

# The **Open PHA**rmacological **C**oncepts **T**riple **S**to**r**e

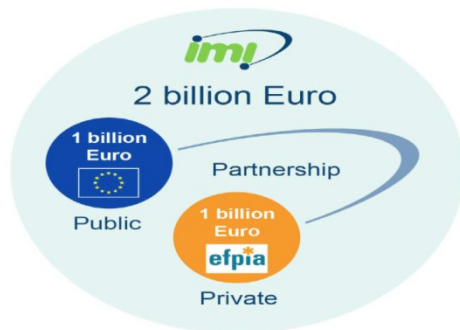
Egon Willighagen

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](http://www.openphacts.org)

**Pfizer Limited – Coordinator**

**Universität Wien – Managing entity**

Technical University of Denmark

University of Hamburg, Center for Bioinformatics

BioSolveIT GmbH

Consorti 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

Number	sum	Nr of 1	Question
15	12	9	<b>All oxido,reductase inhibitors active &lt;100nM in both human and mouse</b>
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?



"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  $<1 \mu\text{M}$ "

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

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**

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						■	



Navigation

- Compound
  - Compound by name
  - Compound by structure
- Target
  - Target by name
- Pharmacology
  - Pharmacology by Enzyme family
  - Pharmacology by Compound
  - Pharmacology by Target

Help and Feedback (+)

API Status (+)

TSV Downloads (+)

Compound by name: **Pharmacology by Compound name**

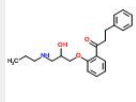
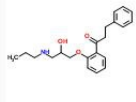
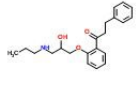
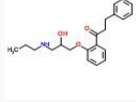
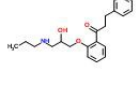
Hint: Type in compound name. E.g. "Aspirin" and select a result

Compound name:  Search...

Filter Provenance:  On  Off

Pharmacology by Compound name search results - Total Records: 122

Prepare tsv file

	Structure	Compound Name	Target Names	Target Organisms	Assay Organism	Assay Description	Activity Type	Relation	Value	Units	SMILES	InChi	InChi Key	Pubmed ID
1		Propafenone	P-glycoprotein 1	Homo sapiens		Compound was tested for inhibition of daunomycin efflux in the resistant human T-lymphoblast cell line CEM vcr 1000.	EC50	=	320	nM	O=C(c1ccccc1O...	InChi=1S/C21H...	JWHAUXFOSRP...	<a href="#">9767638</a>
2		Propafenone	L5178Y (Lymphoma cells)	Mus musculus	Mus musculus	Compound was tested for inhibition of rhodamine 123 efflux in the resistant human T-lymphoblast cell line CEM vcr 1000.					O=C(c1ccccc1O...	InChi=1S/C21H...	JWHAUXFOSRP...	<a href="#">9767638</a>
3		Propafenone				Calculated membrane partition coefficient (Kmemb)	Log Kmemb	=	0.7		O=C(c1ccccc1O...	InChi=1S/C21H...	JWHAUXFOSRP...	<a href="#">15027870</a>
4		Propafenone				Volume of distribution in man (IV dose)	Vdss	=	3.6	L.kg-1	O=C(c1ccccc1O...	InChi=1S/C21H...	JWHAUXFOSRP...	<a href="#">12061889</a>
5		Propafenone				Unbound fraction (plasma)	Fraction fr...	=	0.05		O=C(c1ccccc1O...	InChi=1S/C21H...	JWHAUXFOSRP...	<a href="#">12061889</a>
6														





## Open PHACTS Explorer

- Navigation
- Compound
    - Compound by name
    - Compound by structure
  - Target
    - Target by name
  - Pharmacology
    - Pharmacology by Enzyme family
    - Pharmacology by Compound
    - Pharmacology by Target

Compound by name **Pharmacology by Compound name**

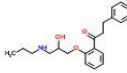
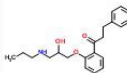
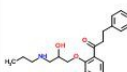
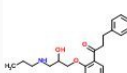
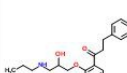
Hint: Type in compound name. E.g. "Aspirin" and select a result

Compound name:

Filter Provenance:  On  Off

Pharmacology by Compound name search results - Total Records: 122

Prepare tsv file

Structure	Compound Name	Target Names	Target Organisms	Assay Organism	Assay Description	Activity Type	Relation	Value	Units	SMILES	InChi	InChi Key	Pubmed ID
	Propafenone	P-glycoprotein 1	Homo sapiens		Compound was tested for inhibition of daunomycin efflux in the resistant human T-lymphoblast cell line CEM vcr1000.	EC50	=	320	nM	<chem>O=C(c1ccccc1OCC)</chem> InChi=1S/C21H27	JWHAUXFOSRPERK		9767638
	Propafenone	L5178Y (Lymphoma cells)	Mus musculus	Mus musculus	Compound was tested for inhibition of rhodamine 123 efflux in mdr 1 transfectant L5178Y VMDRI C.06 mouse lymphoma cells; nd=Not determined					<chem>O=C(c1ccccc1OCC)</chem> InChi=1S/C21H27	JWHAUXFOSRPERK		9767638
	Propafenone				Calculated membrane partition coefficient (Kmemb)	Log Kmemb	=	0.7		<chem>O=C(c1ccccc1OCC)</chem> InChi=1S/C21H27	JWHAUXFOSRPERK		1502720
	Propafenone				Volume of distribution in man (IV dose)	Vdss	=	3,6	L.kg-1	<chem>O=C(c1ccccc1OCC)</chem> InChi=1S/C21H27	JWHAUXFOSRPERK		12061889
	Propafenone				Unbound fraction (plasma)	Fraction free	=	0.05		<chem>O=C(c1ccccc1OCC)</chem> InChi=1S/C21H27	JWHAUXFOSRPERK		12061889

- Help and Feedback
- API Status
- TSV Downloads

## 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



**PHARMATREK**


**TARGETS** 🔍 📄

p38 alpha homo ✕ You have 1 targets selected

**Mitogen-activated protein kinase 14 (Homo sapiens)**  
Amino Acid, Peptide, or Protein

📄 📄 📄

**Mitogen-activated protein kinase 14 (Homo sapiens)**  
Amino Acid, Peptide, or Protein

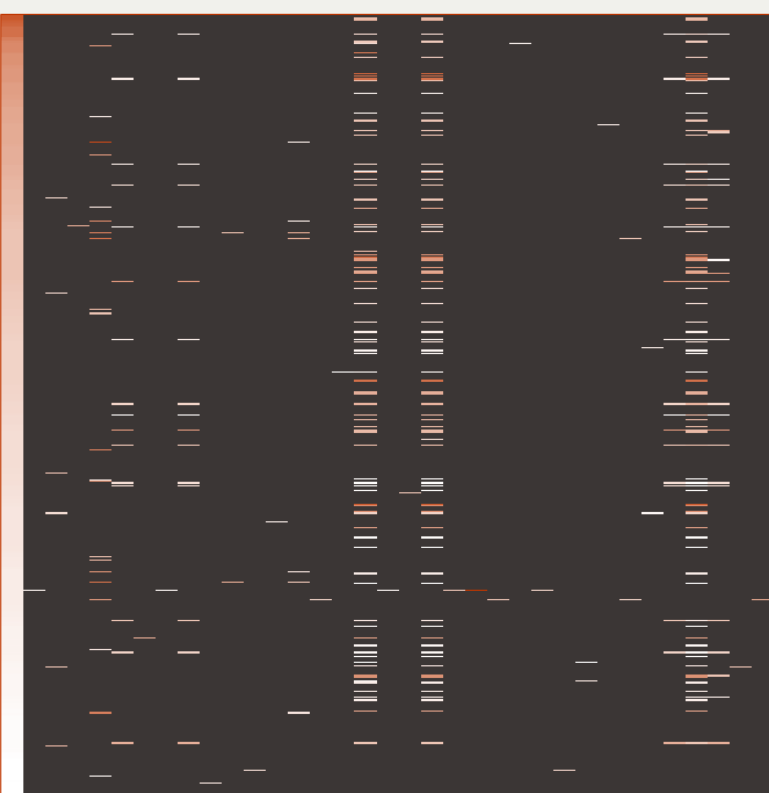
📄 📄 ✕ 📄

**alpha thalassemia/mental retardation syndrome X-linked homolog (human) protein, mouse**  
Amino Acid, Peptide, or Protein

📄 📄 📄

📄 📄 ✕  connect

**Interaction Map**




35 TARGETS  
546 MOLECULES  
Min annotation [8.00]  
Max annotation [10.41]


Expand target space

📄 📄 📄


**LIGANDS** 🔍 📄

 **2(1H)-quinazolinone, 5-(2-chloro-4-fluorophenyl)-1-(2,6-dichlorophenyl)-3,4-dihydro-7-(4-piperidiny)-**

📄 📄 📄 📄

 **2(1H)-quinazolinone, 5-(2-chloro-4-fluorophenyl)-1-(2,6-dichlorophenyl)-3,4-dihydro-7-(4-piperidinyloxy)-**

📄 📄 📄 📄

 **2(1H)-quinazolinone, 5-(2-chlorophenyl)-1-(2,6-dichlorophenyl)-3,4-dihydro-7-(1-piperazinylmethyl)-**



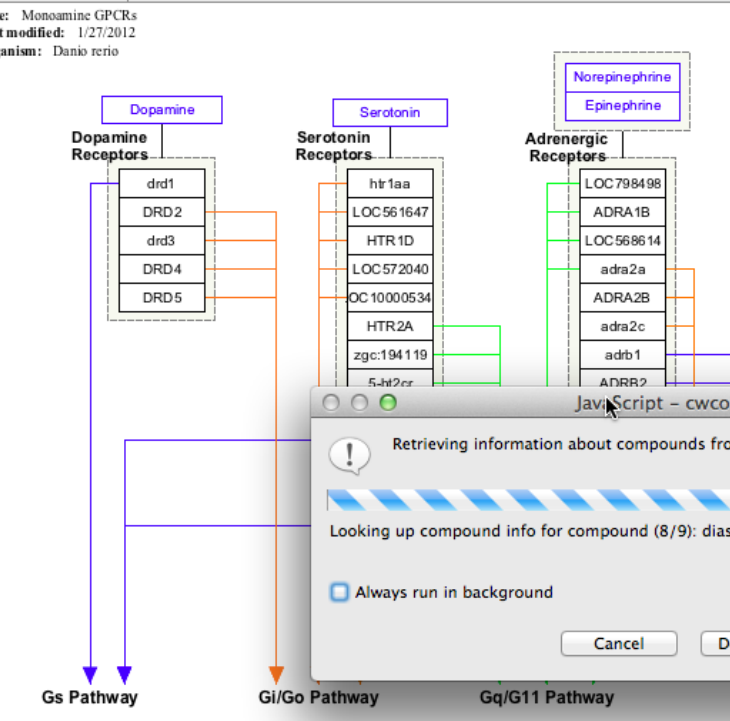
Bioclipse

Chemoinform... <Text Search...

pwdemo.js cwcomp.js mols.sdf WP1389.png aspirin-ops.sdf

OPSdemo Virtual

Title: Monoamine GPCRs  
Last modified: 1/27/2012  
Organism: Danio rerio

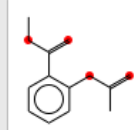
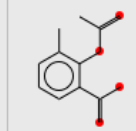
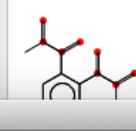


**Dopamine**  
Dopamine Receptors: drd1, DRD2, drd3, DRD4, DRD5

**Serotonin**  
Serotonin Receptors: htr1aa, LOC561647, HTR1D, LOC572040, OC10000534, HTR2A, zgc:194119, 5-htr2c

**Adrenergic**  
Adrenergic Receptors: LOC798498, ADRA1B, LOC568614, adra2a, ADRA2B, adra2c, adrb1, ADRB2

Gs Pathway, Gi/Go Pathway, Gq/G11 Pathway

	2D-structure	mw_freebase	smiles	ro5_violations	logp
5		194.184	O=C(Oc1cccc1...	0	1.495
6		194.184	O=C(Oc1c(cccc...	0	1.906
7			[Al+3].O=C(Oc1...	0	1.19
		358.299	O=C(Oc1cccc1...	0	2.359
		313.305	O=C(Oc2cccc2...	0	2.15

JavaScript - cwcomp.js

Retrieving information about compounds from Open PHACTS

Looking up compound info for compound (8/9): diaspirin

Always run in background

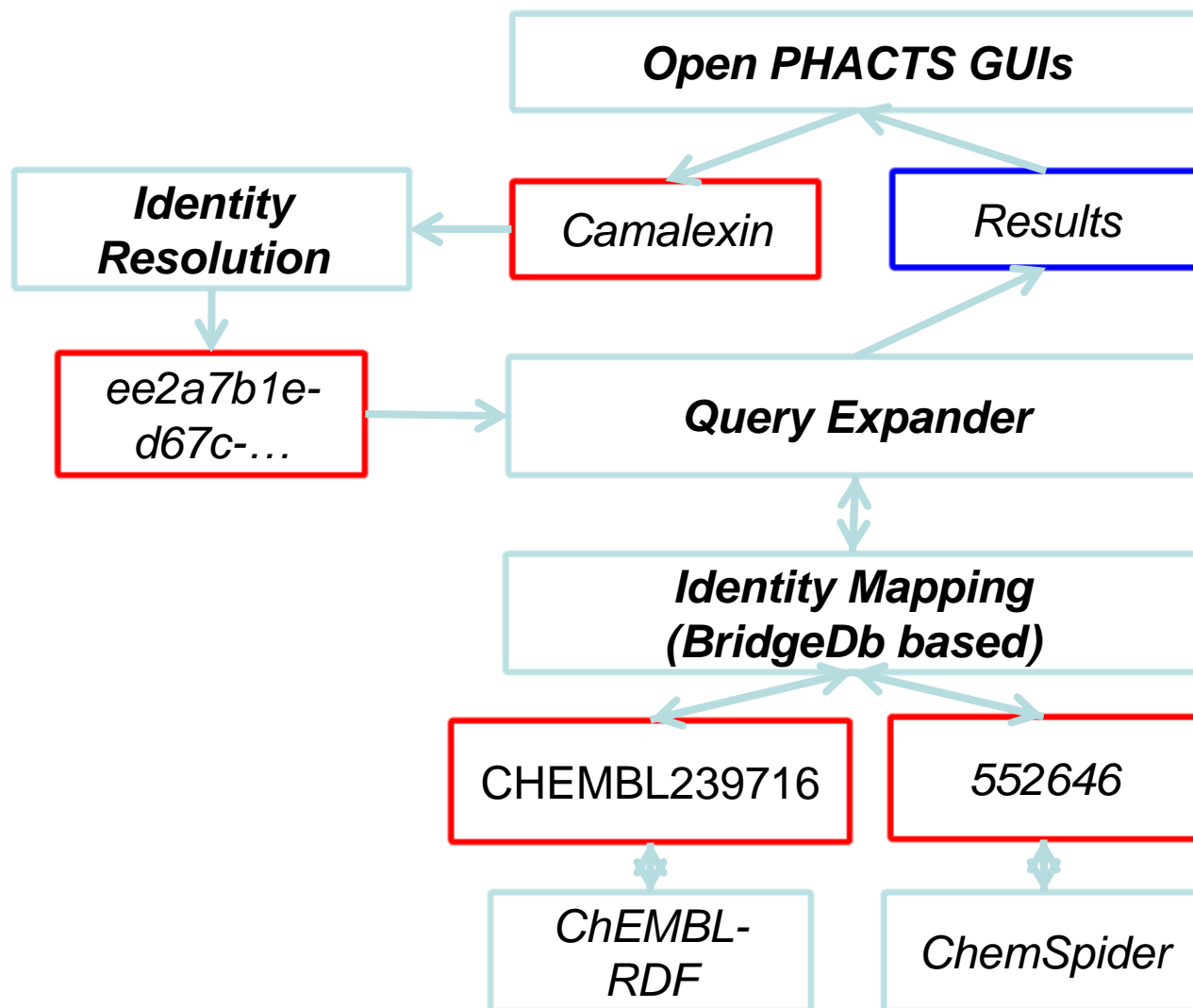
Cancel Details >> Run in Background

Progress JavaScript - cwcomp.js

Retrieving information about compounds from Open PHACTS: Looking up compound info for compound (3/9): aspirin, butalbital and caffeine drug combination

Size (kb): 32.908 Depth: 32 820 x 556 Name: WP1389.png Table Single Molecule Headers

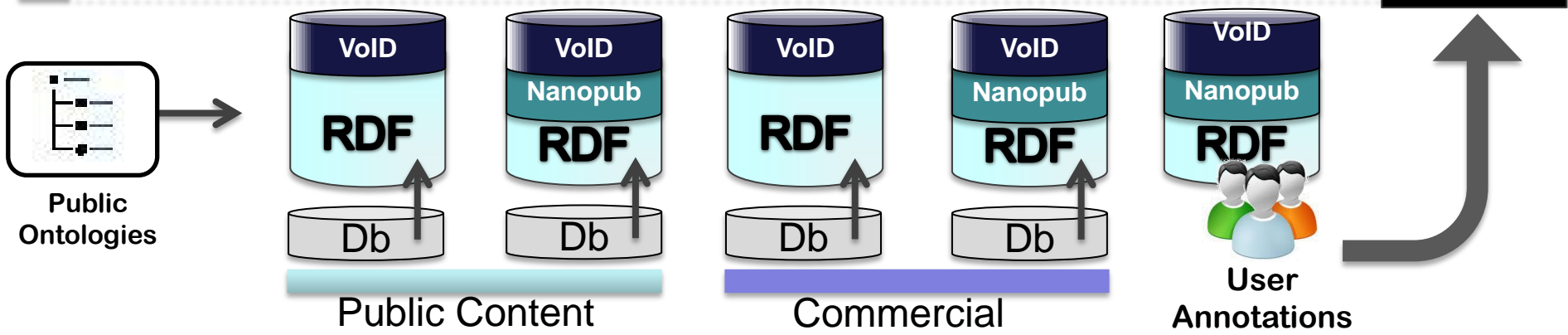
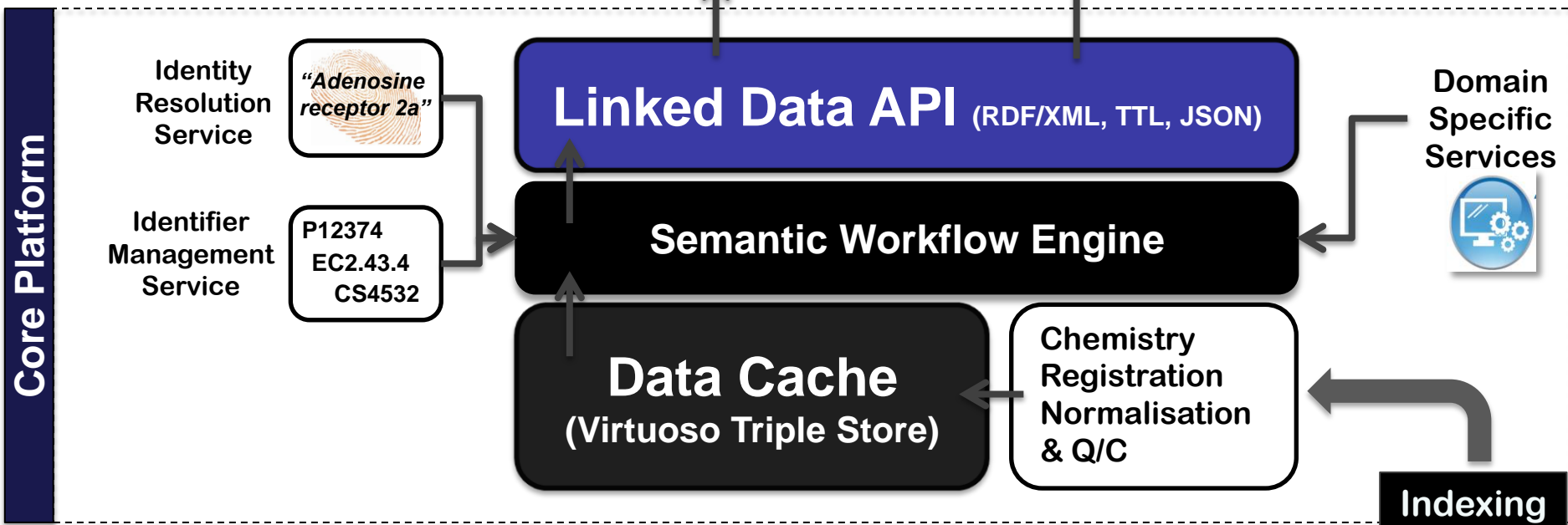
JavaScript - cwcomp.js







# Apps





# 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



## 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

## Associated partners

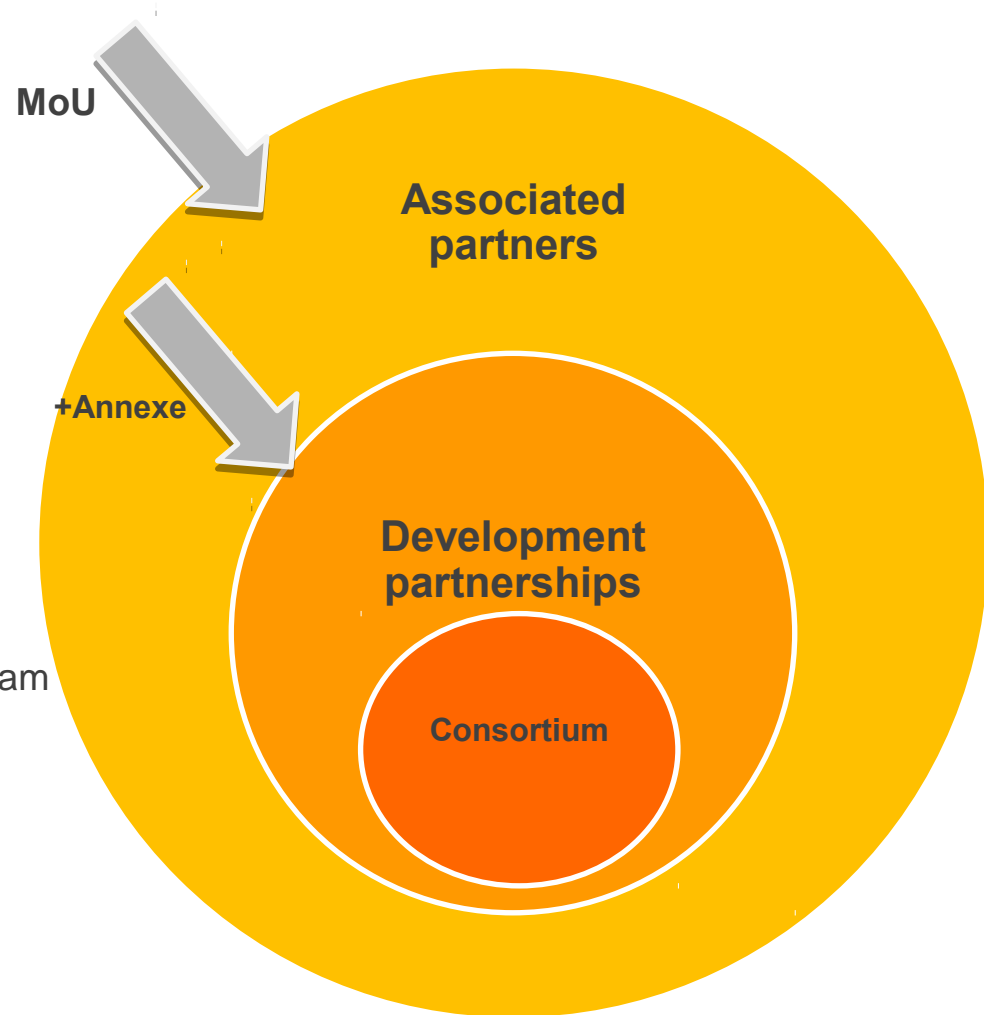
- Organisations, most will join here
- Support, information
- Exchange of ideas, data, technology
- Opportunities to demo at community webinars
- Need MoU

## Development partnerships

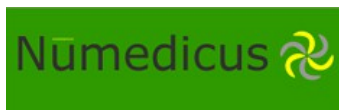
- Influence on API developments
- Opportunities to demo ideas & use cases to core team
- Need MoU and annexe

## Consortium

28 current members



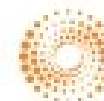
# Associate Partners



INDIANA UNIVERSITY



Sequenomics



THOMSON REUTERS

# Wrap up

- ❖ 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](http://www.openphacts.org)