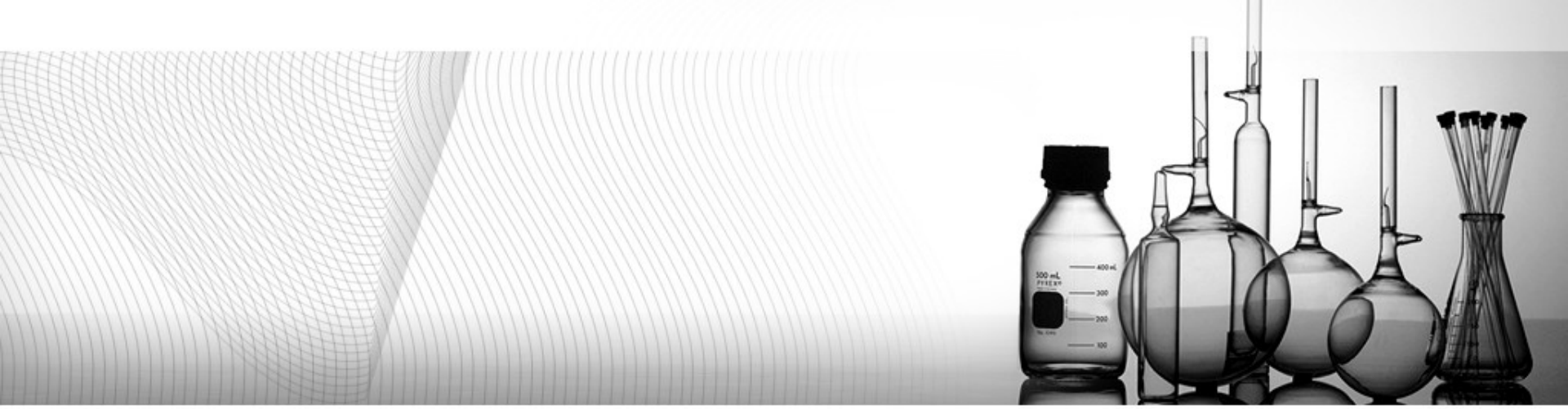


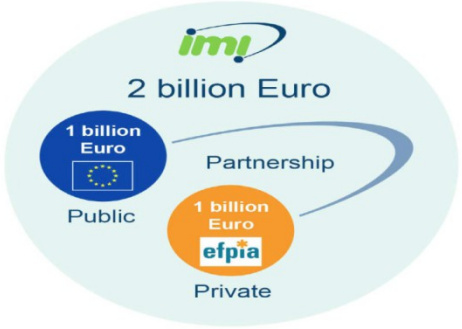
Open PHACTS
Open Pharmacological Space

The Open PHArmacological Concepts Triple Store

Egon Willighagen
Dept of Bioinformatics – BiGCaT, Maastricht University
@egonwillighagen, #oteu13

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")…*
- Delivering services to support

- Work split into clusters:
  - Technical Build (*focus here)*
  - Scientific Drive
  - Community & Sustainability

The Project

Open PHACTS Project Partners

www.openphacts.org

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

Research Questions

Open PHACTS
Open Pharmacological Space

Number	sum	Nr of 1	Question
15	**12**	**9**	**All oxido,reductase 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

What do we need?

Open PHACTS
Open Pharmacological Space

ChEMBL | DrugBank | Gene Ontology | Wikipathways | GeneGo

ChEBI | Uniprot | UMLS | GVKBio

ConceptWiki | ChemSpider | TrialTrove | TR Integrity

The Open PHACTS infrastructure can support many different domains & questions

Quantitative Data Challenges

Open PHACTS
Open Pharmacological Space

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

>5000 types

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
IC50	p.p.m.	51
IC50	ppm	36
IC50	uM-1	25
IC50	nM kg-1	25
IC50	milliequivalent	22
IC50	kJ m-2	20

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

~ 100 units

Data Licensing Solution

Open PHACTS
Open Pharmacological Space

Chose John Wilbanks as consultant

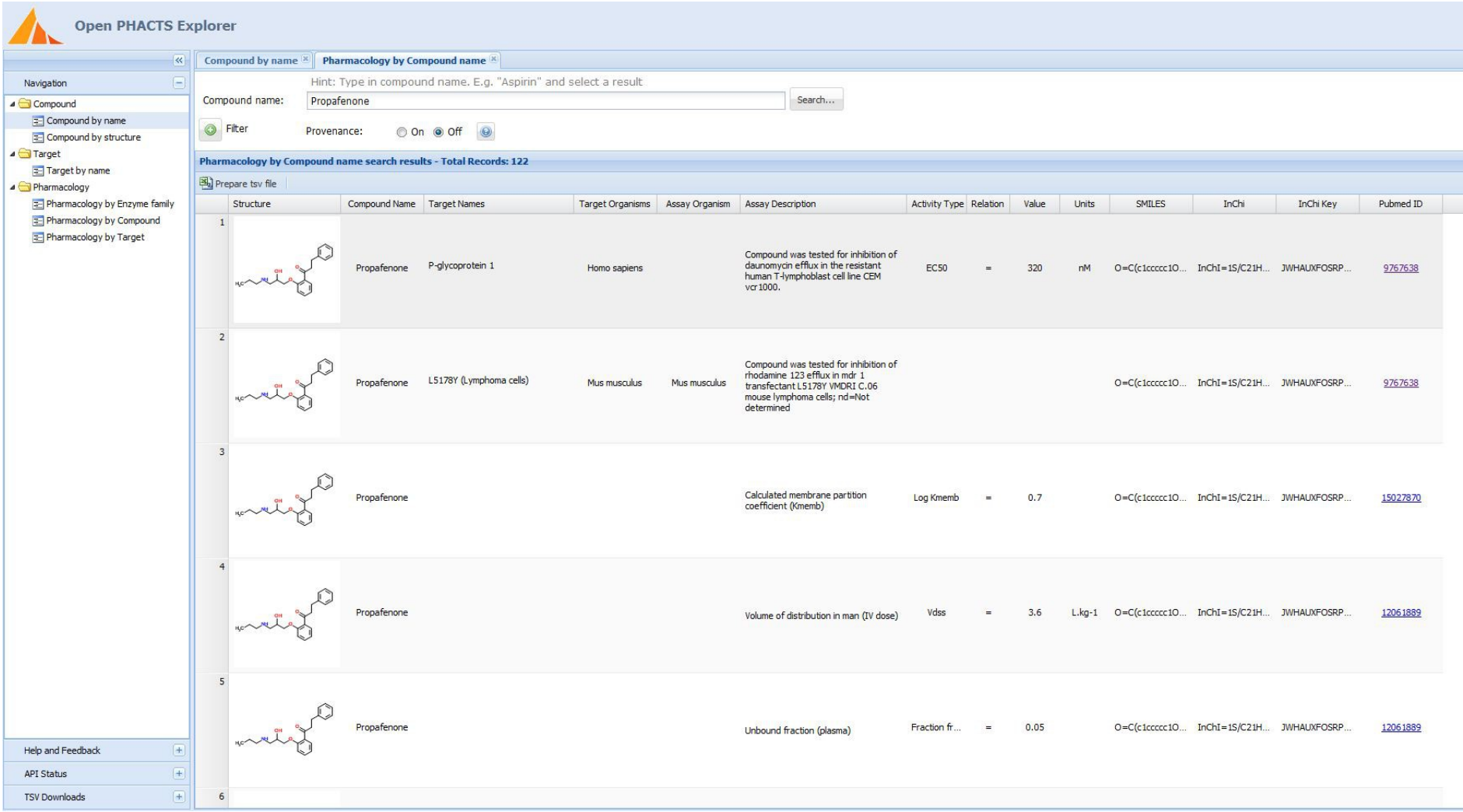
A framework built around STANDARD well-understood Creative Commons licences – and how they interoperate

Deal with the problems by:

- **Interoperable licences**
- **Appropriate terms**
- **Declare expectations to users and data publishers**
- **One size won‘t fit all requirements**

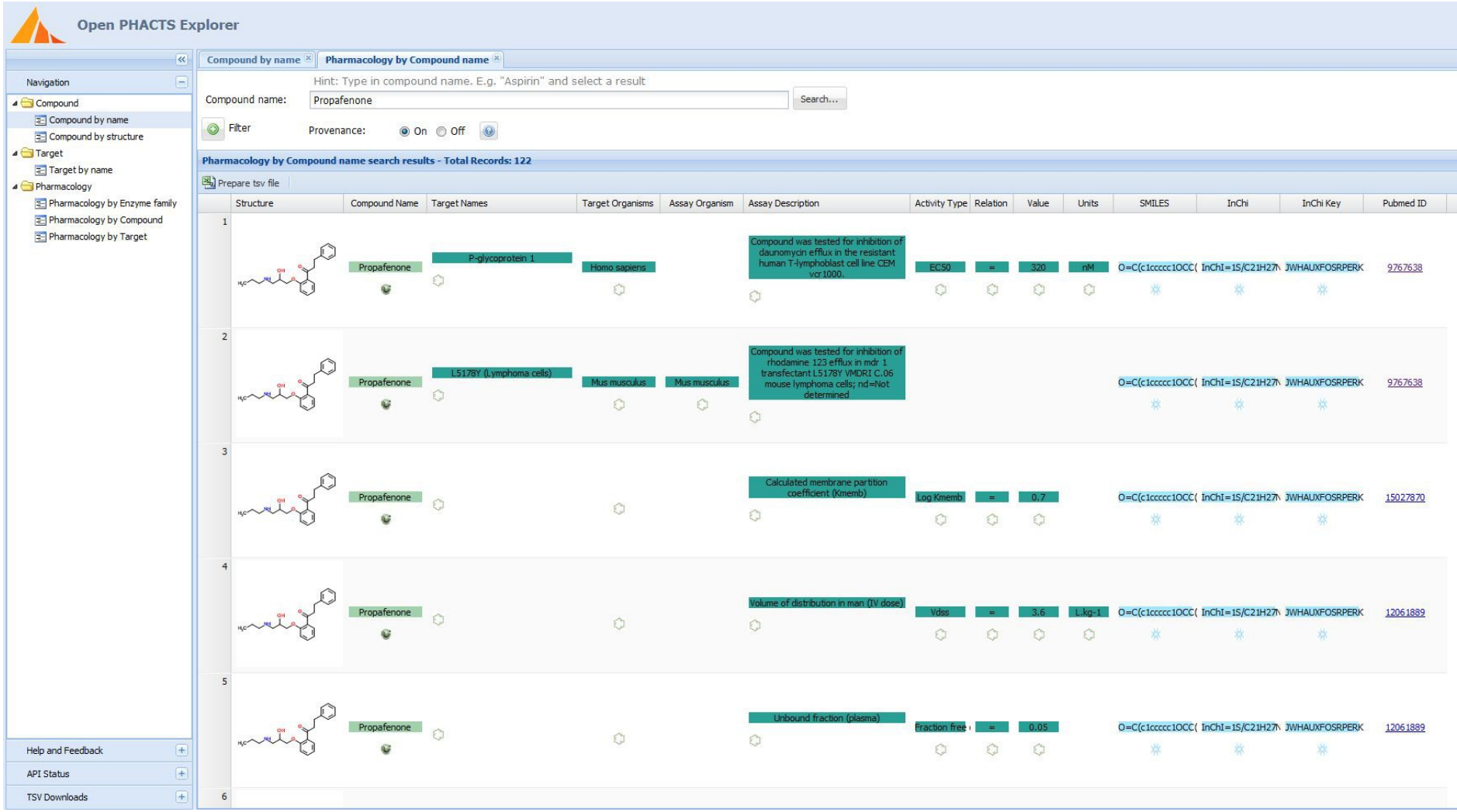
Compatibility chart		Terms that may be used for a derivative work or adaptation						
		BY	BY-NC	BY-NC-ND	BY-NC-SA	BY-ND	BY-SA	PD
Status of original work	PD	■	■	■	■	■	■	■
	BY	■	■	■	■	■	■	
	BY-NC		■	■	■			
	BY-NC-ND							
	BY-NC-SA				■			
	BY-ND							
	BY-SA						■	

Open PHACTS
Open Pharmacological Space



http://explorer.openphacts.org/

Open PHACTS
Open Pharmacological Space

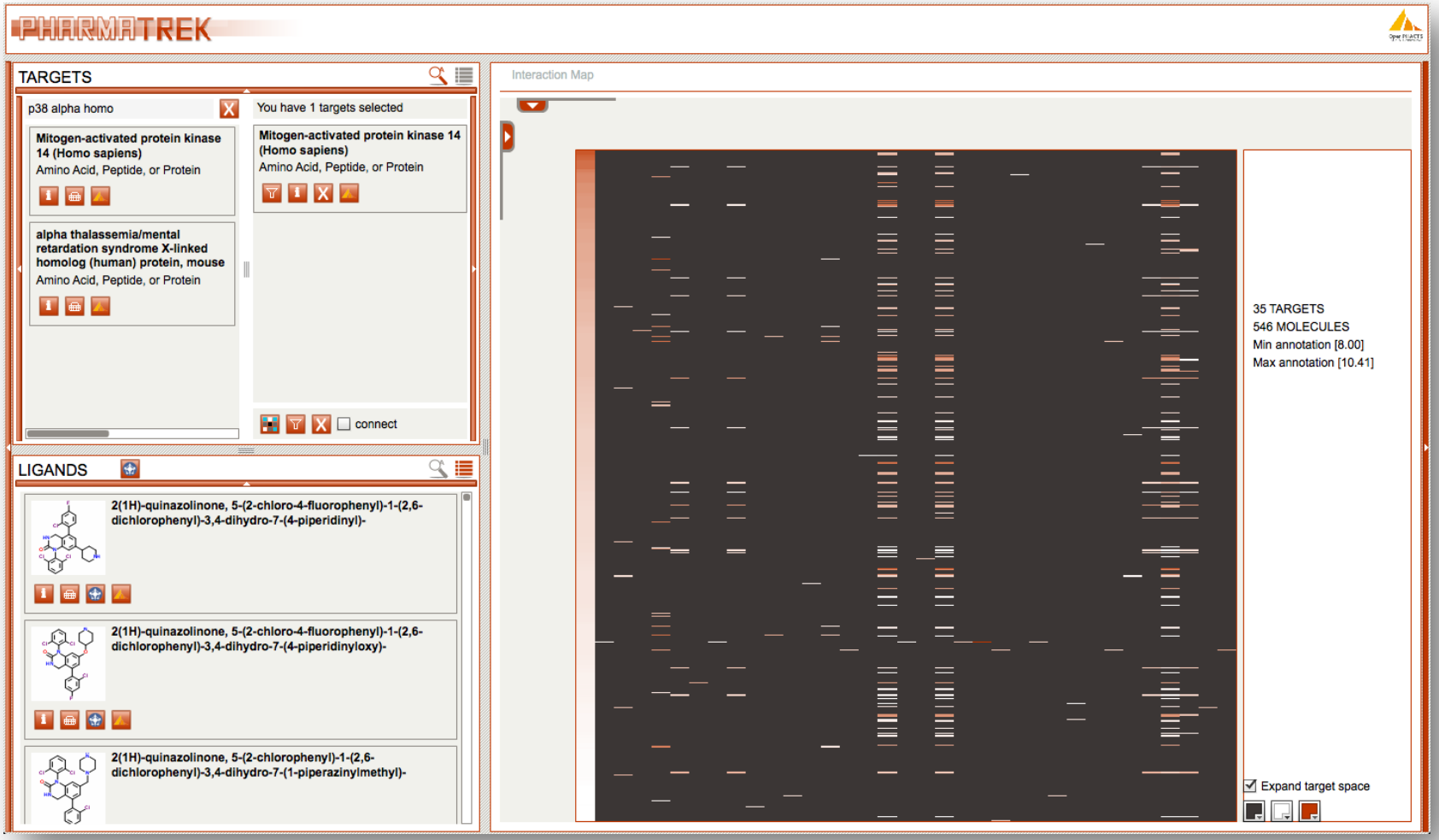


Example applications

Open PHACTS
Open Pharmacological Space

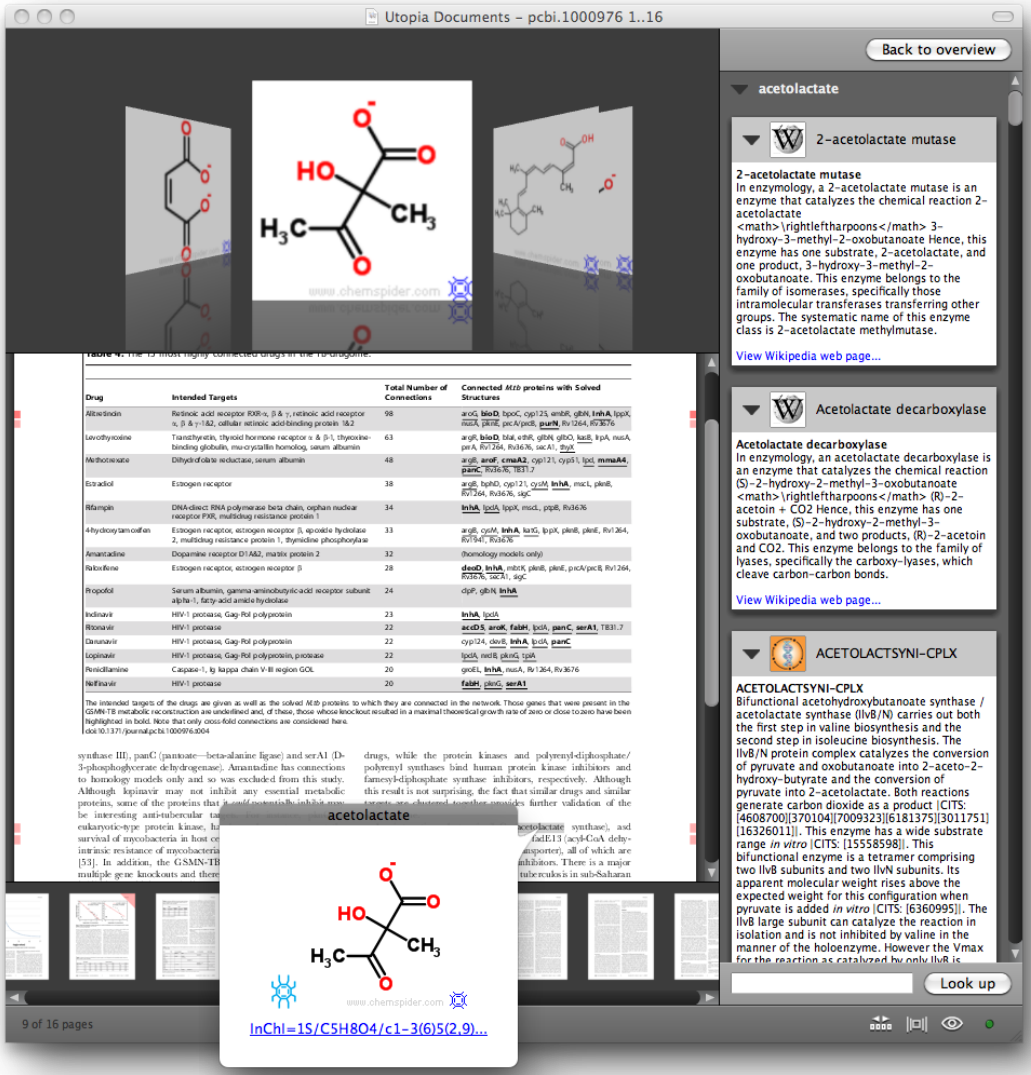
Advanced analytics	
ChemBioNavigator	Navigating at the interface of chemical and biological data with sorting and plotting options
TargetDossier	Interconnecting Open PHACTS with multiple target centric services. Exploring target similarity using diverse criteria
PharmaTrek	Interactive Polypharmacology 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

Open PHACTS
Open Pharmacological Space



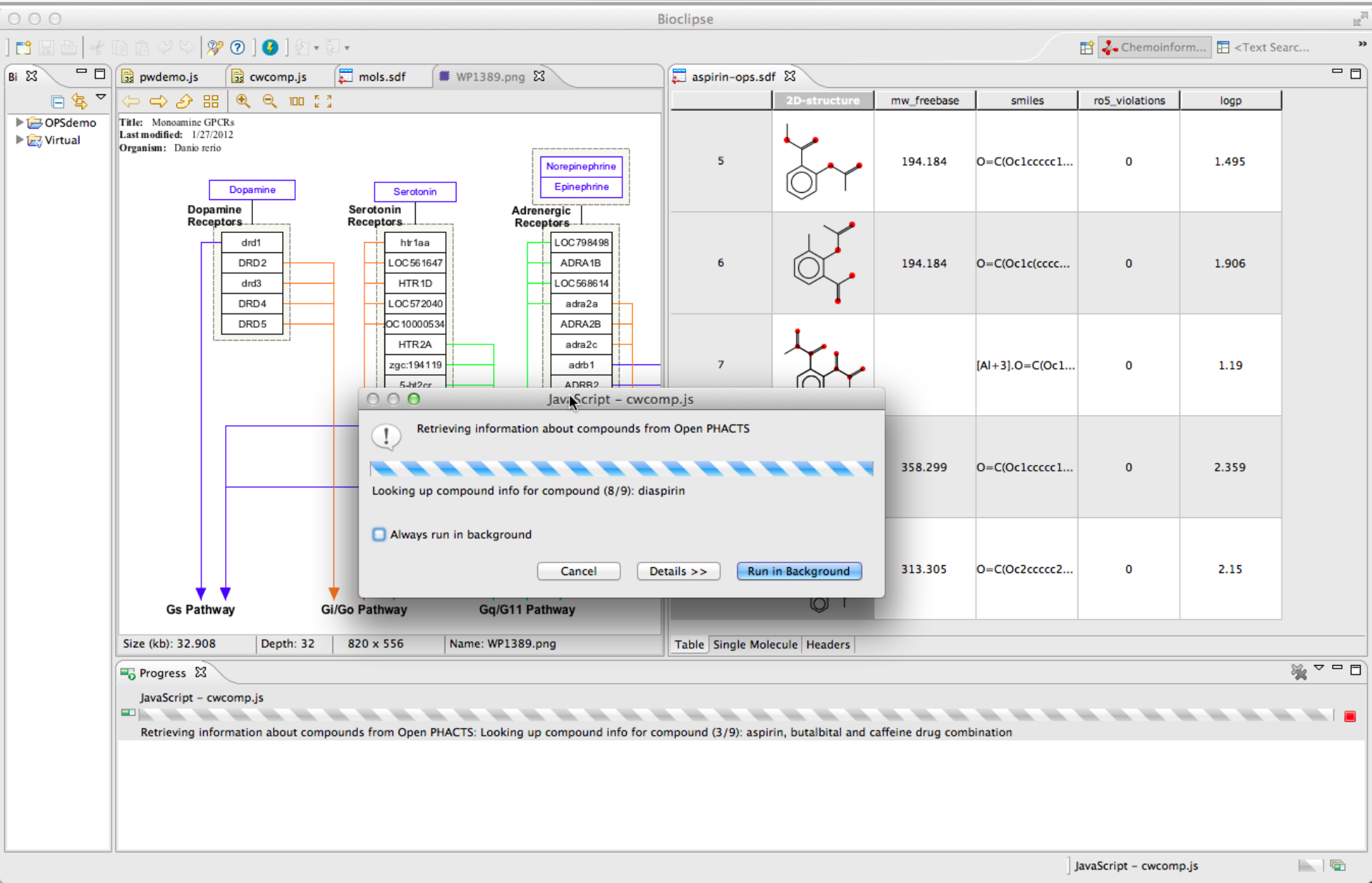
Open PHACTS

Open Pharmacological Space

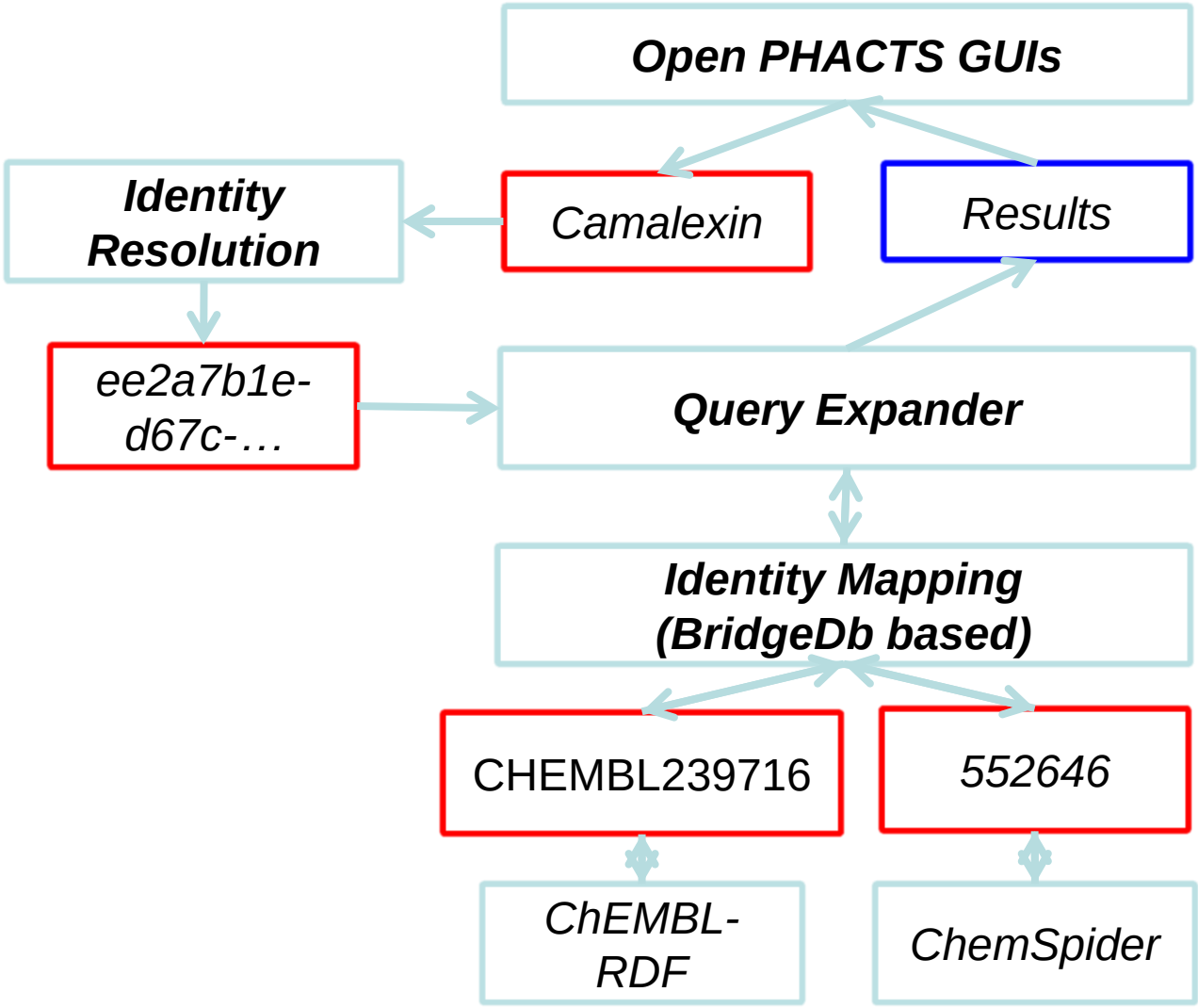


Bioclipse

Open PHACTS
Open Pharmacological Space



Open PHACTS
Open Pharmacological Space

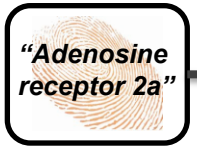


Open PHACTS
Open Pharmacological Space

Apps

Core Platform

Identity Resolution Service



Identifier Management Service

P12374
EC2.43.4
CS4532

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

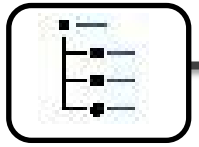
Semantic Workflow Engine

Data Cache
(Virtuoso Triple Store)

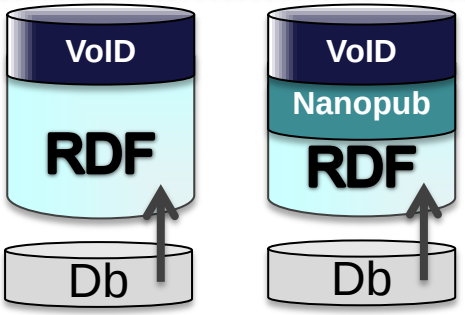
Domain Specific Services

Chemistry
Registration
Normalisation
& Q/C

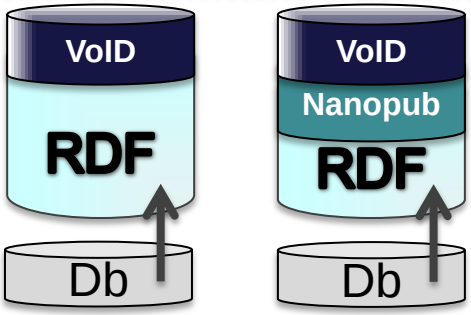
Indexing



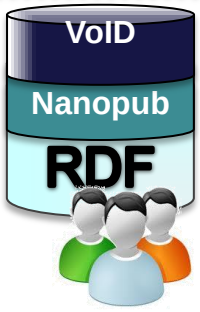
Public Ontologies



Public Content



Commercial



User Annotations

RDF Guidelines

Open PHACTS Working Draft

Guidelines for exposing data as RDF in Open PHACTS

Open PHACTS Working Draft 23 August 2012

This version:
http://www.openphacts.org/specs/2012/WD-rdfguide-20120823/
Latest published version:
http://www.openphacts.org/specs/rdfguide/
Latest editor's draft:
http://www.bigcat.unimaas.nl/~egonw/rdfguide/
Previous version:
none
Editor:

Step 9: who, what, when, why, how

Open PHACTS Working Draft

Dataset Descriptions for the Open Pharmacological Space

Open PHACTS Working Draft 19 October 2012

This version:
http://www.openphacts.org/specs/2012/WD-datadesc-20121019/
Latest published version:
http://www.openphacts.org/specs/datadesc/
Latest editor's draft:
http://www.cs.man.ac.uk/~graya/ops/2012/ED-datadesc/
Previous version:
http://www.openphacts.org/specs/2012/WD-datadesc-20120816/
Editor:

Step 9: who, what, when, why, how

```
# Description of the ChemSpider dataset
:chemSpiderDataset
# General metadata
        a void:Dataset;
        dcterms:title "ChemSpider"@en;
        dcterms:description "ChemSpider's Public Dataset"@en;
        foaf:homepage <http://rdf.chemspider.com/>;
        foaf:page <http://www.chemspider.com/>;
        dcterms:license <http://www.chemspider.com/Disclaimer.aspx>;
        void:uriSpace "http://rdf.chemspider.com/"^^xsd:string;
#Provenance
        dcterms:publisher <http://www.chemspider.com/>;
        dcterms:created "2007-03-01T00:00:00"^^xsd:dateTime;
        dcterms:modified "2012-10-16T00:00:00"^^xsd:dateTime;#Subsets
        void:subset :chemSpiderDataset_chembl_subset,:chemSpiderDataset_drugbank_subset;
#Vocabularies, topics, resources
        void:vocabulary <http://purl.org/dc/elements/1.1/>,
                <http://purl.org/dc/terms/>,
                <http://www.openarchives.org/ore/terms/>,
                <http://www.polymerinformatics.com/ChemAxiom/ChemDomain.owl#>,
                <http://xmlns.com/foaf/0.1/>;
        dcterms:subject <http://dbpedia.org/resource/Molecule>;
        void:exampleResource <http://rdf.chemspider.com/2157>;
#Dataset Access
        void:sparqlEndpoint <http://rdf.chemspider.com/sparql>;
#Update Frequency
        voag:frequencyOfChange freq:continuous;
#Other Metadata
        # Technical features
        void:feature <http://www.w3.org/ns/formats/RDF_XML>;
        # Dataset statistics
        void:triples "1157624328"^^xsd:nonNegativeInteger;
```

Community

Open PHACTS and the scientific 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

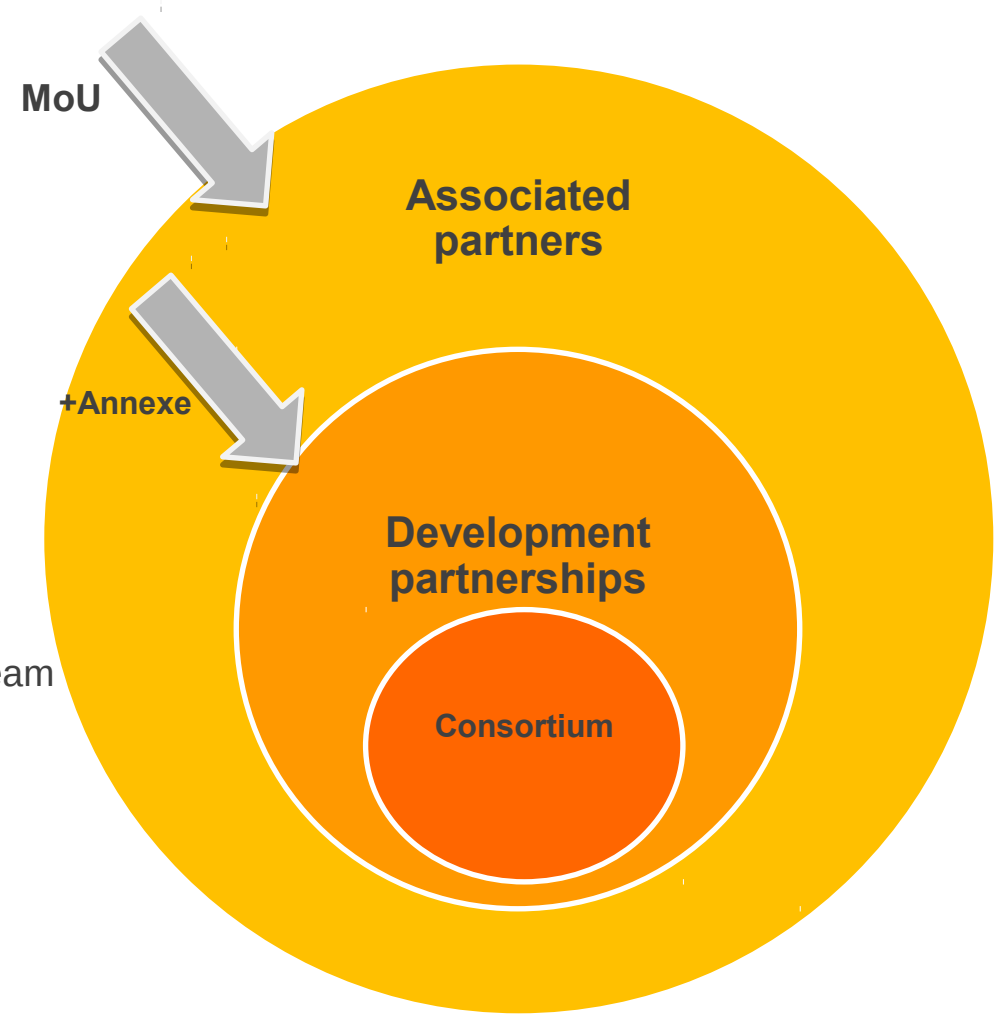
Influence on API developments

Opportunities to demo ideas & use cases to core team

Need MoU and annexe

Consortium

28 current members



Associate Partners

Open PHACTS
Open Pharmacological Space

Sequenomics

Wrap up

Open PHACTS
Open Pharmacological Space

- Open Source and Open Data solutions
  - Used and Developed
- Specifications
  - http://www.openphacts.org/specs/
- Solutions for pharma

- "Shaping the future of Open PHACTS"
  5th Community Workshop, Amsterdam 10-11 Oct

www.openphacts.org