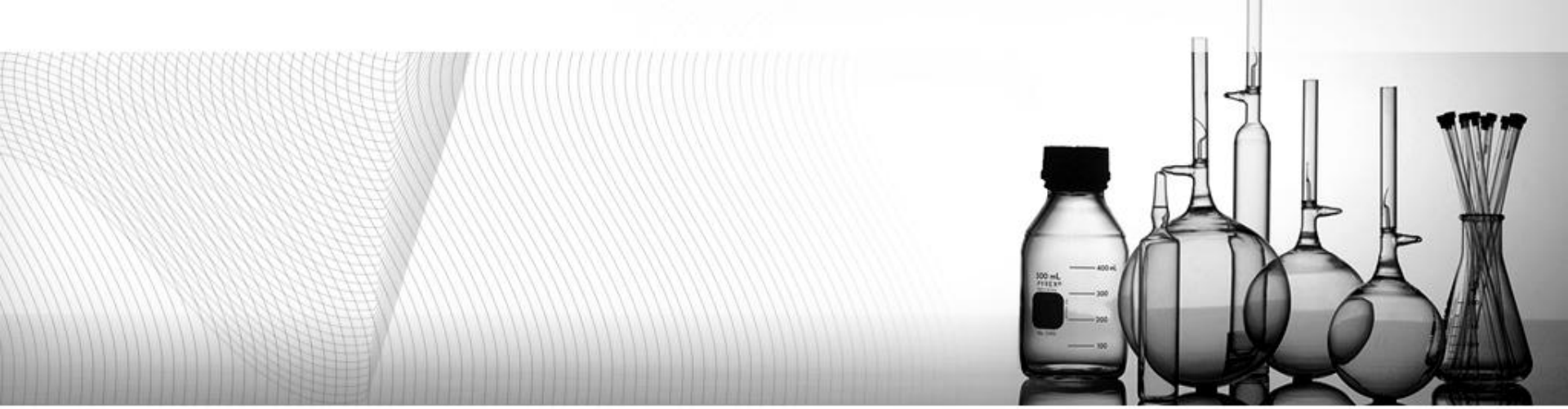


Open PHACTS
Open Pharmacological Space

Opening up Pharmacological Space: The Open PHACTS API

@Chris_Evelo
Dept. Bioinformatics - BiGCaT
Maastricht University

Open PHACTS
Open Pharmacological Space

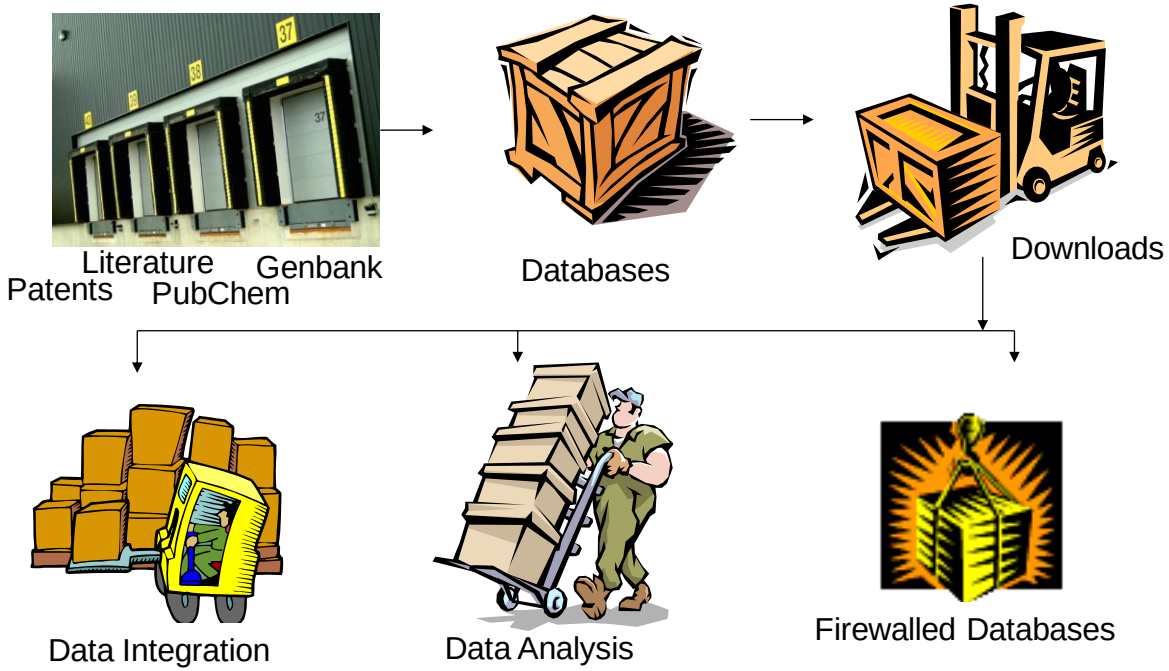
Fundamental issues:

- There is a *lot* of science outside your walls
- It's a chaotic space
- Scientists want to find information quickly and easily
- Often they just "cant get there"
  (or don't even know where "there" is)
- And you have to manage it all (or not)

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

The Innovative Medicines Initiative

- EC funded public-private partnership for pharmaceutical research
- Focus on key problems
  - Efficacy, Safety, Education & Training, **Knowledge Management**

The Open PHACTS Project

- Create a *semantic integration hub ("Open Pharmacological Space")…*
- Deliver services to support on-going drug discovery programs in pharma and public domain
- *Not just another project;* leading academics in semantics, pharmacology and informatics, driven by solid industry business requirements
- 23 academic partners, 8 pharmaceutical companies, 3 biotech companies
- Work split into clusters:
  - Technical Build (*focus here)*
  - Scientific Drive
  - Community & Sustainability

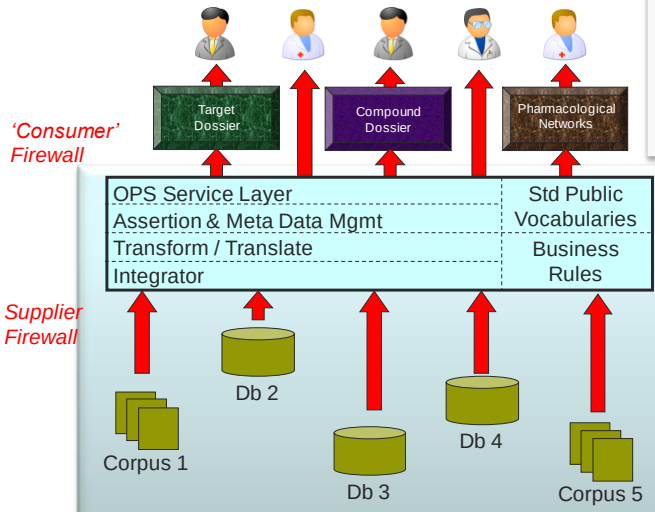
Open PHACTS
Open Pharmacological Space

Major Work Streams

Build: OPS service layer and resource integration

Drive: Development of exemplar work packages & Applications

Sustain: Community engagement and long-term sustainability



Work Stream 2: Exemplar Drug Discovery Informatics tools
Develop exemplar services to test OPS Service Layer
Target Dossier (Data Integration)
Pharmacological Network Navigator (Data Visualisation)
Compound Dossier (Data Analysis)

Work Stream 1: Open Pharmacological Space (OPS) Service Layer

Standardised software layer to allow public DD resource integration

- Define standards and construct OPS service layer
- Develop interface (API) for data access, integration and analysis
- Develop secure access models

Existing Drug Discovery (DD) Resource Integration

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

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 clinical data for a given list of compounds defined by their chemical structure (with options to match stereochemistry or not).
41	13	8	A project is considering Protein Kinase C Alpha (PRKCA) as a target. What are all the compounds known to modulate the target directly? What are the compounds that may modulate the target directly? i.e. return all cmpds active in assays where the resolution is at least at the level of the target family (i.e. PKC) both from structured assay databases and the literature.
44	13	8	Give me all active compounds on a given target with the relevant assay data
46	13	8	Give me the compound(s) which hit most specifically the multiple targets in a given pathway (disease)
59	14	8	Identify all known protein-protein interaction inhibitors

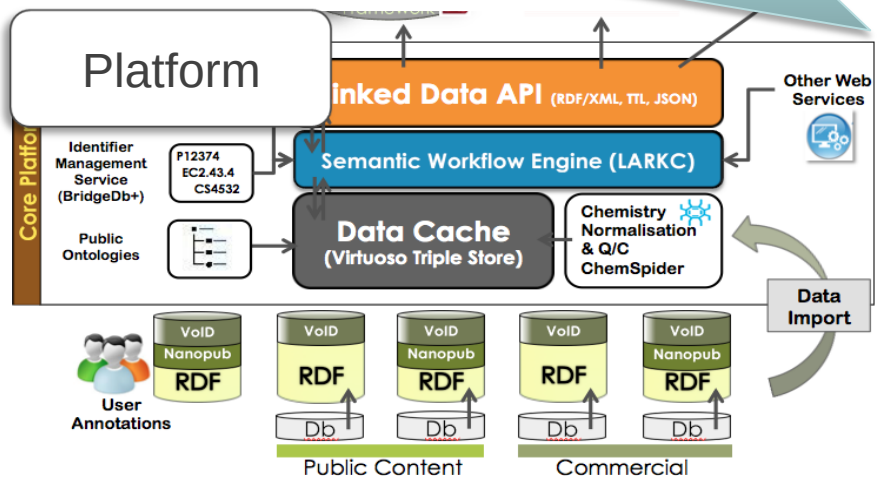
Paper in DDT http://www.sciencedirect.com/science/article/pii/S135964613001542

Open PHACTS Scientific Services

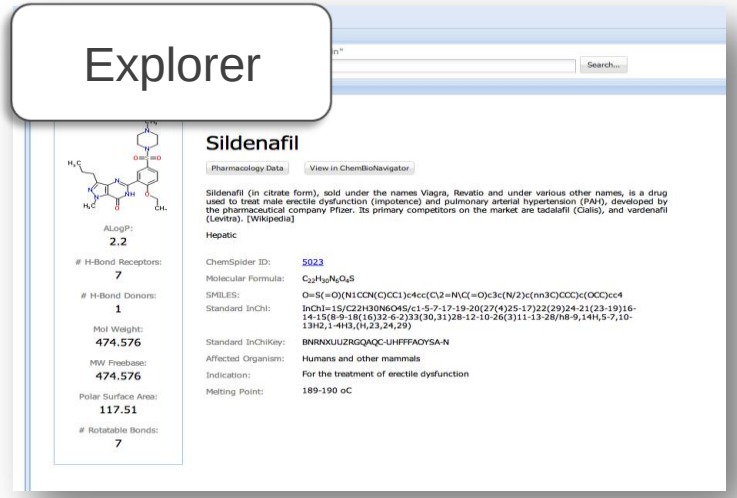
Open PHACTS
Open Pharmacological Space

"Provenance Everywhere"

Platform



Explorer



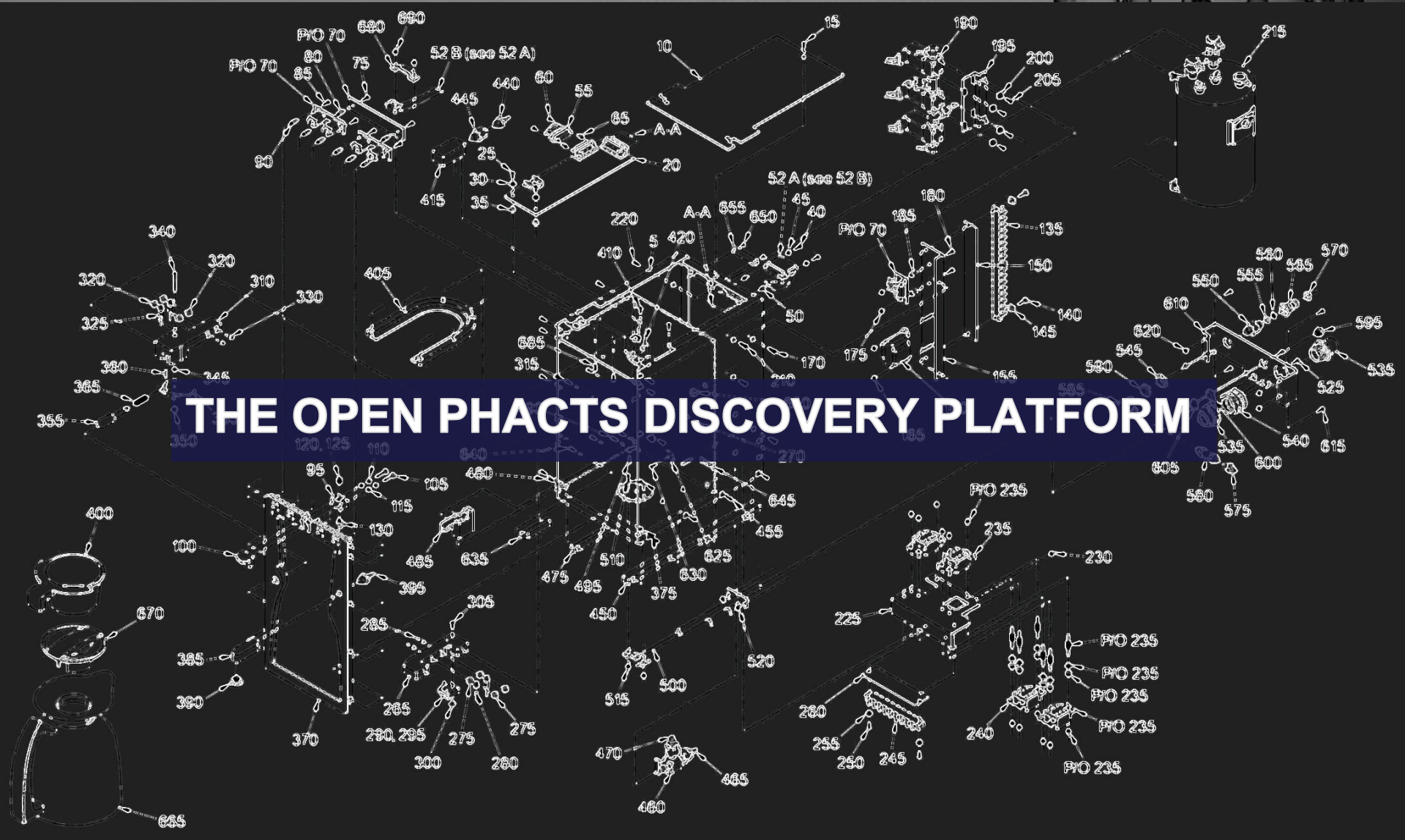
Apps

API

Standards

Open PHACTS

Open Pharmacological Space



THE OPEN PHACTS DISCOVERY PLATFORM

Open PHACTS
Open Pharmacological Space

Contrary to popular belief:

If you produce RDF from two different data sources and put them in a triple store… ***Magic does not happen, and it does not automatically become linked data!***

You need to link:

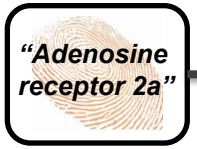
- Terms to ontologies
- Ontologies to ontologies
- Identifiers to identifiers
- Text to known concepts
- Chemicals to known structures

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)

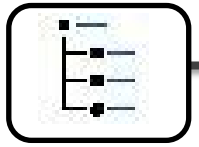
Semantic Workflow Engine

Data Cache
(Virtuoso Triple Store)

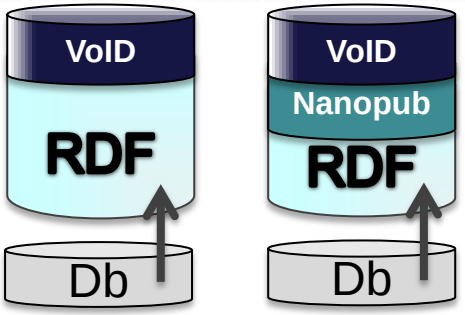
Domain Specific Services

Chemistry
Registration
Normalisation
& Q/C

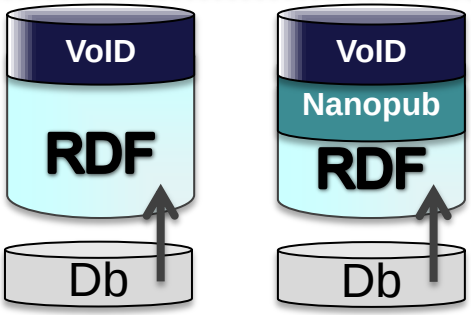
Indexing



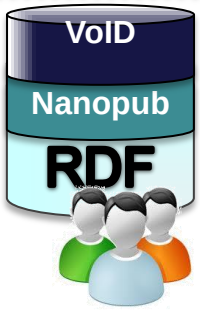
Public Ontologies



Public Content



Commercial



User Annotations

Open PHACTS
Open Pharmacological Space

Present Content

Source	Initial Records	Triples	Properties
Chembl	1,149,792 ~1,091,462 cmpds ~8845 targets	146,079,194	17 cmpds 13 targets
DrugBank	19,628 ~14,000 drugs ~5000 targets	517,584	74
UniProt	536,789	156,569,764	78
ENZYME	6,187	73,838	2
ChEBI	35,584	905,189	2
GO/GOA	38,137	24,574,774	42
ChemSpider/ACD	1,194,437	161,336,857	22 ACD, 4 CS
ConceptWiki	2,828,966	3,739,884	1
WikiPathways	Just added		

Open PHACTS
Open Pharmacological Space

Quantitative Data Challenges

STANDARD_TYPE	UNIT_COUNT
AC50	7
Activity	421
EC50	39
IC50	46
ID50	42
Ki	23
Log IC50	4
Log Ki	7
Potency	11
log IC50	0

STANDARD_TYPE	STANDARD_UNITS	COUNT(*)
IC50	nM	829448
IC50	ug.mL-1	41000
IC50		38521
IC50	ug/ml	2038
IC50	ug ml-1	509
IC50	mg kg-1	295
IC50	molar ratio	178
IC50	ug	117
IC50	%	113
IC50	uM well-1	52

>5000 types

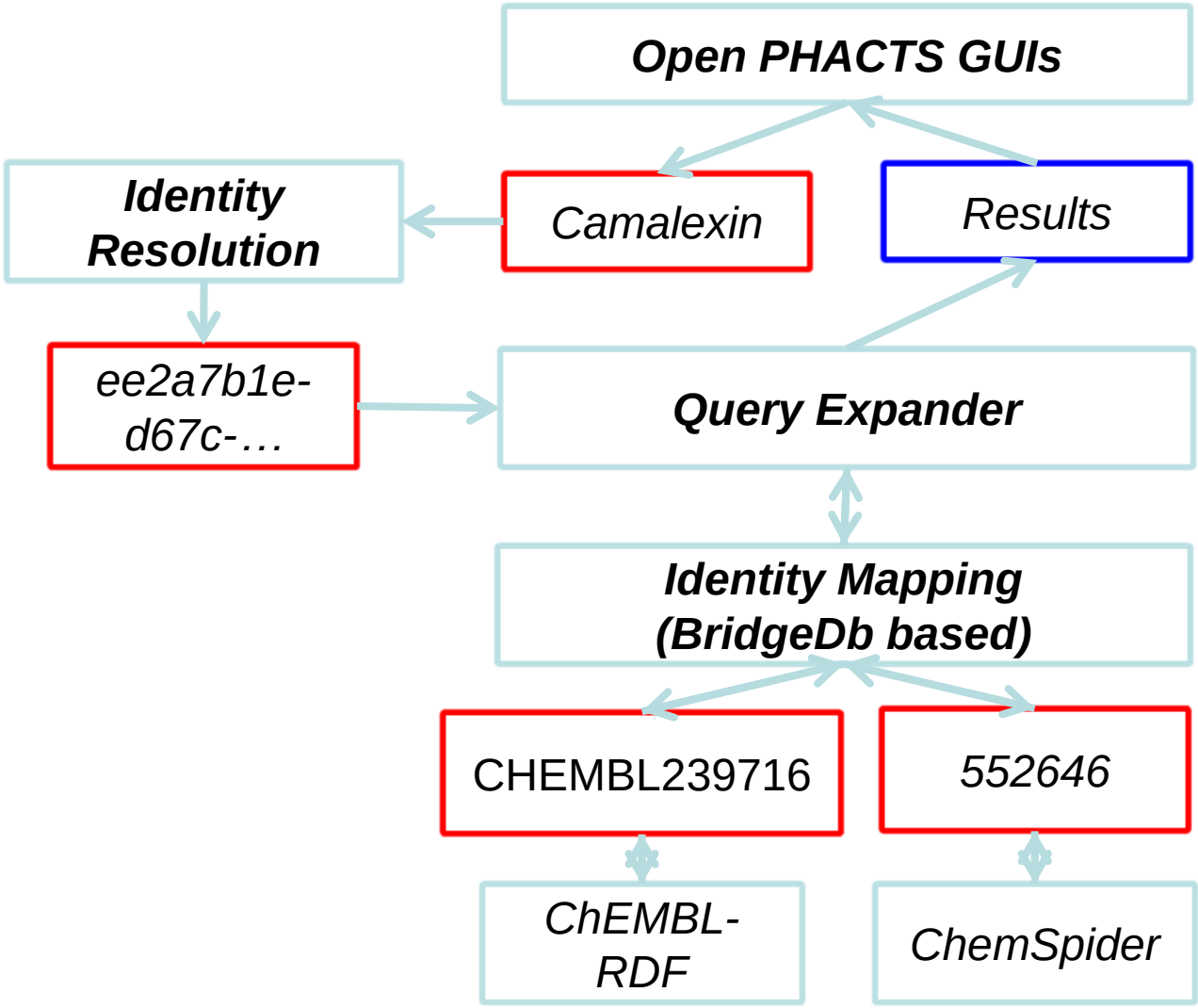
Implemented using the Quantities, Dimension, Units, Types Ontology (http://www.qudt.org/)

~ 100 units

Open PHACTS
Open Pharmacological Space

Let the IMS take the strain….

Open PHACTS
Open Pharmacological Space



Open PHACTS
Open Pharmacological Space

Chemistry Registration

- Existing chemistry registration system uses standard ChemSpider deposition system: includes low-level structure validation and manual curation service by RSC staff.

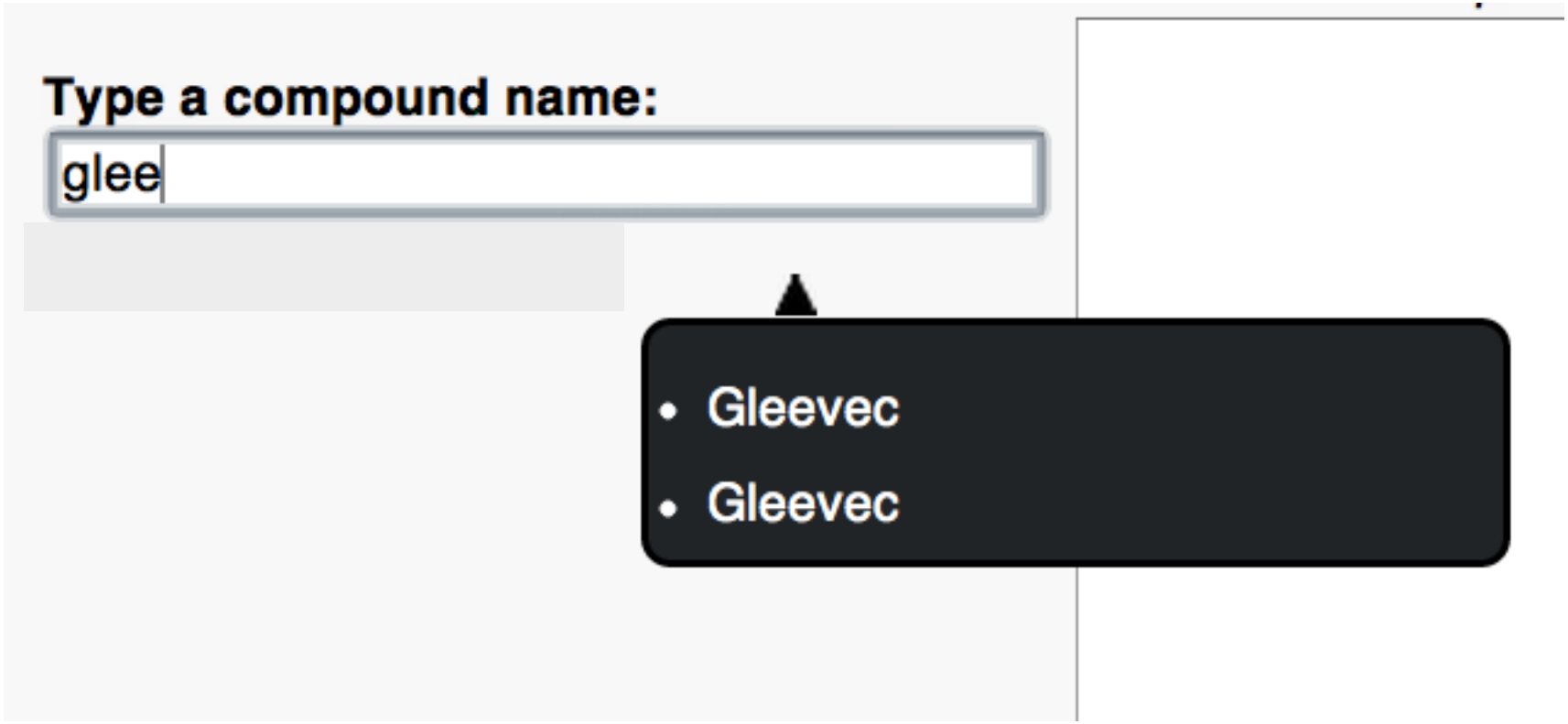
- New Registration System in Development
  - Utilizes ChemSpider Validation and Standardization platform including collapsing tautomers
  - Utilizes FDA rule set as basis for standardization (GSK lead)
  - Will generate Open PHACTS identifier (OPS ID)

Chemistry Registration Normalisation & Q/C

ChemSpider
The free chemical database

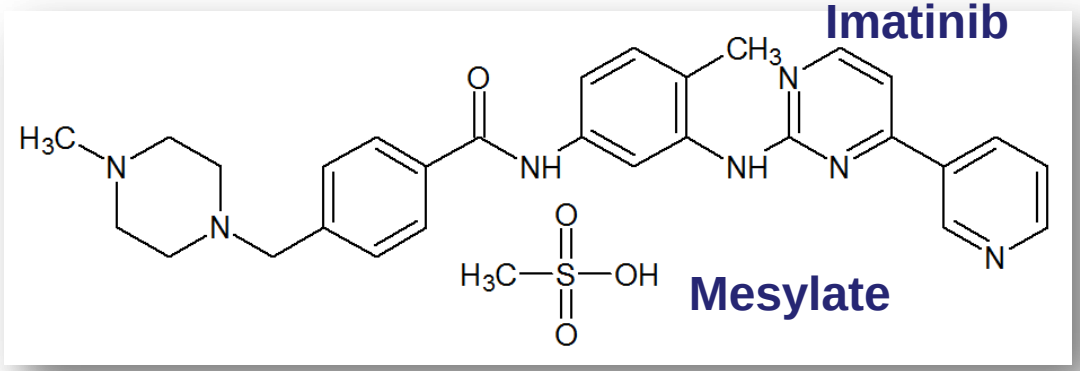
Open PHACTS
Open Pharmacological Space

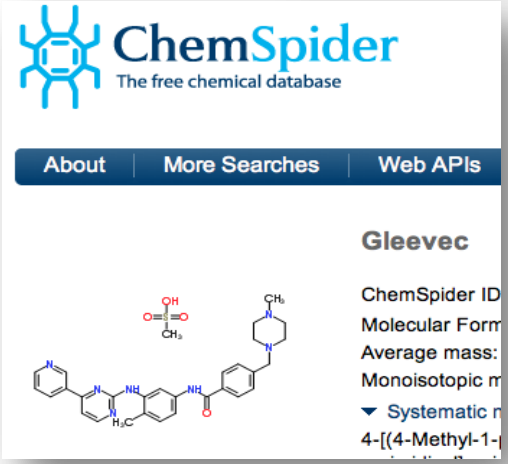
Its easy to integrate, difficult to integrate well:



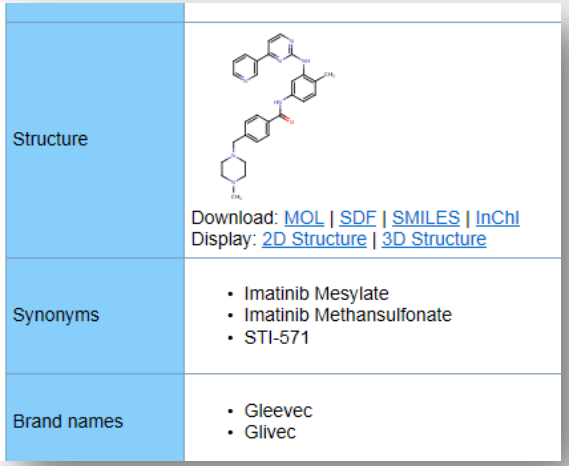
Open PHACTS
Open Pharmacological Space

What Is Gleevec?

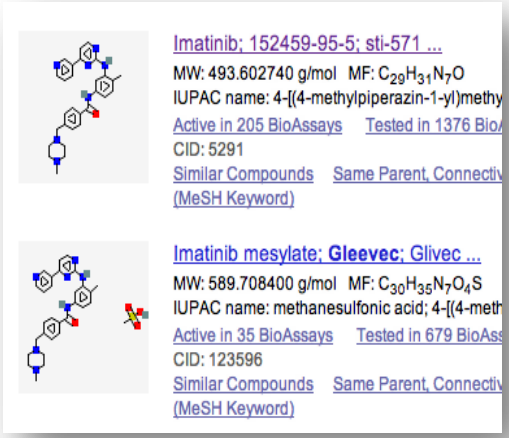




ChemSpider



Drugbank



PubChem

Open PHACTS
Open Pharmacological Space

Dynamic Equality

chemspider:gleevec

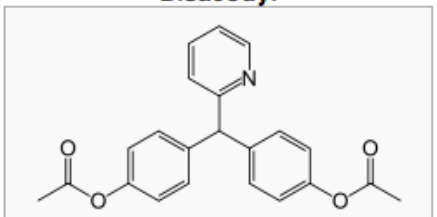
drugbank:gleevec

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

Open PHACTS

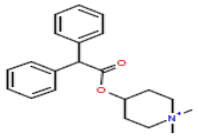
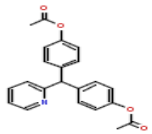
Open Pharmacological Space

Bisacodyl



Systematic (IUPAC) name

(pyridin-2-ylmethanediyl)dibenzene-4,1-diyl diacetate
OR
4,4'-(pyridin-2-ylmethylene)bis(4,1-phenylene) diacetate

Compound	Synonym
 CHEMBL168067	Bisacodyl
 CHEMBL942	Bisacodyl

Concept Wiki

Home Search

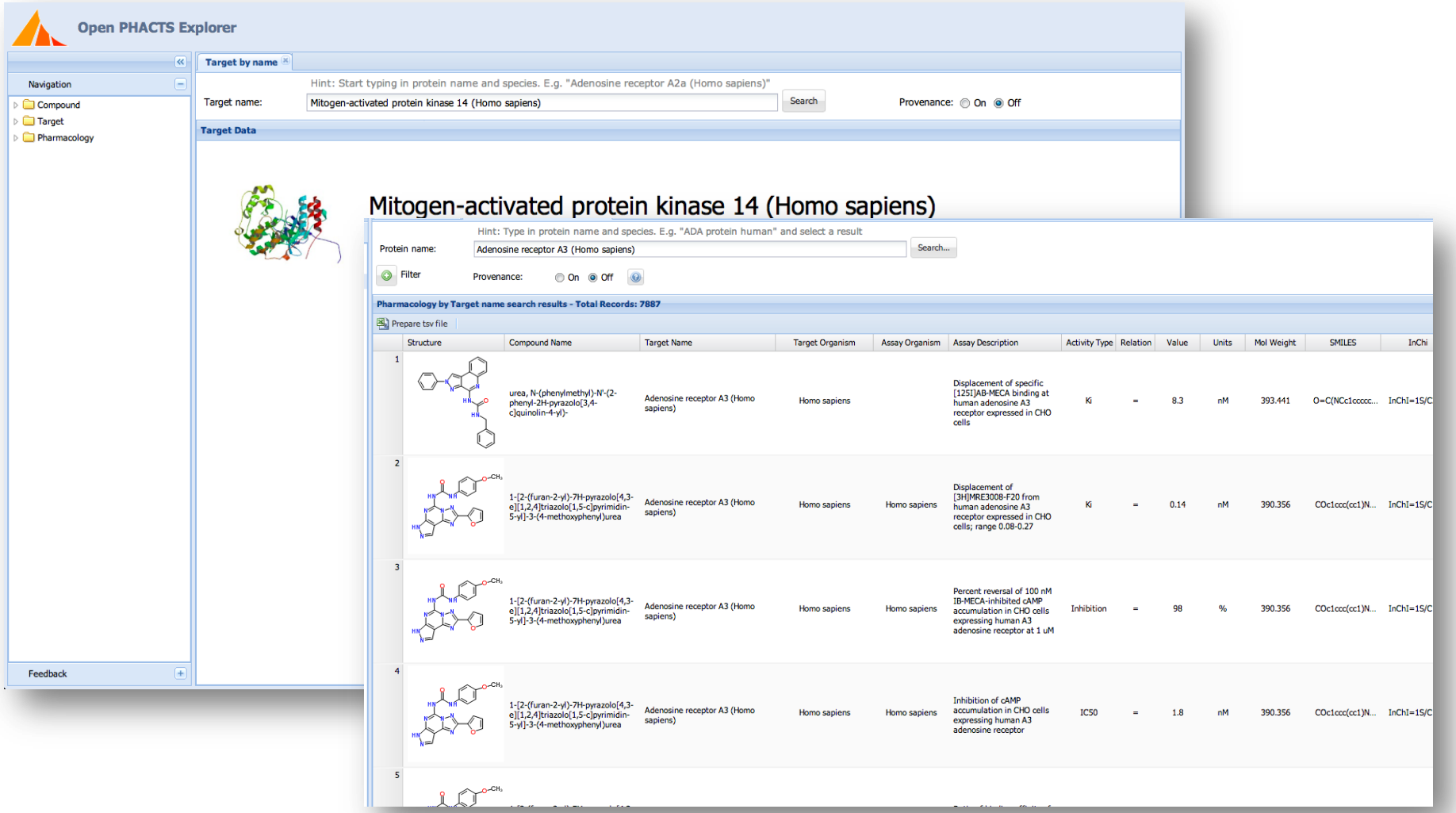
Terminology

Save Add Click on any of the labels to start editting

Term	Authority
Milnacipran	**Community, ChemSpider**
Enter A Term Here	Community
(1R,2S)-2-(Aminomethyl)-N,N-diethyl-1-phenylcyclopropanecarboxamide	Community, ChemSpider
cis-(±)-2-(Aminomethyl)-N,N-diethyl-1-phenylcyclopropanecarboxamide	Community, ChemSpider
cyclopropanecarboxamide, 2-(aminomethyl)-N,N-diethyl-1-phenyl-, (1R,2S)-	Community, ChemSpider
92623-85-3	Community, ChemSpider

Open PHACTS

Open Pharmacological Space



explorer.openphacts.org

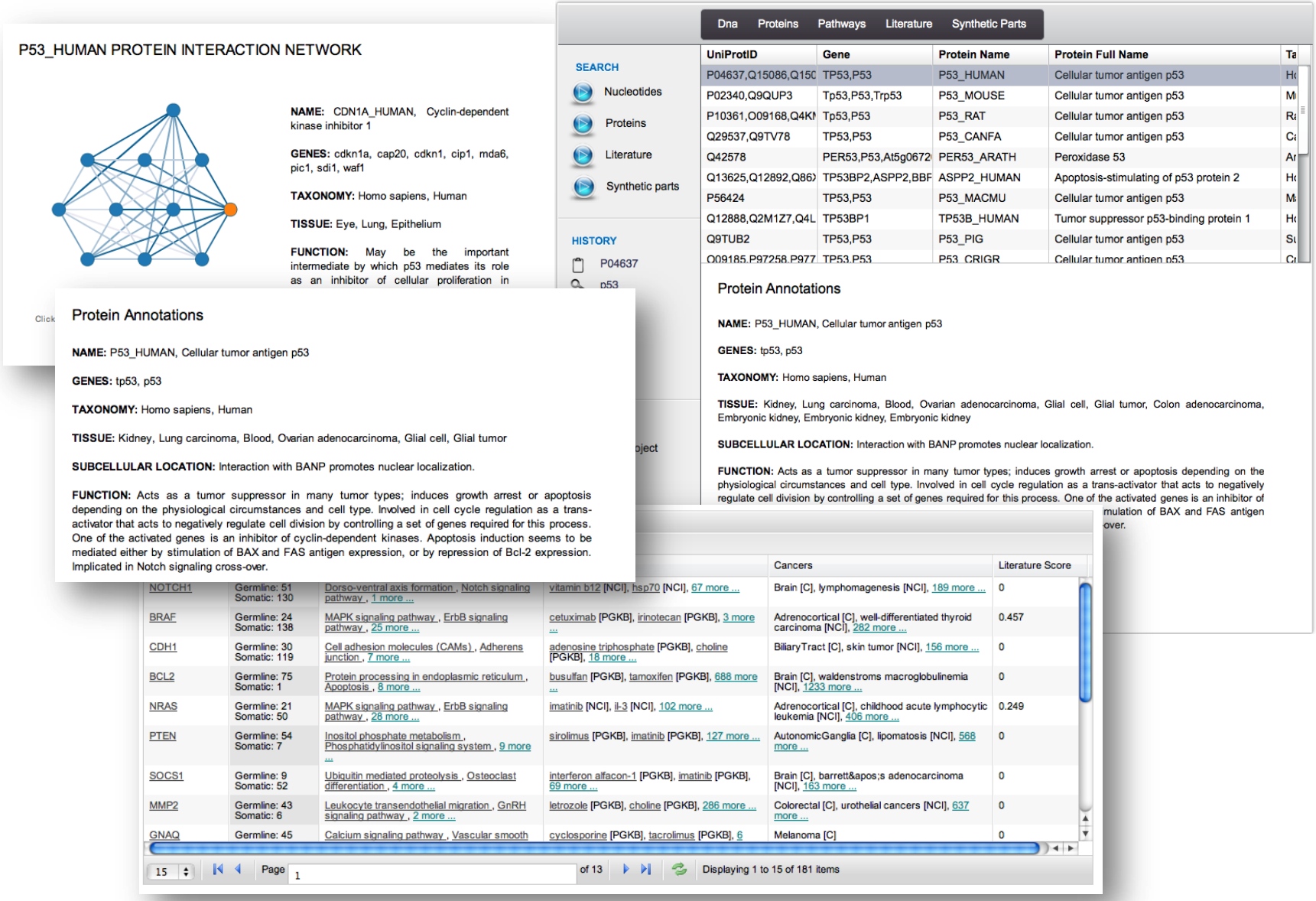
Example applications

Open PHACTS
Open Pharmacological Space

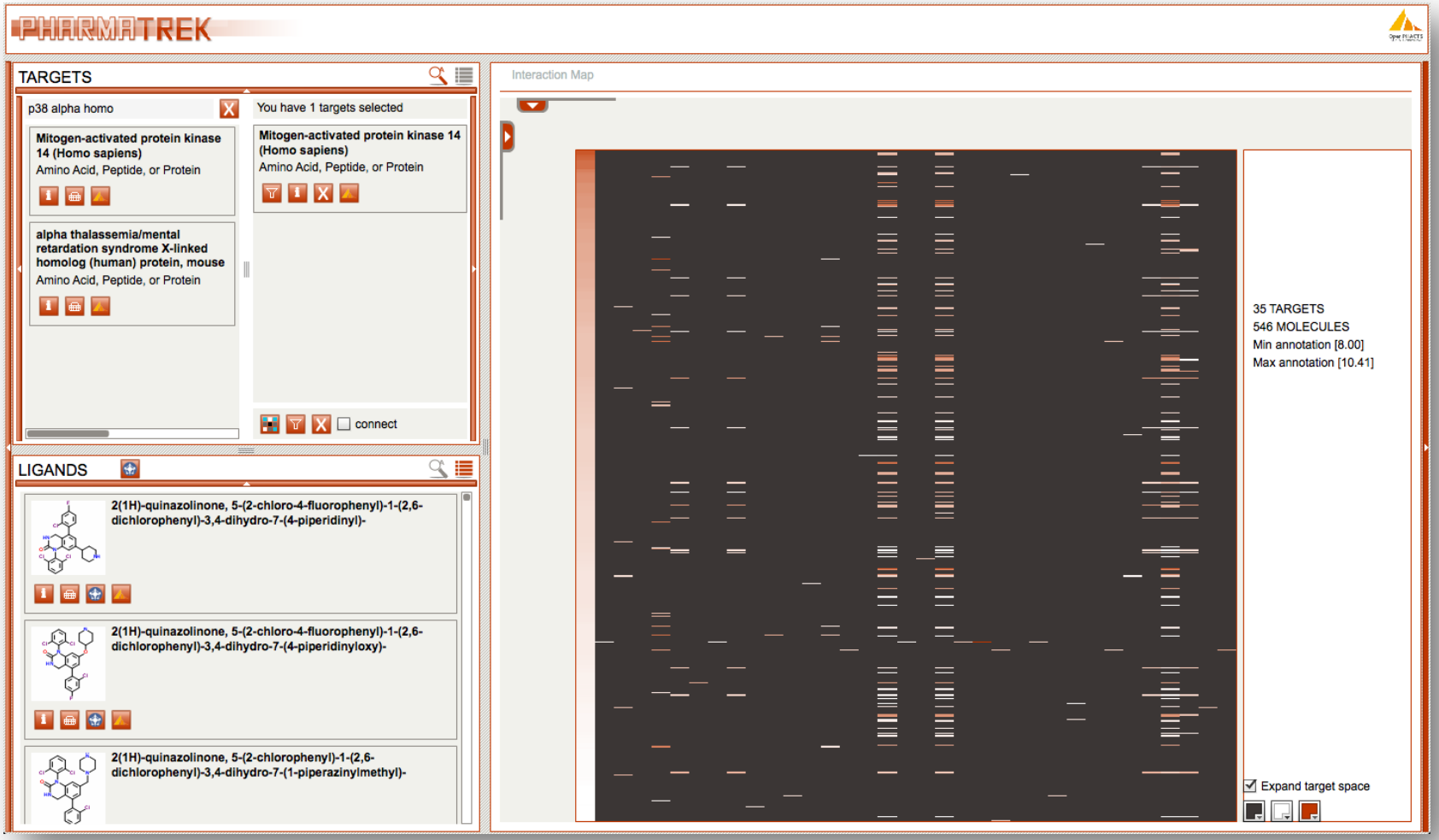
Advanced analytics	
ChemBioNavigator	Navigating at the interface of chemical and biological data with sorting and plotting options
TargetDossier	Interconnecting Open PHACTS with multiple target centric services. Exploring target similarity using diverse criteria
PharmaTrek	Interactive Polypharmacology space of experimental annotations
UTOPIA	Semantic enrichment of scientific PDFs
Predictions	
GARFIELD	Prediction of target pharmacology based on the Similar Ensemble Approach
eTOX connector	Automatic extraction of data for building predictive toxicology models in eTOX project

Target dossier (CNIO)

Front-end framework to visualize biological data

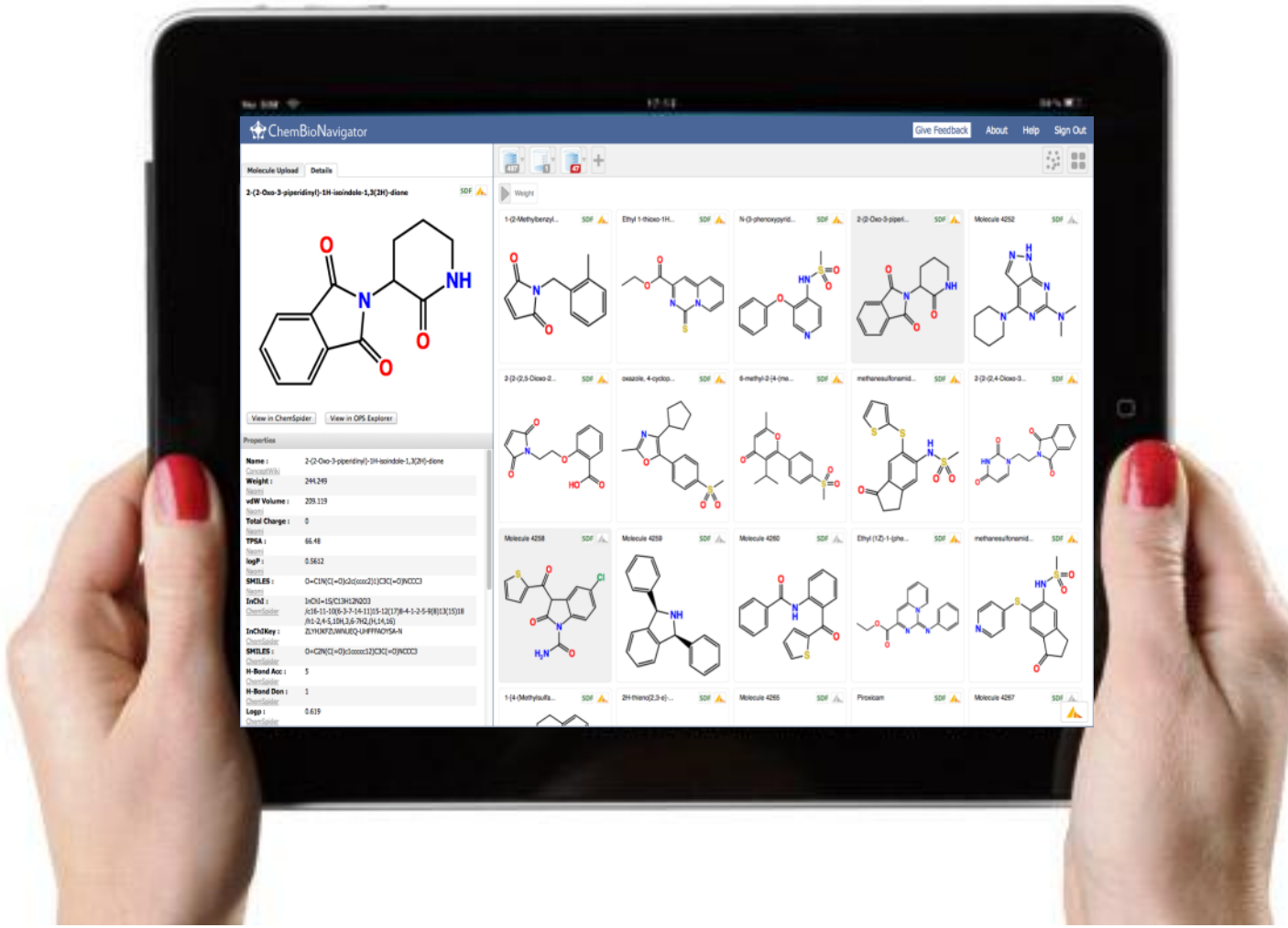


Open PHACTS
Open Pharmacological Space



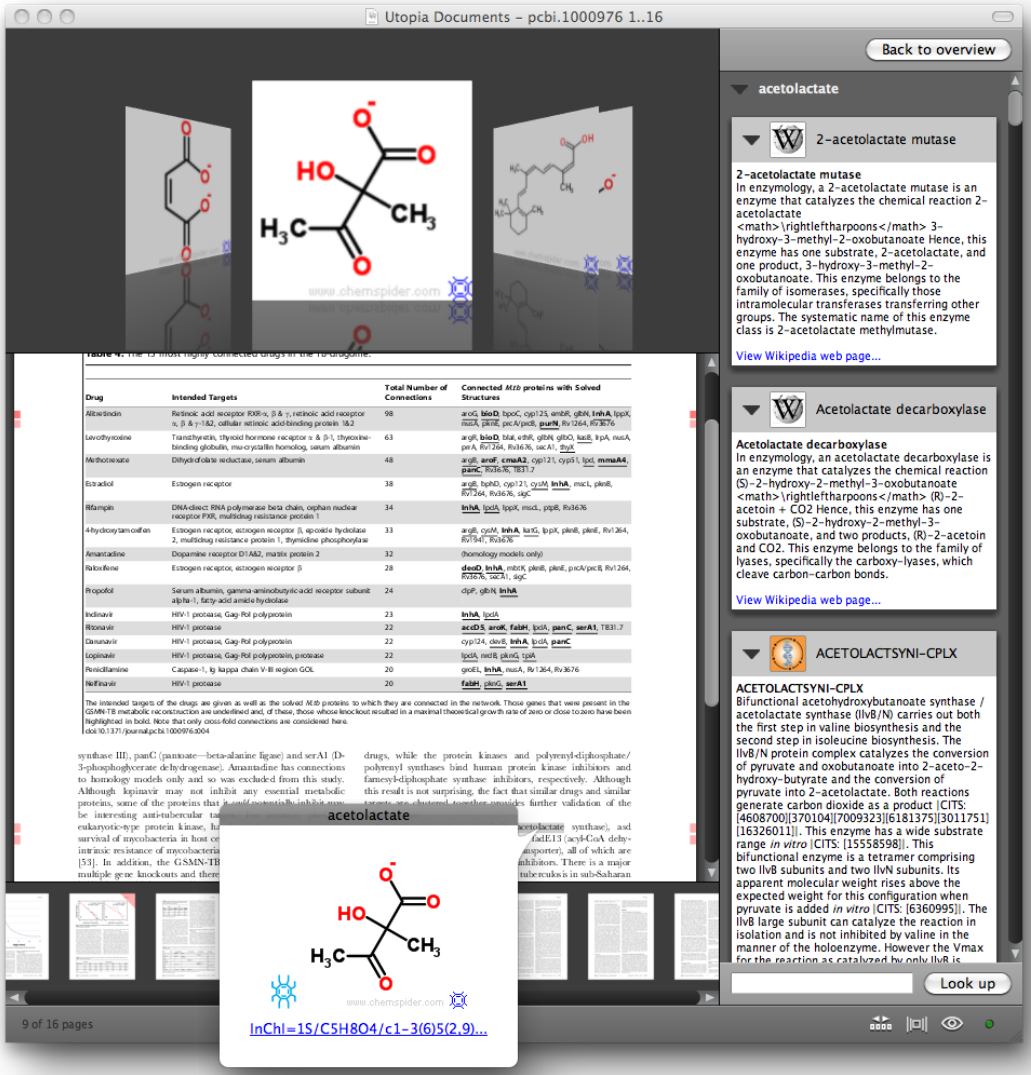
Open PHACTS

Open Pharmacological Space



Open PHACTS

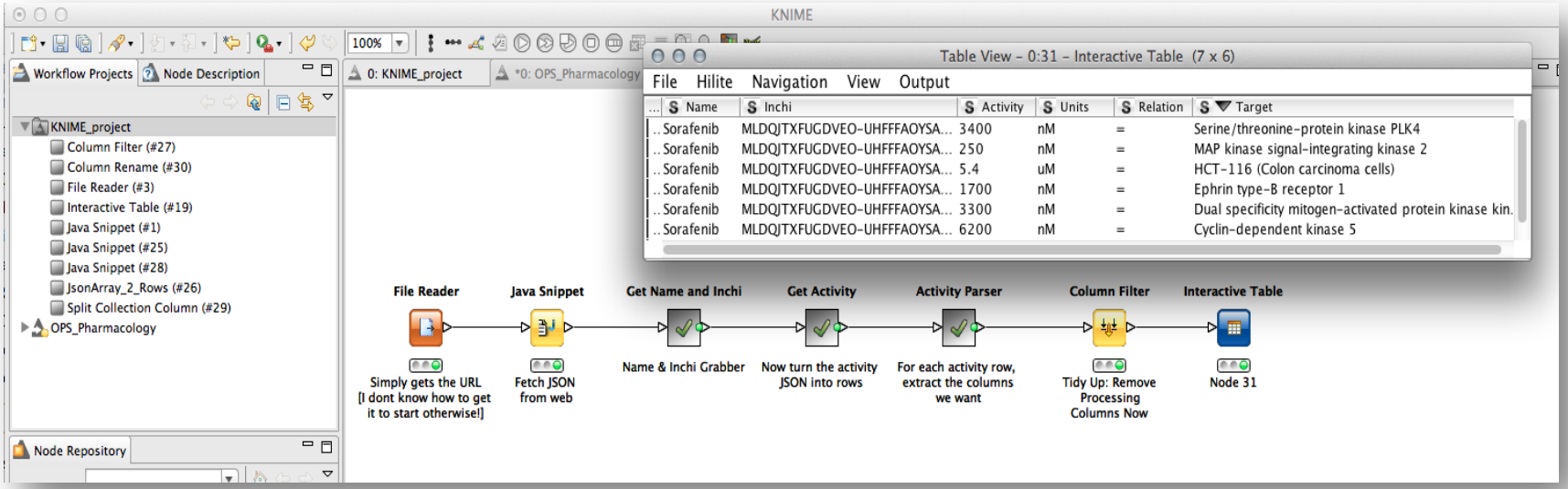
Open Pharmacological Space

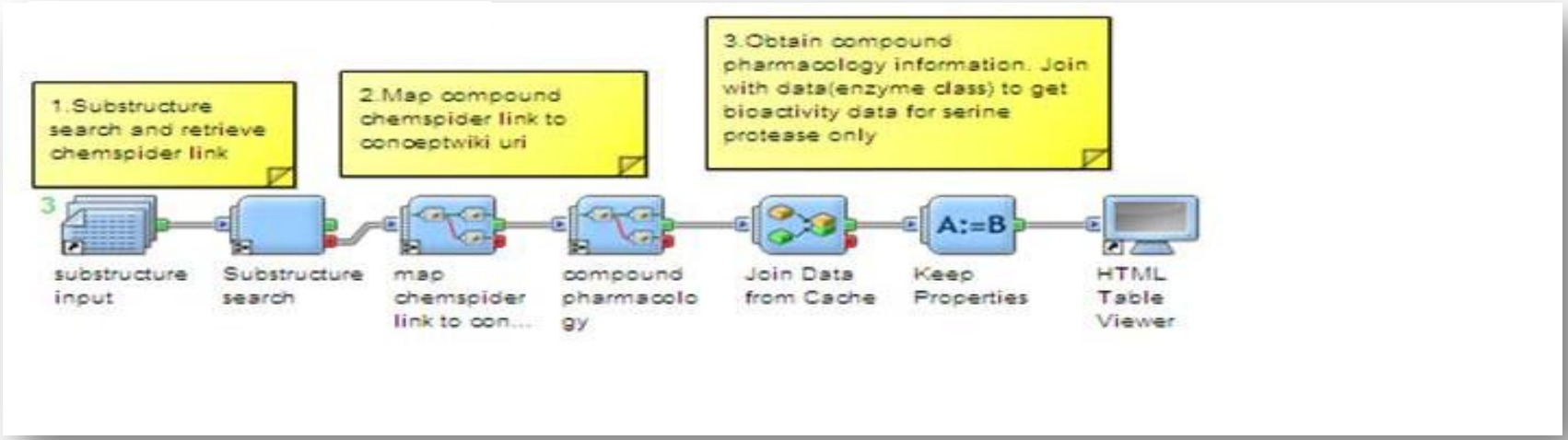


Open PHACTS
Open Pharmacological Space

Open PHACTS

Open Pharmacological Space





Open PHACTS

Open Pharmacological Space

Chrome File Edit View History Bookmarks Window Help 100% Sa

Open PHACTS - YouTube

www.youtube.com/user/OpenPHACTS/videos

Playlists

What to watch
My subscriptions
Social 44

SUBSCRIPTIONS

OpenPHACTS

Browse channels
Manage subscriptions

openphacts.org

Open PHACTS

Subscribed 29

1 in your circles subscribed

Videos Discussion About

Uploads Date added (newest - oldest)

Helium for Excel - using the Open PHACTS API

20 views 3 weeks ago

PharmaTrek: Polypharmacology Browser

288 views 3 months ago

The Open PHACTS/SciBite iPharm Demo App (version 0.01)

67 views 4 months ago

Open PHACTS Explorer - Search Compounds by Structure

69 views 4 months ago

Open PHACTS Explorer - View Pharmacology by Target

56 views 4 months ago

Open PHACTS Explorer - Searching Compounds by Name

91 views 4 months ago

Open PHACTS Explorer - Search Compounds by Name

114 views 4 months ago

Open PHACTS Explorer - View Pharmacology by Compound

43 views 4 months ago

Open PHACTS Explorer - View Pharmacology by Enzyme Family

Open PHACTS Explorer - Search for Target Information by Name

33 views 4 months ago

The Open PHACTS Explorer

610 views 4 months ago

Open PHACTS 6 Month Lashup Demo

1,173 views 1 year ago

www.youtube.com/user/OpenPHACTS

Open PHACTS
Open Pharmacological Space

A Precompetitive Knowledge Framework

Pharma Needs

- Sustainability
  Stability
  Security
- Management / Governance
- Data Mining Services/Algorithms
- Community KD Innovation

Integration

- Mapping & Populating
- Architecture
- Interfaces & Services

Inputs

- Vocabularies & Identifiers (URIs)
- Content Structured & Unstructured

Leveraging Our Community

Open PHACTS
Open Pharmacological Space

Associated partners

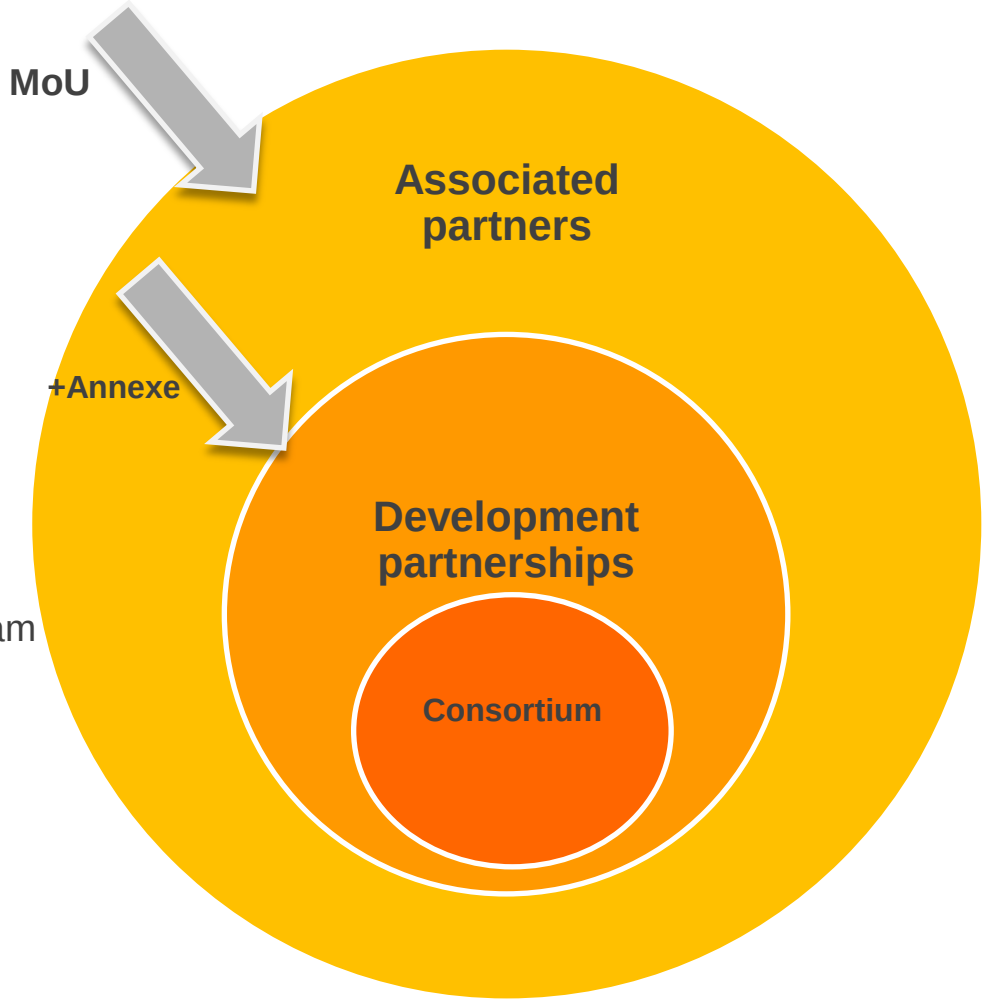
Organisations, most will join here
Support, information
Exchange of ideas, data, technology
Opportunities to demo at community webinars
Need MoU

Development partnerships

Influence on API developments
Opportunities to demo ideas & use cases to core team
Need MoU and annexe

Consortium

28 current members

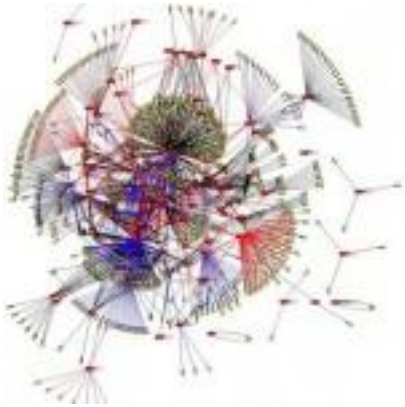


Sustaining Impact

Open PHACTS
Open Pharmacological Space

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

- …and more resource for future development



Open PHACTS
Open Pharmacological Space

The Open PHACTS community ecosystem

Pfizer
universität wien
DTU
Universität Hamburg
BioSolve
accelrys
ACD/Labs
AUREUS SCIENCES
bel
CERTARA
BIO2RDF
Parc de Salut MAR
LUMC
Leids Universitair Medisch Centrum
RSC Advancing the Chemical Sciences
VU University Amsterdam
codeN
Concept Web Alliance
eagle
ENTAGEN
ACCELERATING INSIGHT
CODDES
Erasmus MC
Genetta SOFT
cnio
The University of Manchester
Maastricht University
AQnowledge
Semantics for Science
Global Health Equity Foundation
GVK BIO
ian Harrow Consulting
IOFA Consult
inte:ligand
USC
universitätbonn
AstraZeneca
gsk GlaxoSmithKline
School of Informatics and Computing
Indiana University Bloomington
IRB Barcelona
IUPHAR-DB
The IUPHAR Database of Receptors and Channels
Johannes Gutenberg Universität Mainz
LDBC
MGH 1811
mobileflow
ESTEVE
NOVARTIS
MERCK
Lundbeck
The National Center for Biomedical Ontology
Numedicus
ontoforce
orphanet
Lilly
nbic
SIB Swiss Institute of Bioinformatics
connecteddiscovery
Parc Científic de Barcelona
PHORTOS CONSULTANTS
PRESTWICK CHEMICAL
The medicinal chemistry company
Sage BIONETWORKS
The Scripps Research Institute
EMBL-EBI
janssen
OPENLINK SOFTWARE
Making Technology Work For You
SEQUENOMICS
syngenta
thehyve
THOMSON REUTERS
rdfinformatics
VIVO
connect • share • discover

Open PHACTS

Open Pharmacological Space

A UK-based not-for-profit member owned company

Publish

Maintain and develop the Open PHACTS Platform

'Research based' organisation

Many different types of member

Develop and contribute to data standards

Innovate

Develop and build the Open PHACTS community

Provide stable API services

Become partner in other consortia

Precompetitive

Develop and build new value-added analytical methods

Promote and build interoperable data beyond Open PHACTS

Open PHACTS
Open Pharmacological Space

Becoming part of the Open PHACTS Foundation

A UK-based not-for-profit member owned company

Members

- membership offers early access to platform updates and releases
- the opportunity to steer research and development directions
- receive technical support
- work with the ecosystem of developers and semantic data integrators around Open PHACTS

- tiered membership
- familiar business and governance model

Open PHACTS Project Partners

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

pmu@openphacts.org

Open PHACTS
Open Pharmacological Space

- **Access to a wide range of interconnected data** – easily jump between pharmacology, chemistry, disease, pathways and other databases without having to perform complex mapping operations
- **Query by data type, not by data source** (“Protein Information” not “Uniprot Information)
- **API queries that seamlessly connect data** (for instance the Pharmacology query draws data from Chembl, ChemSpider, ConceptWiki and Drugbank)
- **Strong chemistry representation** – all chemicals reprocessed via Open PHACTS chemical registry to ensure consistency across databases
- **Built using open community standards**, not an ad-hoc solution. Developed in conjuction with 8 major pharma (so your app will speak their language!)
- **Simple, flexible data-joining** (join compound data ignoring salt forms, join protein data ignoring species)
- **Provenance everywhere** – every single data point tagged with source, version, author, etc
- **Nanopublication-enabled**. Access to a rich dataset of established and emerging biomedical “assertions”
- **Professionally Hosted** (Continually Monitored)
- **Developer-friendly** JSON/XML methods. **Consistent API** for multiple services
- **Seamless data upgrades.** We manage updates so you don’t have to
- **Community-curation** tools to enhance and correct content
- Access to **a rich application network** (many different App builders)
- **Toolkits** to support many different languages, workflow engines and user applications
- **Private and secure**, suitable for confidential analyses
- **Active & still growing** through a unique public-private partnership