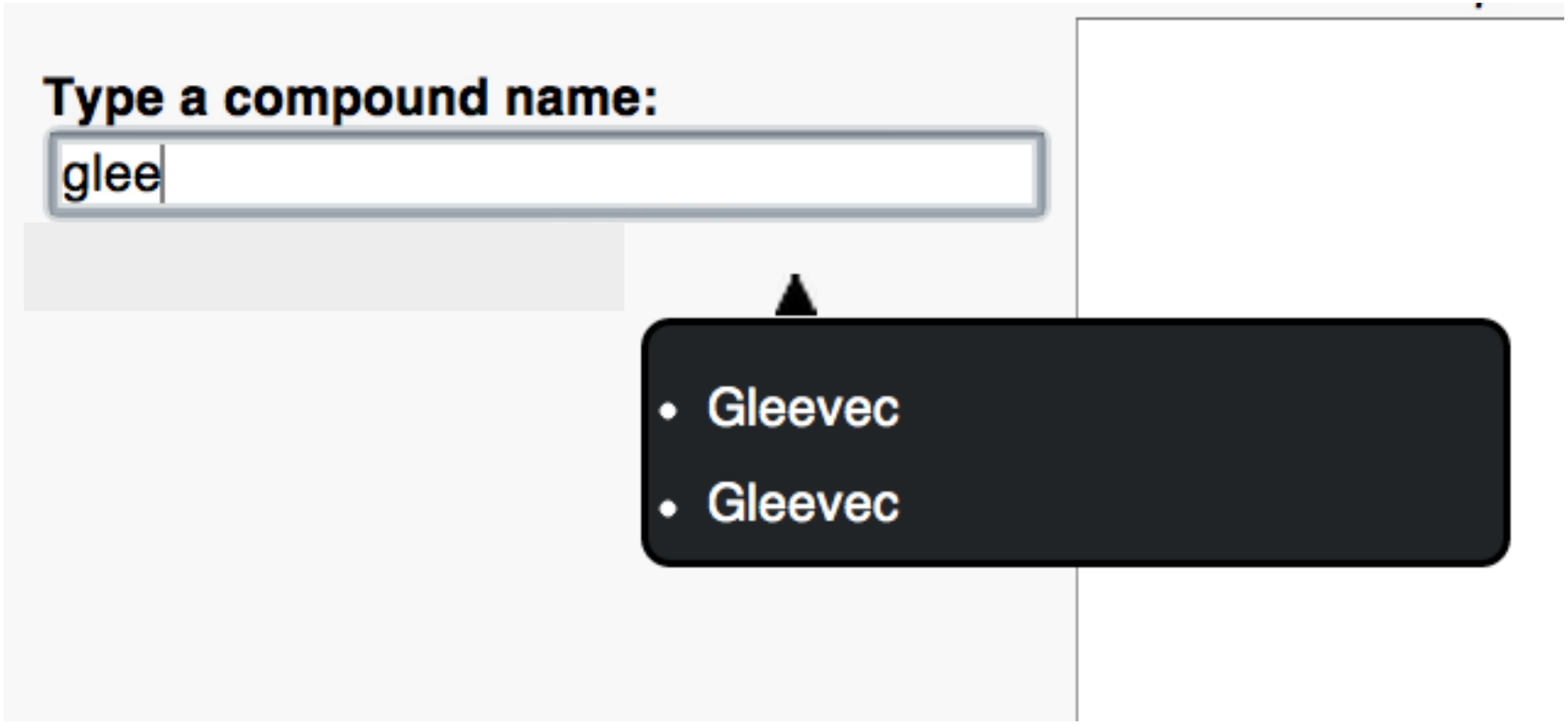




Provenance everywhere

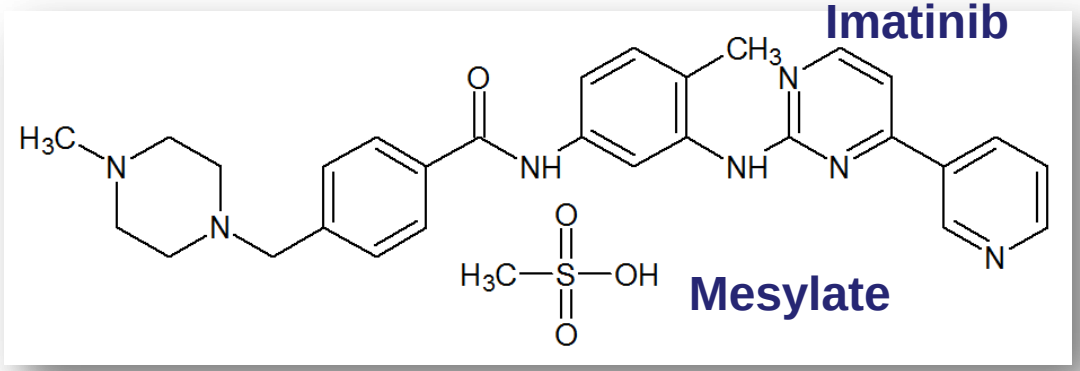
Open PHACTS
Open Pharmacological Space

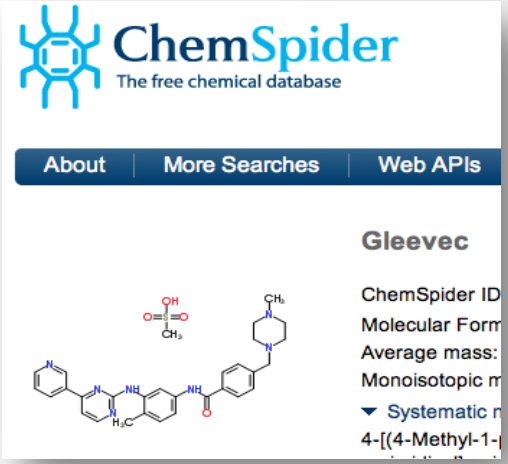
Its easy to integrate, difficult to integrate well:



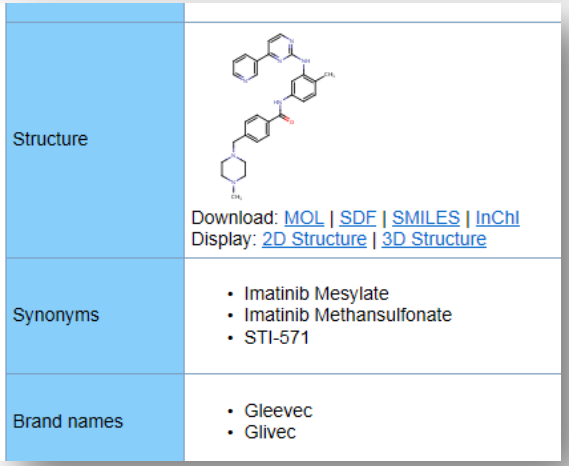
Open PHACTS
Open Pharmacological Space

What Is Gleevec?

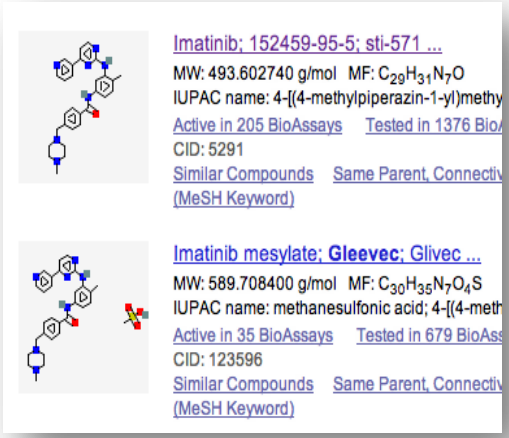




ChemSpider



Drugbank



PubChem

Open PHACTS
Open Pharmacological Space

Dynamic Equality

chemspider:gleevec

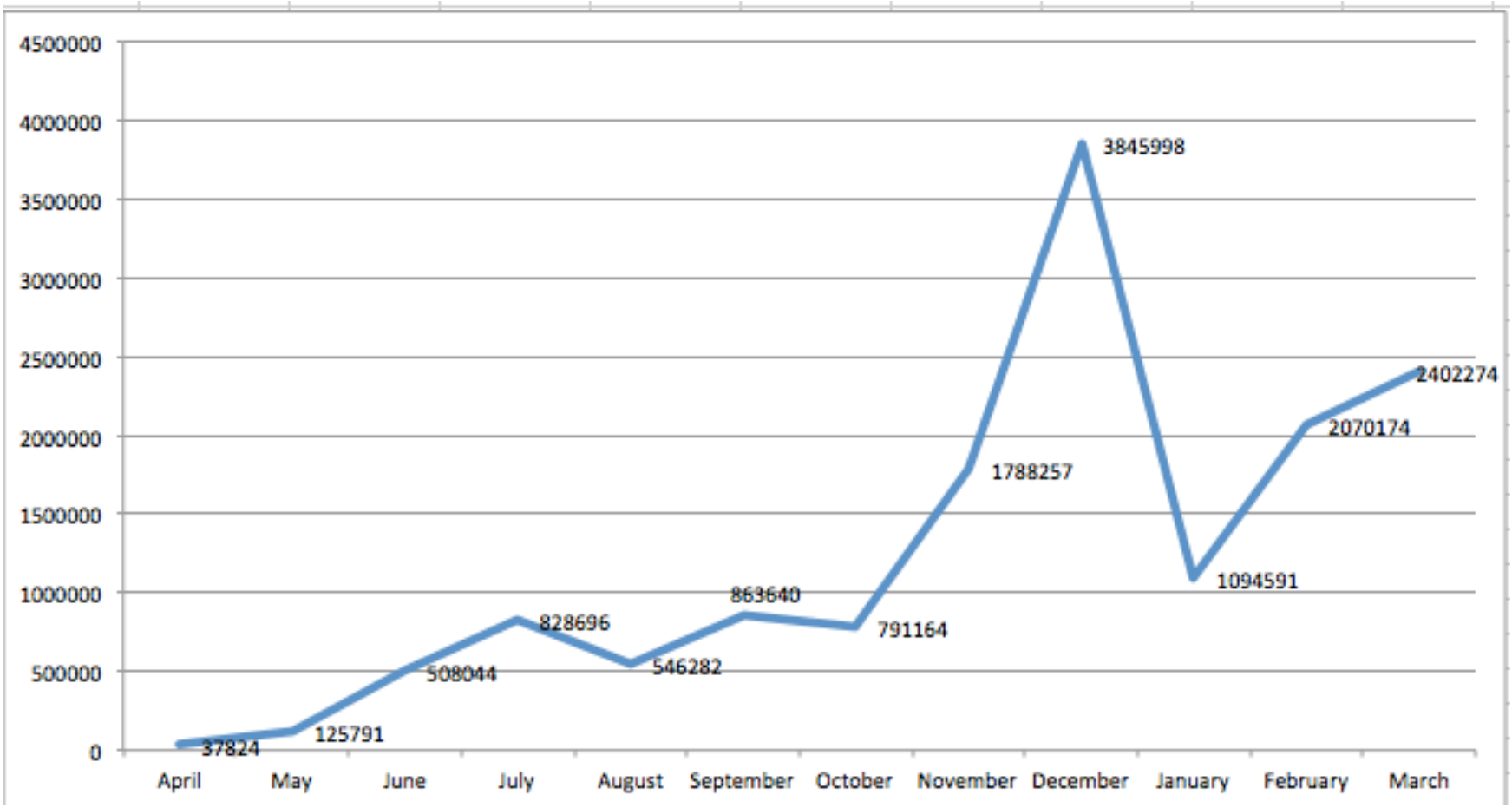
drugbank:gleevec

```
LinkSet#1 {
  chemspider:gleevec hasParent imatinib  ...
  drugbank:gleevec exactMatch imatinib  ...
}
```

APPS

Open PHACTS
Open Pharmacological Space

API Hits (April 2013 – March 2014)



Open PHACTS

Open Pharmacological Space

Open PHACTS

Browse and search the data within the Open PHACTS Discovery Platform.

Developed by the University of Manchester and University of Vienna

ChemBioNavigator

Visualise the chemical and biological space of a molecule group in a chemically-aware manner.

Developed by the University of Hamburg and BioSolveIT GmbH

PHARMATREK

Navigate pharmacological space in a flexible and interactive way.

Developed by the Consorci Mar Parc de Salut de Barcelona (PSMAR)

SciBite

Connects the latest news and events in Pharma and Biotech directly to pharmacology data within the Open PHACTS platform.

Developed by SciBite Limited

utopia

Allows the semantic enrichment of scientific articles in PDF format.

Developed by the University of Manchester

GARField

Intuitive predicts target pharmacology based on the Similar Ensemble Approach.

Developed by the Technical University of Denmark

collector

Extracts data to build QSAR predictive models with data from the eTOX project.

Developed by PSMAR as part of the eTOX project

accelrys Pipeline Pilot

A repository of useful Pipeline Pilot components and workflows has been developed.

Open PHACTS - Pipeline Pilot Community

KNIME

A KNIME repository of components and workflows has been developed.

Open PHACTS - KNIME Community

Excel

Queries the Open PHACTS API from Microsoft's Excel spreadsheet software.

Developed by the University of Vienna

AQnowledge Semantics for Science

Identifies significant entities in scientific text, and provides links to Open PHACTS Explorer.

Developed by AQnowledge

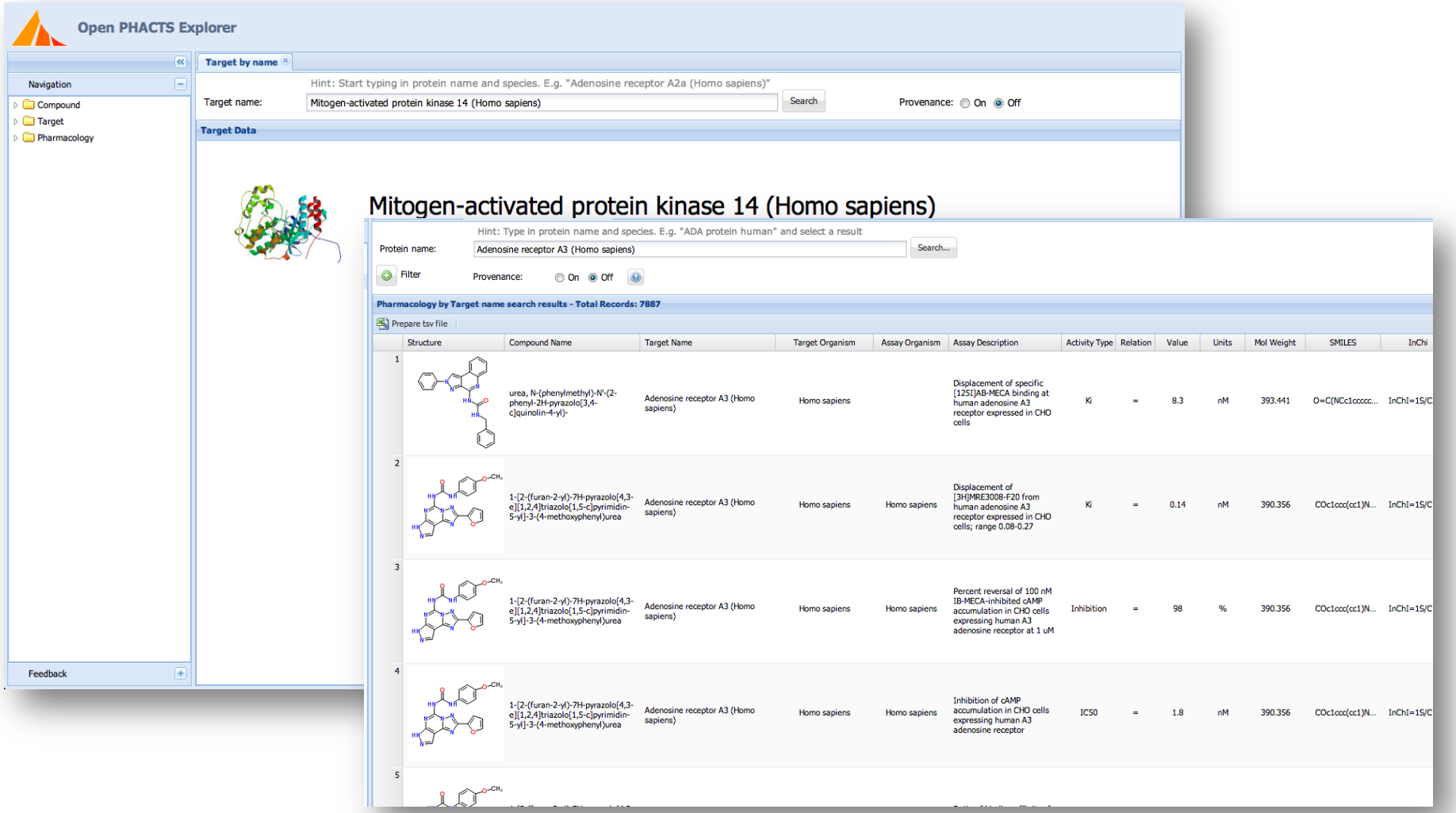
he

Helium for Excel Community Edition contains three functions that use the Open PHACTS API.

Developed by Ceiba Solutions

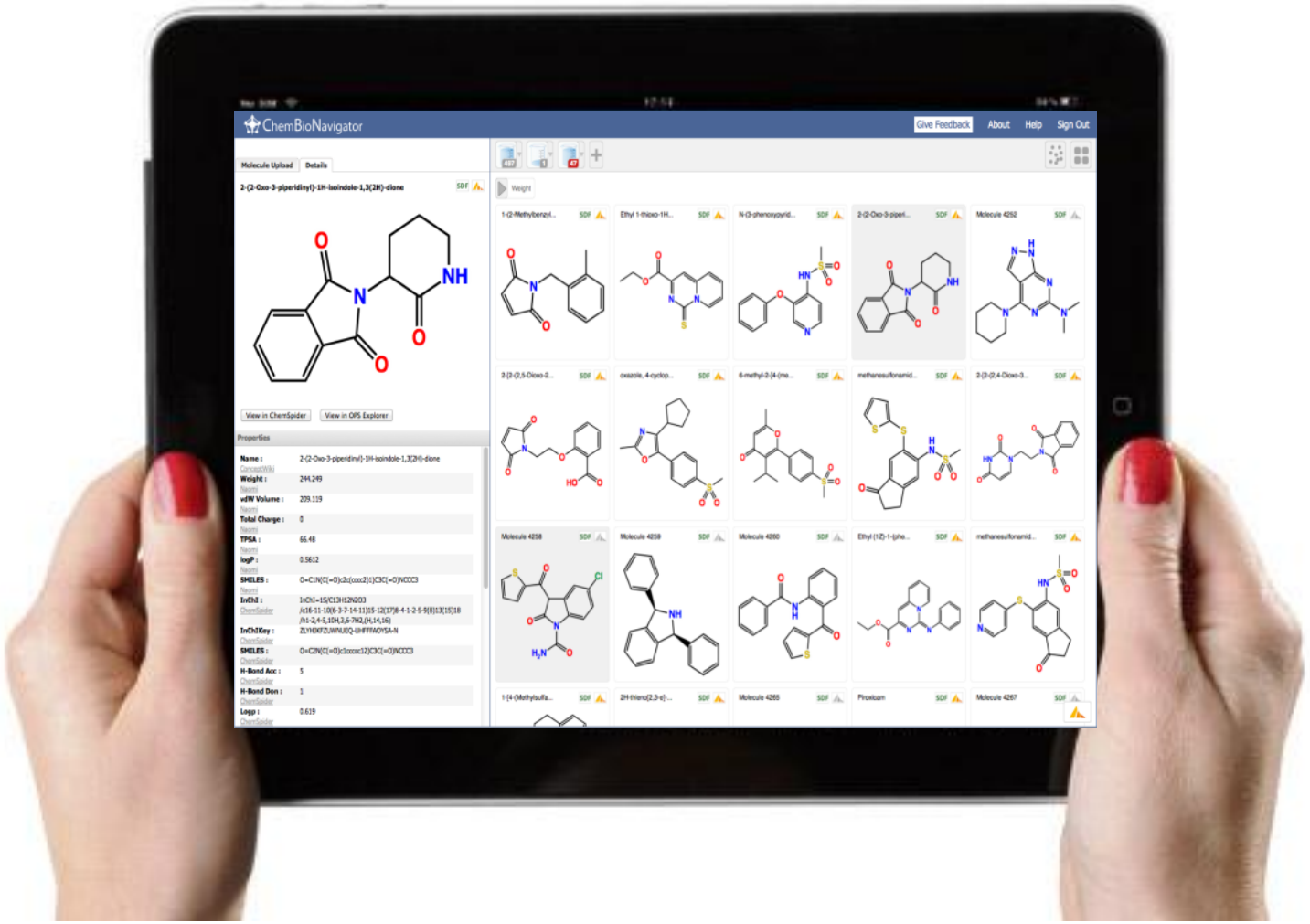
Open PHACTS

Open Pharmacological Space

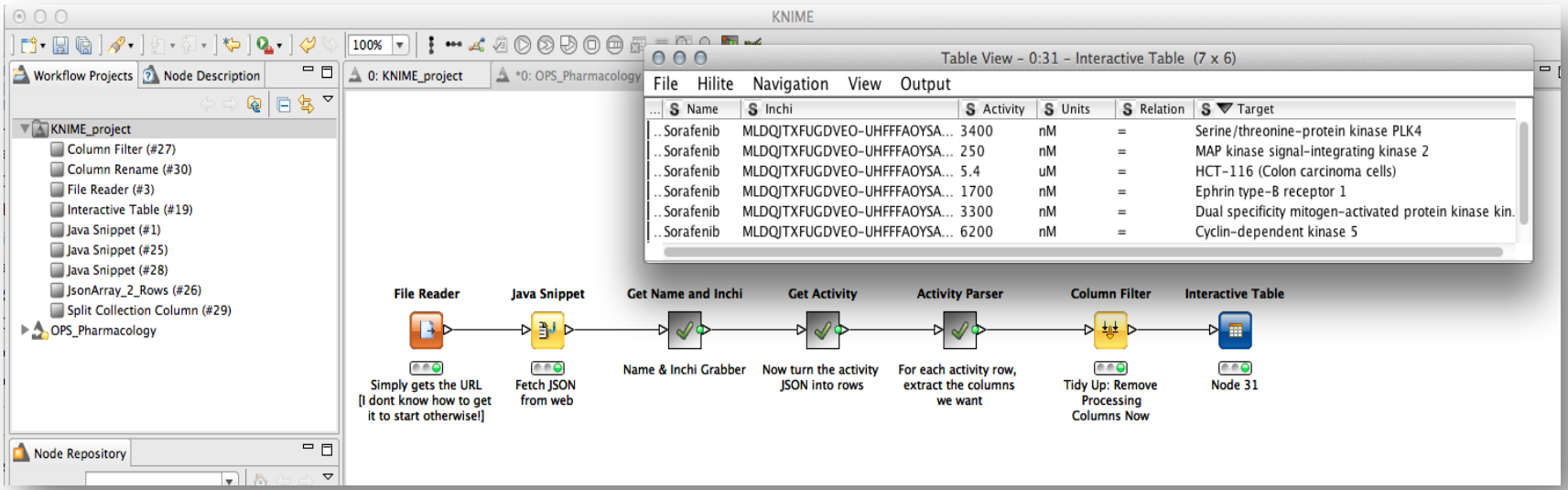


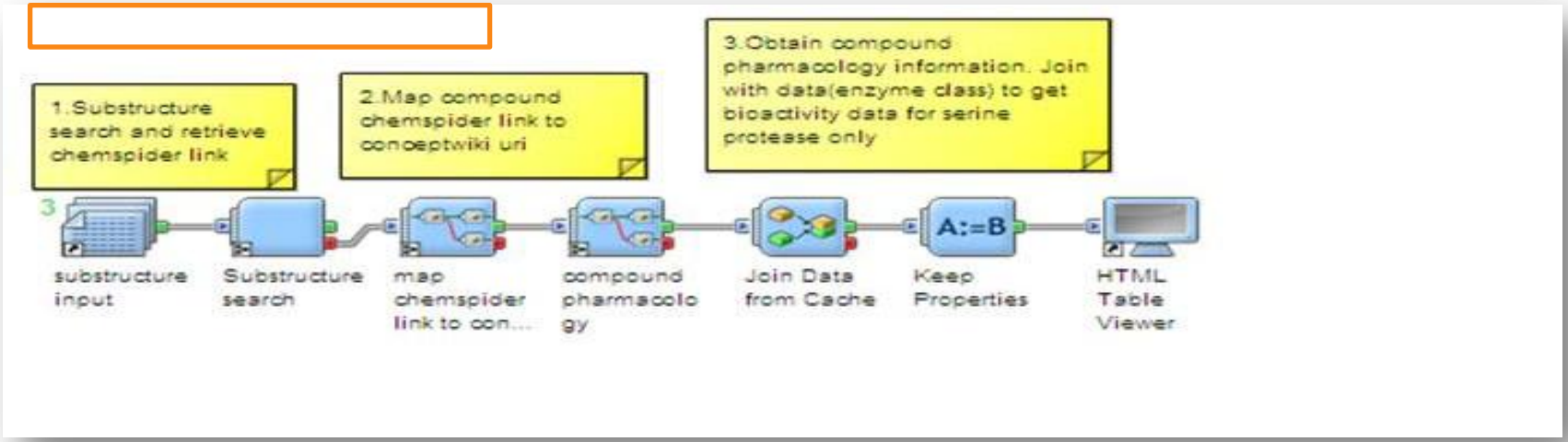
http://explorer.openphacts.org

ChemBioNavigtor



Open PHACTS
Open Pharmacological Space





Open PHACTS
Open Pharmacological Space

Open PHACTS Use Case: Neuroscience / Oncology

- Which compounds are associated with YYY and related targets to design a focused set?

First:

Pipeline Pilot
- YYY target protein sequence (Open PHACTS)

Pipeline Pilot
- top similar sequences (BLAST)

Pipeline Pilot
- compounds against all these (Open PHACTS)

Later:

Pipeline Pilot
- **in house** compounds against all these

Pipeline Pilot
- Public + in house compounds, report

Open PHACTS
Open Pharmacological Space

THE FUTURE

Open PHACTS
Open Pharmacological Space

Pharma Companies

Life Science Companies

Data Providers

Connecting Communities

App Developers

Academic Research

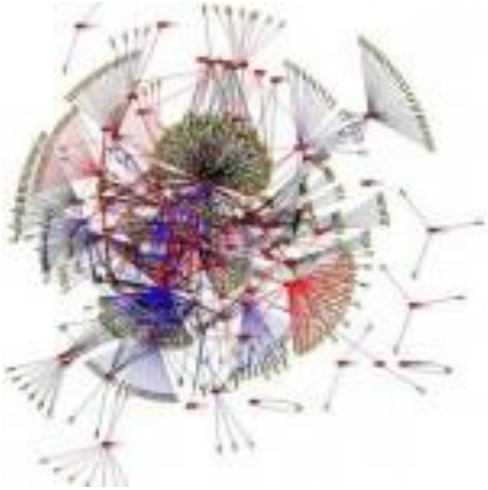
Next Gen IT

Sustaining Impact

Open PHACTS
Open Pharmacological Space

- "Software is free like puppies are free - they both need money for maintenance"

- …and more resource for future development

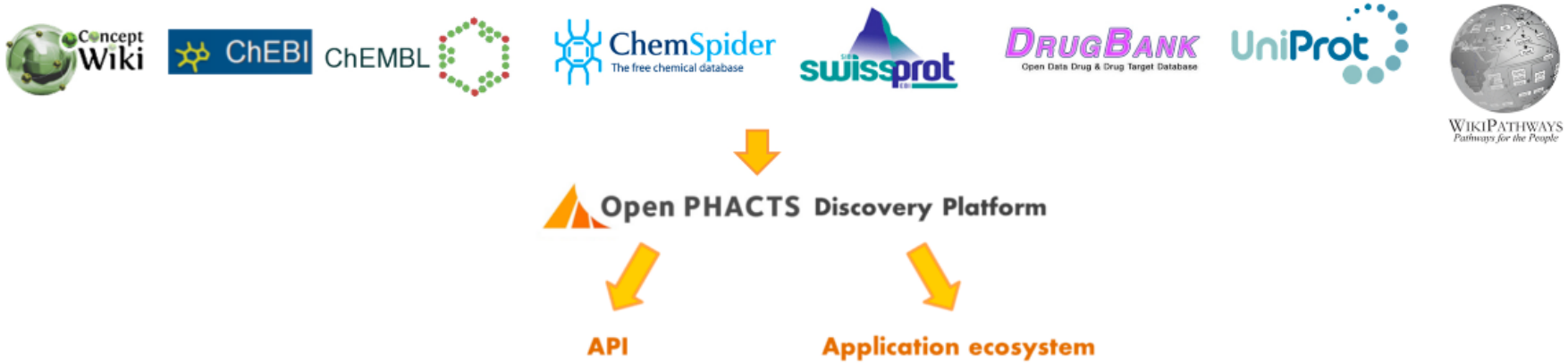


The Open PHACTS Foundation

OPF is a not-for-profit membership organisation, supporting the Open PHACTS Discovery Platform:
A sustainable, open, vibrant and interoperable information infrastructure for applied life science research and development.

To reduce the barriers to drug discovery in industry, academia and for small businesses, the Open PHACTS Discovery Platform provides tools and services to interact with multiple integrated and publicly available data sources. To integrate this data, extensive cross-referencing of scientific concepts is needed across all databases.

The Open PHACTS Foundation ensures the sustainability of the Open PHACTS Discovery Platform infrastructure and acts as a hub for relevant scientific research and development.



Open PHACTS
Open Pharmacological Space

Pfizer Limited – Coordinator
Universität Wien – Managing entity
Technical University of Denmark
University of Hamburg, Center for Bioinformatics
BioSolveIT GmbH
Consorci Mar Parc de Salut de Barcelona
Leiden University Medical Centre
Royal Society of Chemistry
Vrije Universiteit Amsterdam

Spanish National Cancer Research Centre
University of Manchester
Maastricht University
Aqnowledge
University of Santiago de Compostela
Rheinische Friedrich-Wilhelms-Universität Bonn
AstraZeneca
GlaxoSmithKline
Esteve

Novartis
Merck Serono
H. Lundbeck A/S
Eli Lilly
Netherlands Bioinformatics Centre
Swiss Institute of Bioinformatics
ConnectedDiscovery
EMBL-European Bioinformatics Institute
Janssen
OpenLink
The Open PHACTS Foundation

pmu@openphacts.org
@Open_PHACTS
Open PHACTS

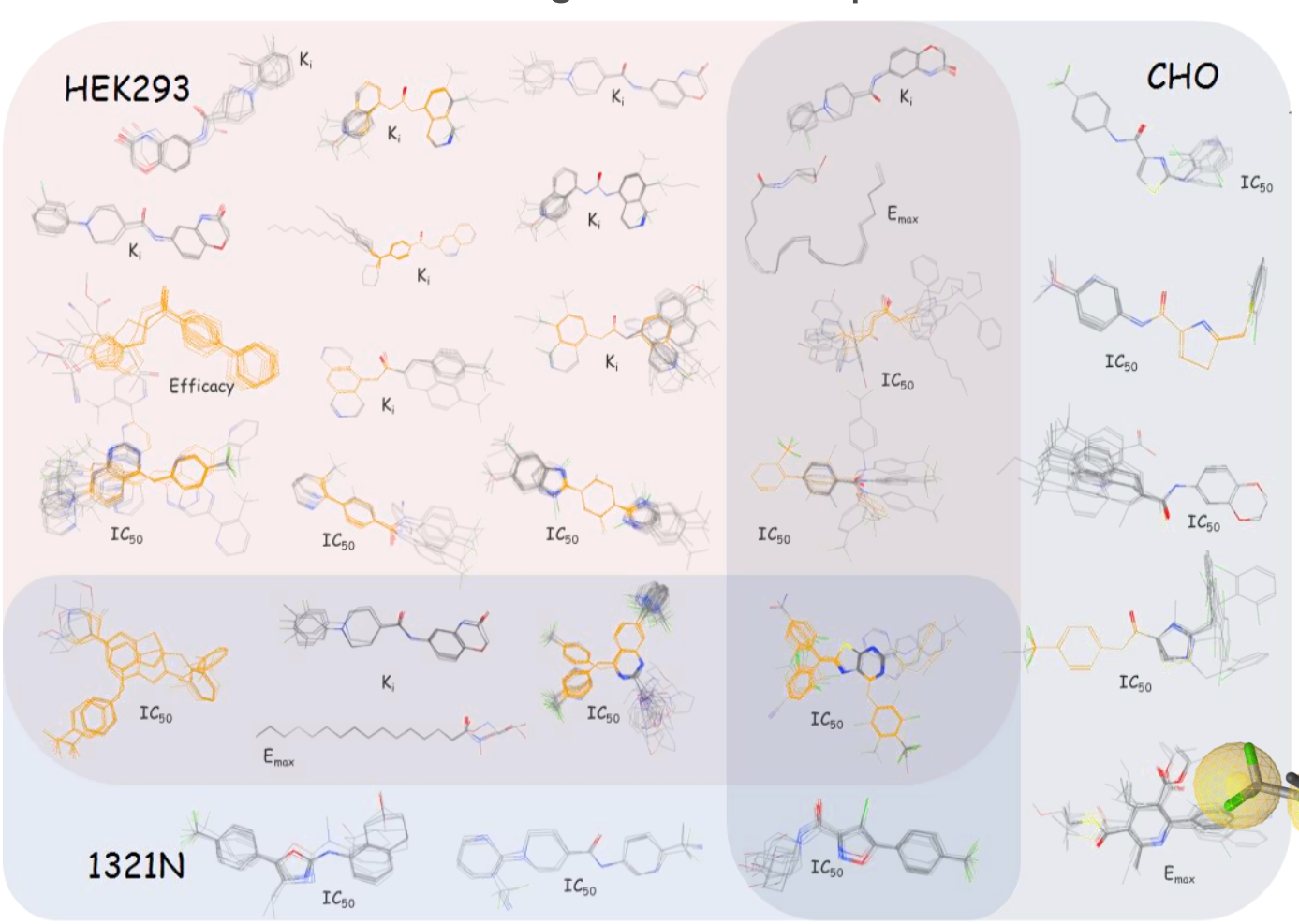
Backup

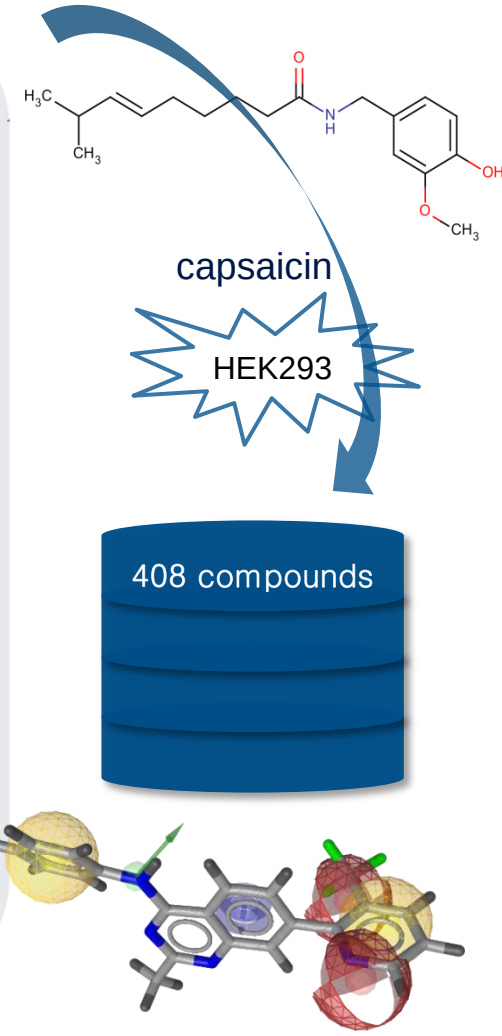
Present Content

Statistics of Datasets Loaded into Open PHACTS Version 1.3

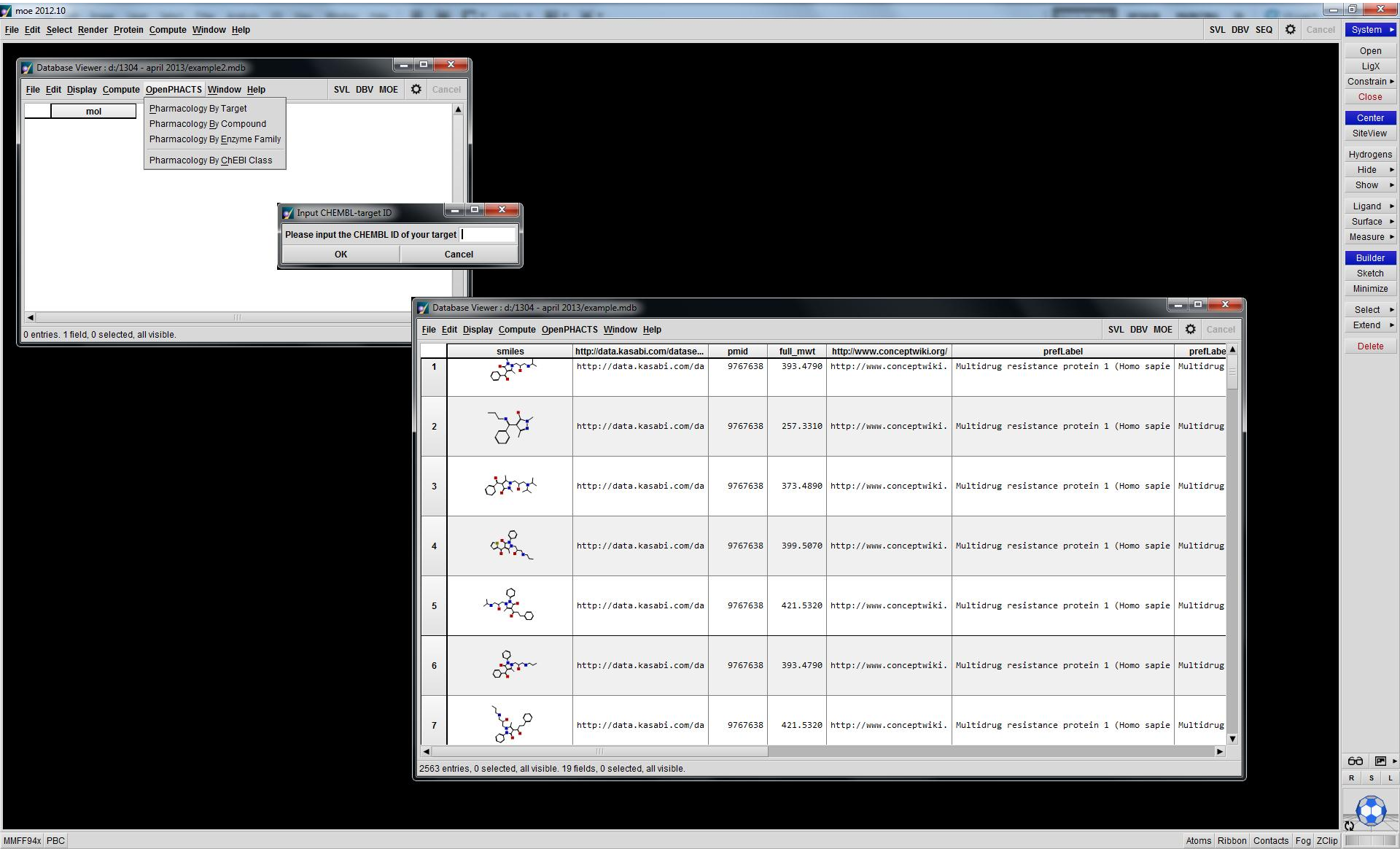
Source	Version	Supplier	Downloaded	Initial Records	Triples	Properties
Chembl	Chembl 16 RDF	EBI	25 June 2013	1,247,403 (~1,236,686 compounds, 9844 targets, 6243 target components, 873 protein classes)	304,420,681	77
DrugBank	Aug 2008	Bio2Rdf (www4.wiwiss.fu-berlin.de)	08 Aug 2012	19,628(~14,000 drugs, 5000 targets)	517,584	74
SwissProt, UniParc, UniRef	2013_06	SIB	2013_06		533,394,147	82
ENZYME	2013_07	SIB	2013_07	6,187	47,661	2
ChEBI	Release 104	EBI	19 June 2013	40,575	40,575	2
GeneOntology	Jan 21, 2013	GO	21 Jan 2013	38,137	1,265,273	26
GOA	2013	GO	09 Sept 2013	various species	23,489,501	15
WikiPathways	v0.?1_20130710	Maastricht	10 July 2013	946	1,449,981	34
ChemSpider		Open PHACTS Chemistry Registry (OCRS)	Nov 11, 2013		tbc	
ConceptWiki	version 1.3	NBIC	09 Sept 2013	2,828,966	3,739,884	1

hTRPV1 → 2328 ligands from Open PHACTS





http://www.openphacts.org


moe 2012.10
File Edit Select Render Protein Compute Window Help
SVL DBV SEQ
Cancel
System
Open
LigX
Constrain
Close
Center
SiteView
Hydrogens
Hide
Show
Ligand
Surface
Measure
Builder
Sketch
Minimize
Select
Extend
Delete
Database Viewer : d:/1304 - april 2013/example2.mdb
File Edit Display Compute OpenPHACTS Window Help
SVL DBV MOE
Cancel
mol
Pharmacology By Target
Pharmacology By Compound
Pharmacology By Enzyme Family
Pharmacology By ChEBI Class
0 entries. 1 field, 0 selected, all visible.
Input CHEMBL-target ID
Please input the CHEMBL ID of your target
OK
Cancel
Database Viewer : d:/1304 - april 2013/example.mdb
File Edit Display Compute OpenPHACTS Window Help
SVL DBV MOE
Cancel
smiles
http://data.kasabi.com/datase...
pmid
full_mwt
http://www.conceptwiki.org/
prefLabel
prefLabe
1 http://data.kasabi.com/da 9767638 393.4790 http://www.conceptwiki. Multidrug resistance protein 1 (Homo sapie Multidrug
2 http://data.kasabi.com/da 9767638 257.3310 http://www.conceptwiki. Multidrug resistance protein 1 (Homo sapie Multidrug
3 http://data.kasabi.com/da 9767638 373.4890 http://www.conceptwiki. Multidrug resistance protein 1 (Homo sapie Multidrug
4 http://data.kasabi.com/da 9767638 399.5070 http://www.conceptwiki. Multidrug resistance protein 1 (Homo sapie Multidrug
5 http://data.kasabi.com/da 9767638 421.5320 http://www.conceptwiki. Multidrug resistance protein 1 (Homo sapie Multidrug
6 http://data.kasabi.com/da 9767638 393.4790 http://www.conceptwiki. Multidrug resistance protein 1 (Homo sapie Multidrug
7 http://data.kasabi.com/da 9767638 421.5320 http://www.conceptwiki. Multidrug resistance protein 1 (Homo sapie Multidrug
2563 entries, 0 selected, all visible. 19 fields, 0 selected, all visible.
MMFF94x PBC
Atoms Ribbon Contacts Fog ZClip