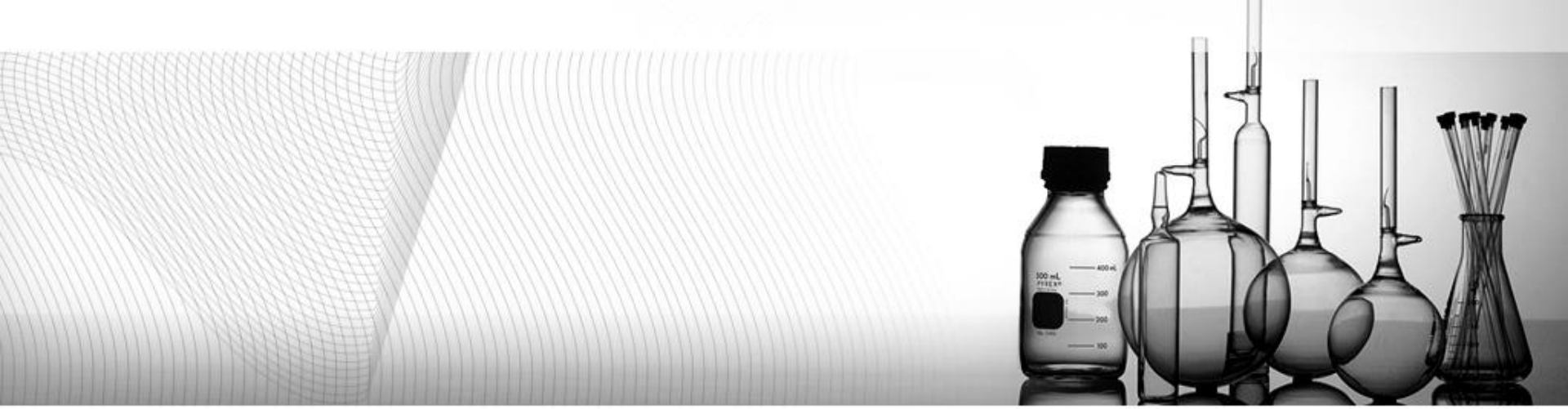


Open PHACTS
Open Pharmacological Space

The Open PHACTS Project: Progress and Future Sustainability

Lee Harland & Bryn Williams-Jones
Open PHACTS / ConnectedDiscovery

Tom Plasterer
AstraZeneca/Open PHACTS Rep

Open PHACTS
Open Pharmacological Space

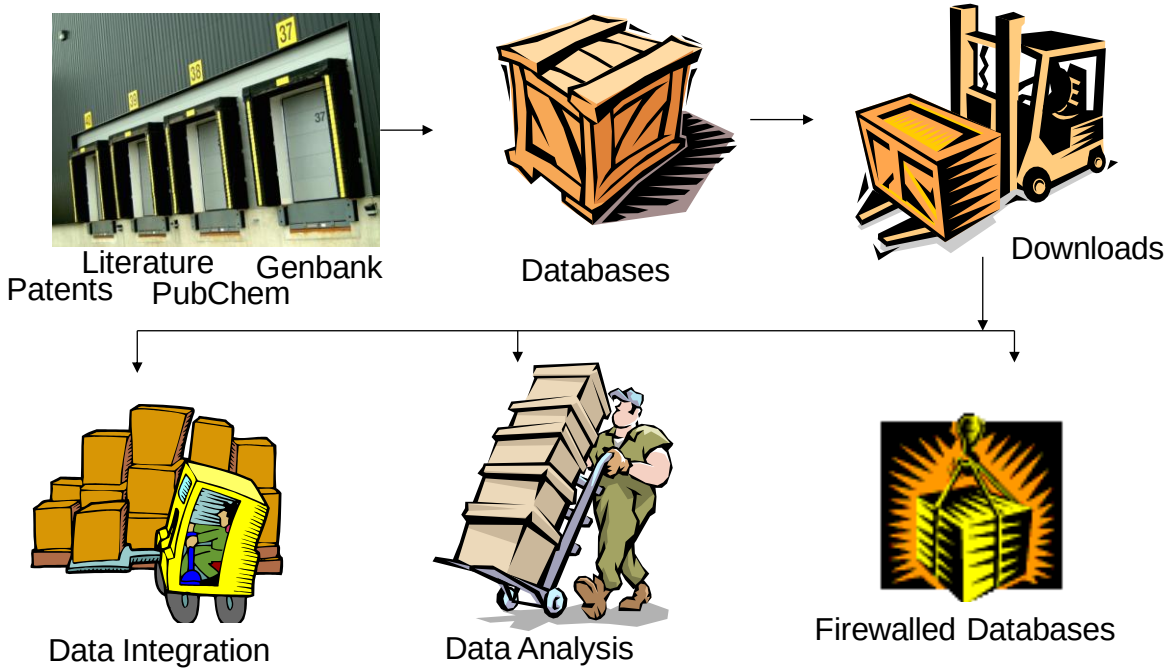
Fundamental issue:

- 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")…*
- Delivering 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 biotechs
- Work split into clusters:
  - Tehnical Build
  - Scientific Drive
  - Community & Sustainability

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 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

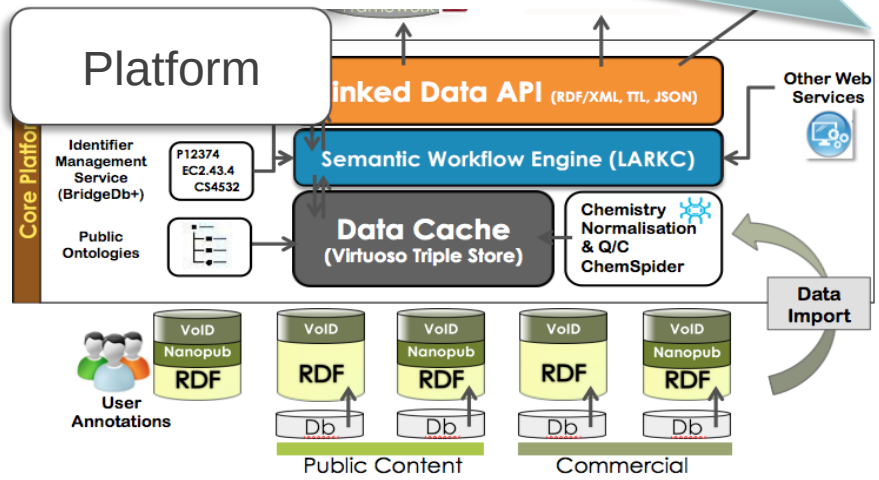
Kamal Azzaoui[1], Edgar Jacoby[14], Stefan Senger[2], Emiliano Cuadrado Rodríguez[3], Mabel Loza[3], Barbara Zdrazil[4], Marta Pinto[4], Antony J. Williams[5], Victor de la Torre[6], Jordi Mestres[7], Manuel Pastor[7], Olivier Taboureau[8], Matthias Rarey[9], Christine Chichester[10], Steve Pettifer[11],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

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

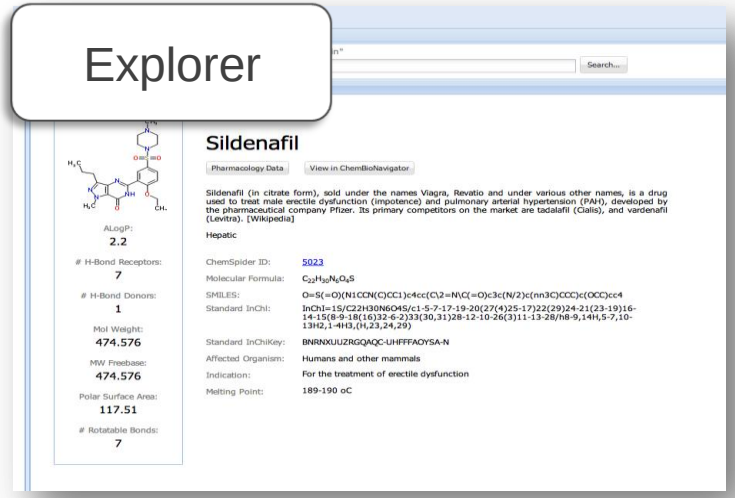
Open PHACTS
Open Pharmacological Space

“Provenance Everywhere”

Platform



Explorer



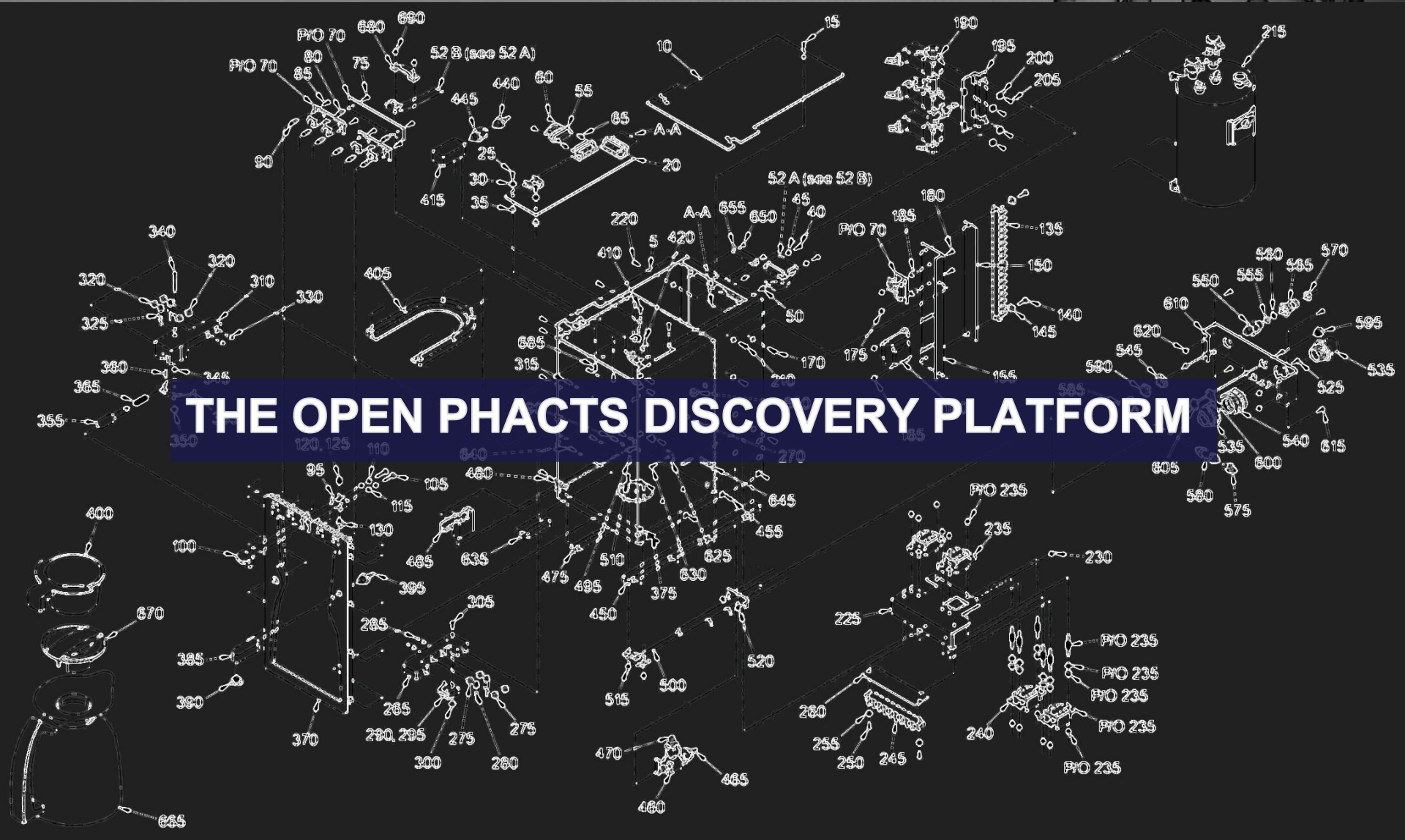
Apps

API

Standards

Open PHACTS

Open Pharmacological Space



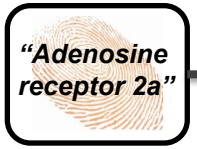
THE OPEN PHACTS DISCOVERY PLATFORM

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)

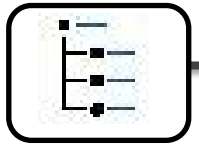
Semantic Workflow Engine

Data Cache
(Virtuoso Triple Store)

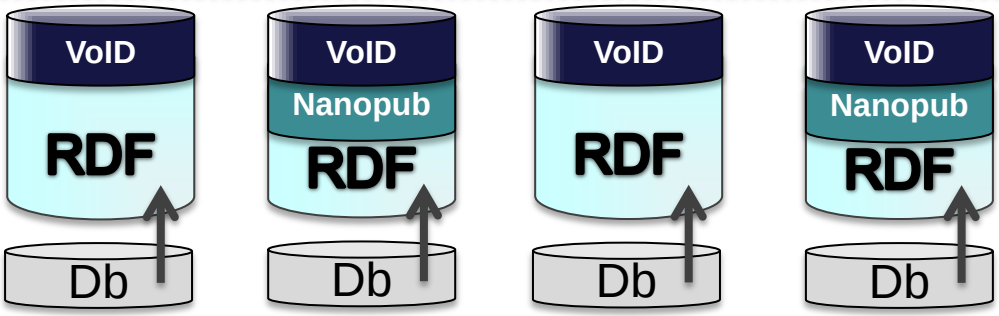
Domain Specific Services

Chemistry
Registration
Normalisation
& Q/C

Indexing



Public Ontologies



Open PHACTS
Open Pharmacological Space

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

Data Licensing Solution

Chose John Wilbanks as consultant

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

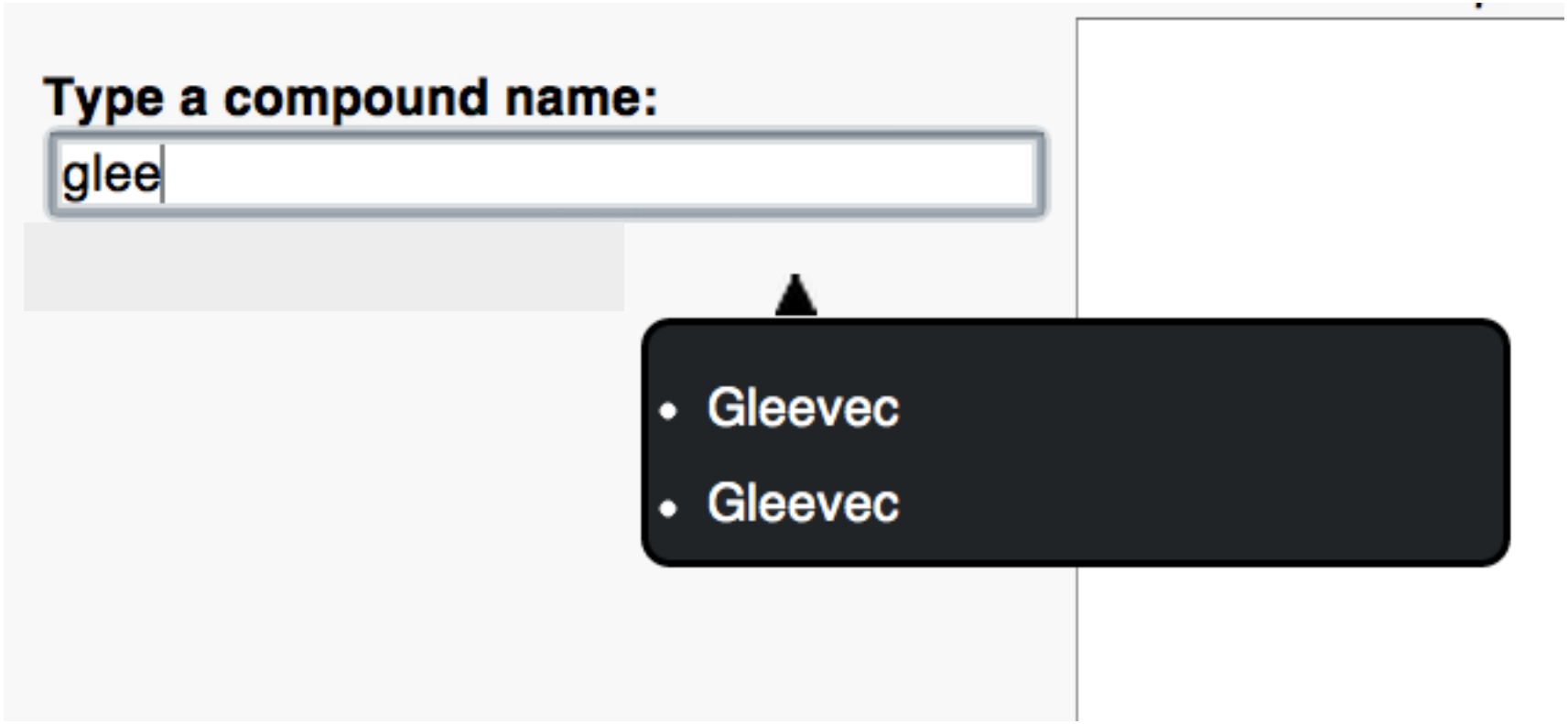
Deal with the problems by:

- 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						■	

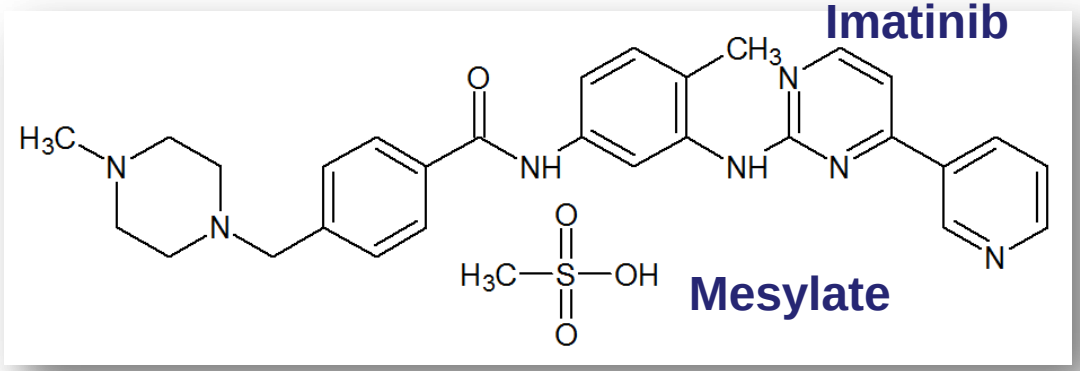
Open PHACTS
Open Pharmacological Space

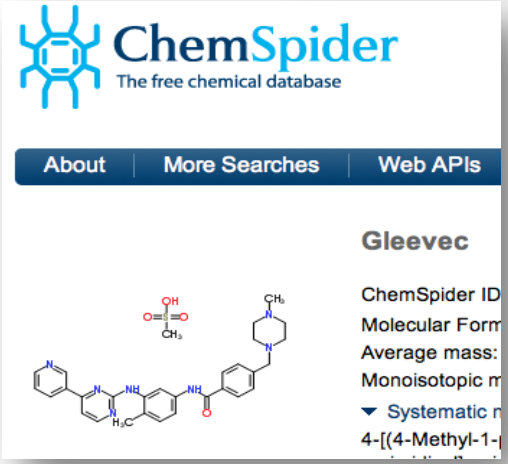
Its easy to integrate, difficult to integrate well:



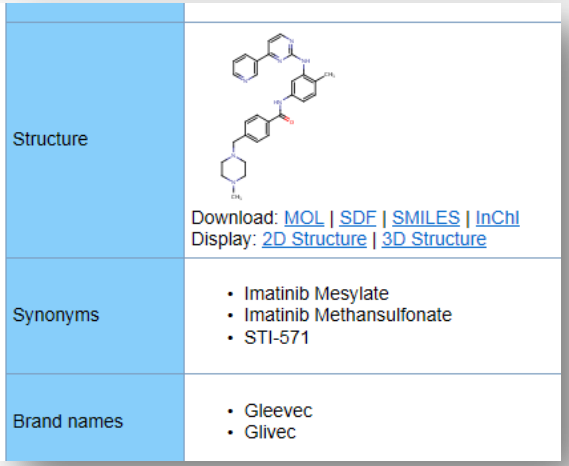
Open PHACTS
Open Pharmacological Space

What Is Gleevec?

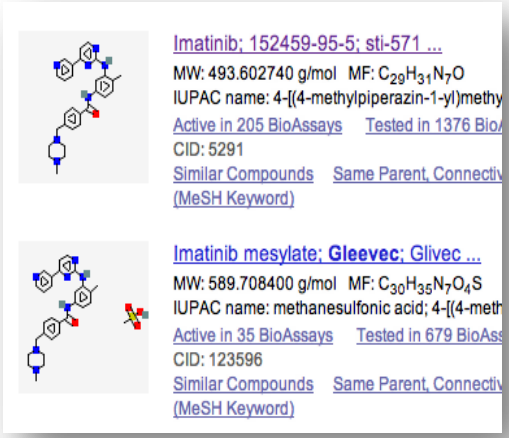




ChemSpider



Drugbank



PubChem

Open PHACTS
Open Pharmacological Space

Dynamic Equality

chemspider:gleevec

drugbank:gleevec

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

Play!

https://dev.openphacts.org/

OpenPHACTS API

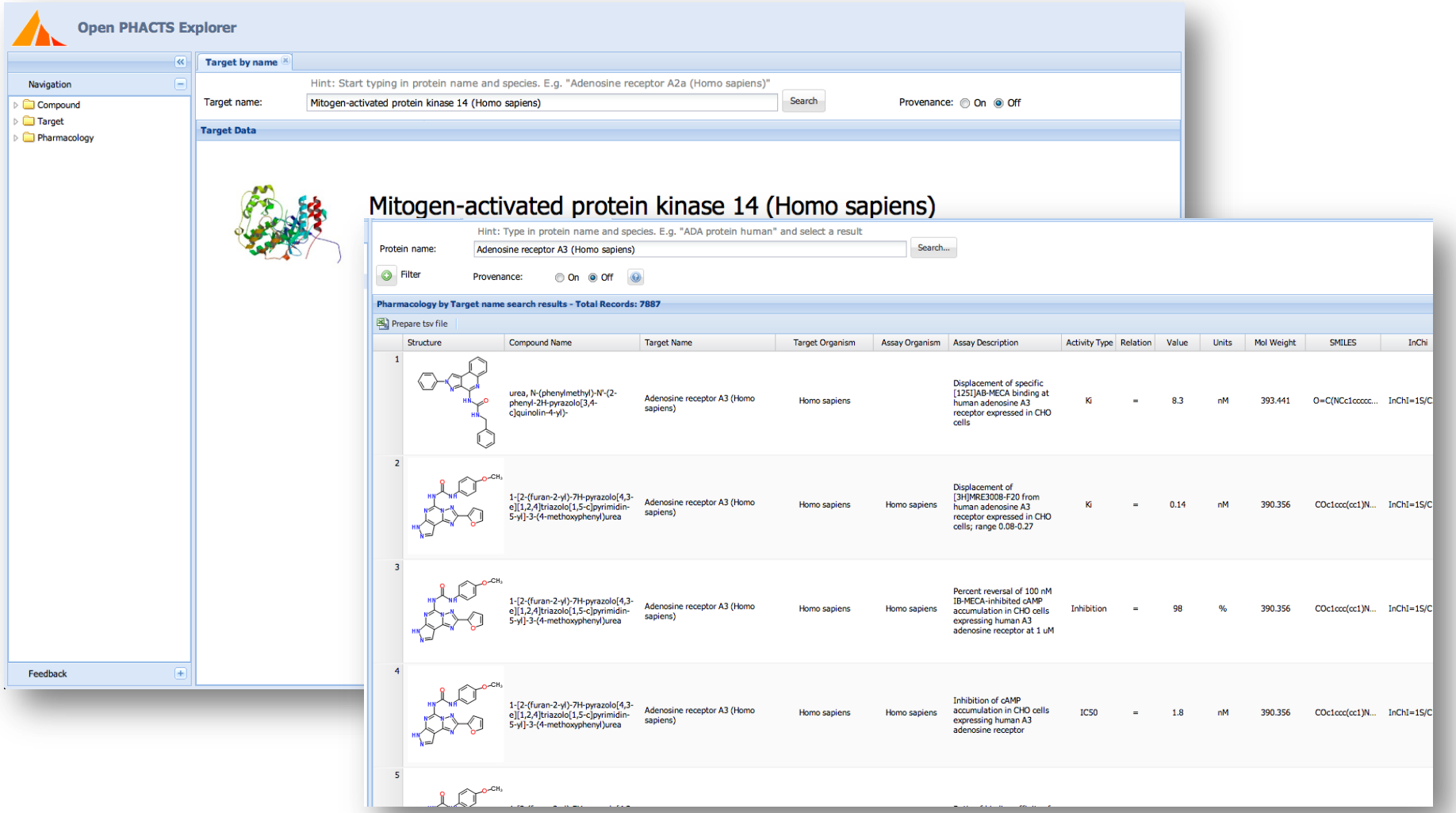
Chemical Structure Exact Search	/structure/exact	GET
InchiKey to URL	/structure	GET
Inchi to URL	/structure	GET
Chemical Structure Similarity Search	/structure/similarity	GET
SMILES to URL	/structure	GET
Chemical Structure Substructure Search	/structure/substructure	GET
Get concept description	/getConceptDescription	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.

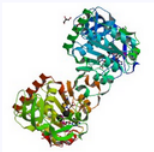
APPS

Open PHACTS

Open Pharmacological Space



http://explorer.openphacts.org



Quinone oxidoreductase (Homo sapiens) (4 of 4 results loaded)

Filter Results

Create TSV

Compound	Target	Assay	Activity						
Name	Organism	Organism	Description	Type	Relation	Value	Units	Mol Weight	SMILES	InChi
2,5-bis({6-[ethyl(2-methoxybenzyl)amino]hexyl}amino)cyclohexa-2,5-diene-1,4-dione	<		Binding affinity to NQQ1	Km	=	12700	nM	632.876	CCN(CCCCCCNC1=CC(=O)C(=CC1=O)NCCCCCCN(CC)CC2=CC=CC=C2OC)CC3=CC=CC=C3OC	InChI=1S/C38H56N4O4/c1-5-41(29-31-19-11-13-21-37(31)45-3)25-17-9-7-15-23-39-33-27-36(44)34(28-35(33)43)40-24-16-8-10-18-26-42(6-2)30-32-20-12-14-22-38(32)46-4/h11-14,19-22,27-28,39-40H,5-10,15-18,23-26,29-30H2,1-4H3
2'-Hydroxy-4-bromochalcone	<	Homo sapiens	Induction of quinone reductase activity in human MCF7 cells	Activity				303.151	C1=CC=C(C(=C1)C(=O)/C=C/C2=CC=C(C=C2)Br)O	InChI=1S/C15H11BrO2/c16-12-8-5-11(6-9-12)7-10-15(18)13-3-1-2-4-14(13)17/h1-10,17H/b10-7+
2'-hydroxychalcone	<	Homo sapiens	Induction of quinone reductase activity in human MCF7 cells	Activity				224.255	C1=CC=C(C=C1)/C=C/C(=O)C2=CC=CC=C2O	InChI=1S/C15H12O2/c16-14-9-5-4-8-13(14)15(17)11-10-12-6-2-1-3-7-12/h1-11,16H/b11-10+
(2E)-1-(2-hydroxyphenyl)-3-(pyridin-2-yl)prop-2-en-1-one	<	Homo sapiens	Induction of quinone reductase activity in human MCF7 cells	Activity				225.243	C1=CC=C(C(=C1)C(=O)/C=C/C2=CC=CC=N2)O	InChI=1S/C14H11NO2/c16-13-7-2-1-6-12(13)14(17)9-8-11-5-3-4-10-15-11/h1-

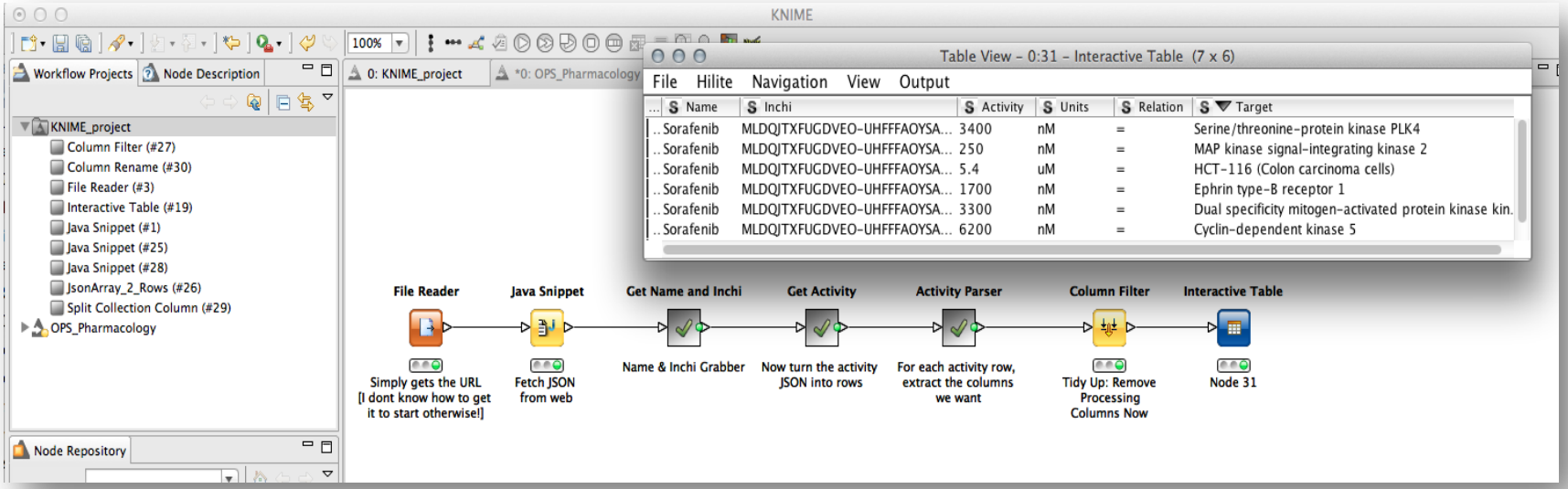
Open PHACTS
Open Pharmacological Space

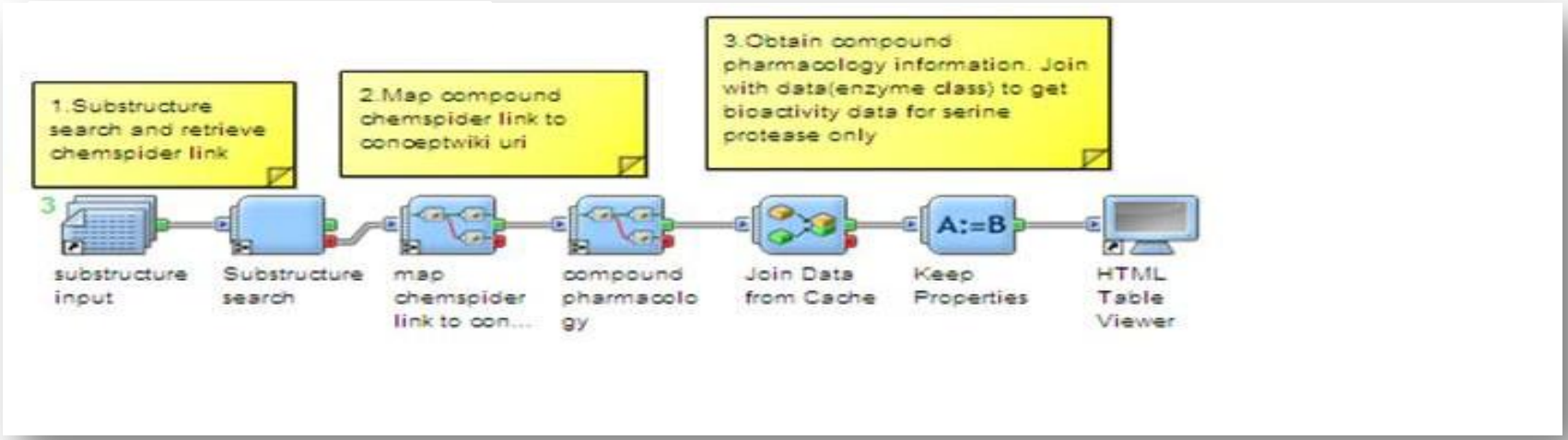
Example applications

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

Uptake at AstraZeneca: a Use Case

Applying BioAssay Ontology to facilitate HTS analysis

Linda Zander Balderud
Ola Engkvist

Chemistry Innovation Centre, Discovery Sciences
AstraZeneca

Assay Informatics project
Benefits in Adopting BioAssay Ontology (BAO)

- Common language for assay annotation
- Improved project success analyses based on assay technologies
- Better understand the impact of technology artifacts like frequent hitters
- Assay design and screening cascade support during assay development in early projects
- Improved capability to perform combined data mining of internal and public data

FLIPR Tetra High Throughput Cellular Screening System (from Molecular Devices)

The BioAssay Ontology (BAO)

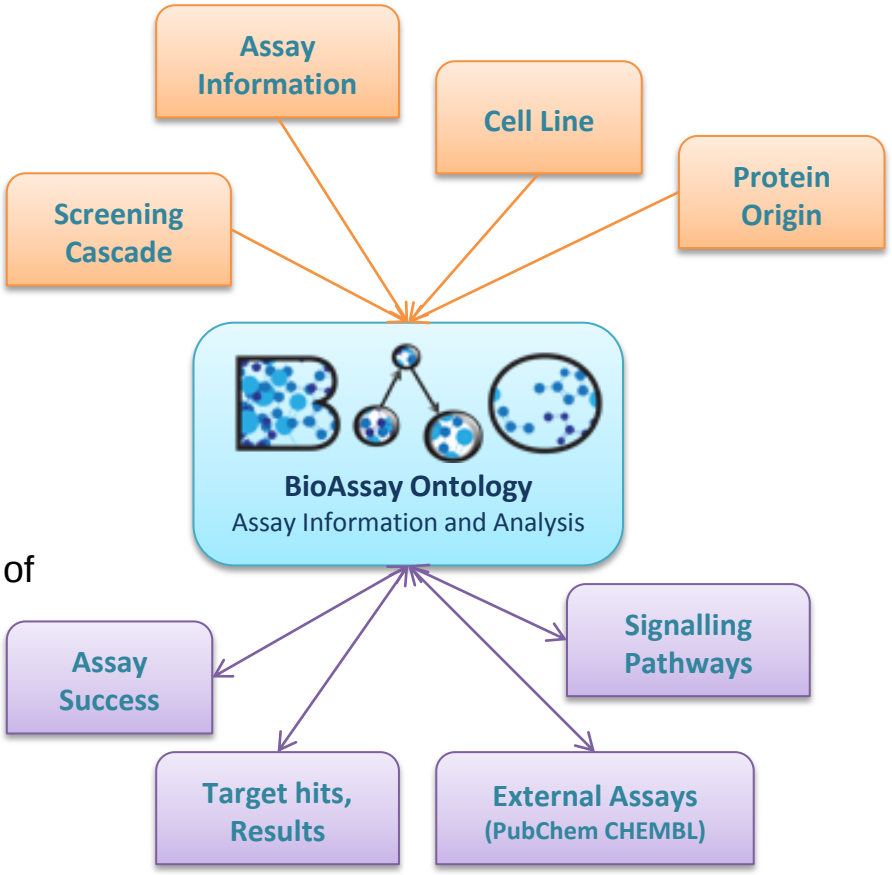
Computational Science, University of Miami, USA

Domain:

- Assay design
- Assay format
- Detection technology
- Meta target
- Endpoint
- Perturbagen

BioAssay Ontology imports:

- **NCBI taxonomy** - organism names and IDs
- **Uniprot** - protein target names and IDs
- **Unit Ontology** - concentration and time unit terms
- **Ontology of Biomedical Investigation** – descriptions of biological assays
- **Gene Ontology -** biological processes
- **Cell Line Ontology** - cell line names
- **CL** – cell types
- **UBERON** – anatomical entities
- **PATO** – cell phenotype
- **SAR connect** – target classifications



Migration to BAO
Annotation of HTS assays

Manual annotation of protocols

HTS assay: reporter gene assay

- Assay method: reporter gene method: beta lactamase induction
- Detection technology: FRET
- Bioassay: beta lactamase assay

- *Assay kit: LiveBLAzer FRET - B/G Loading Kit*
- *Wavelength: ex 405 em 460, 535*

- Biological process
- Disease

HTS assay: FLIPR

- Assay method: molecular redistribution determination assay
- Detection technology: fluorescence intensity
- Bioassay: calcium redistribution assay

- Assay kit: Fluo-8 No Wash Calcium Assay Kit
- Wavelength: ex 480 em 530

- Biological process
- Disease

Over 900 PubChem assays have been annotated by the BioAssay Ontology team

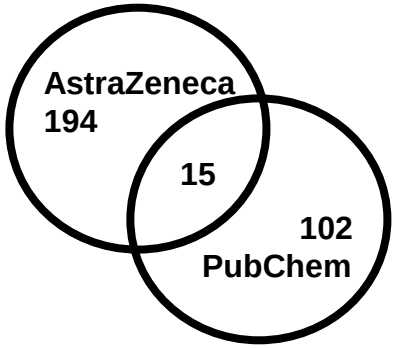
Assay Development Support

Comparison study between AstraZeneca and PubChem HTS assays

412 in-house HTS assays since 2005 have been annotated according to the BioAssay Ontology. The assay design and technology of the annotated assays were analyzed together with 239 primary assays from PubChem. The analyzed PubChem assays are biochemical assays, assays detected by luminescence and/or assays using GPCR targets.

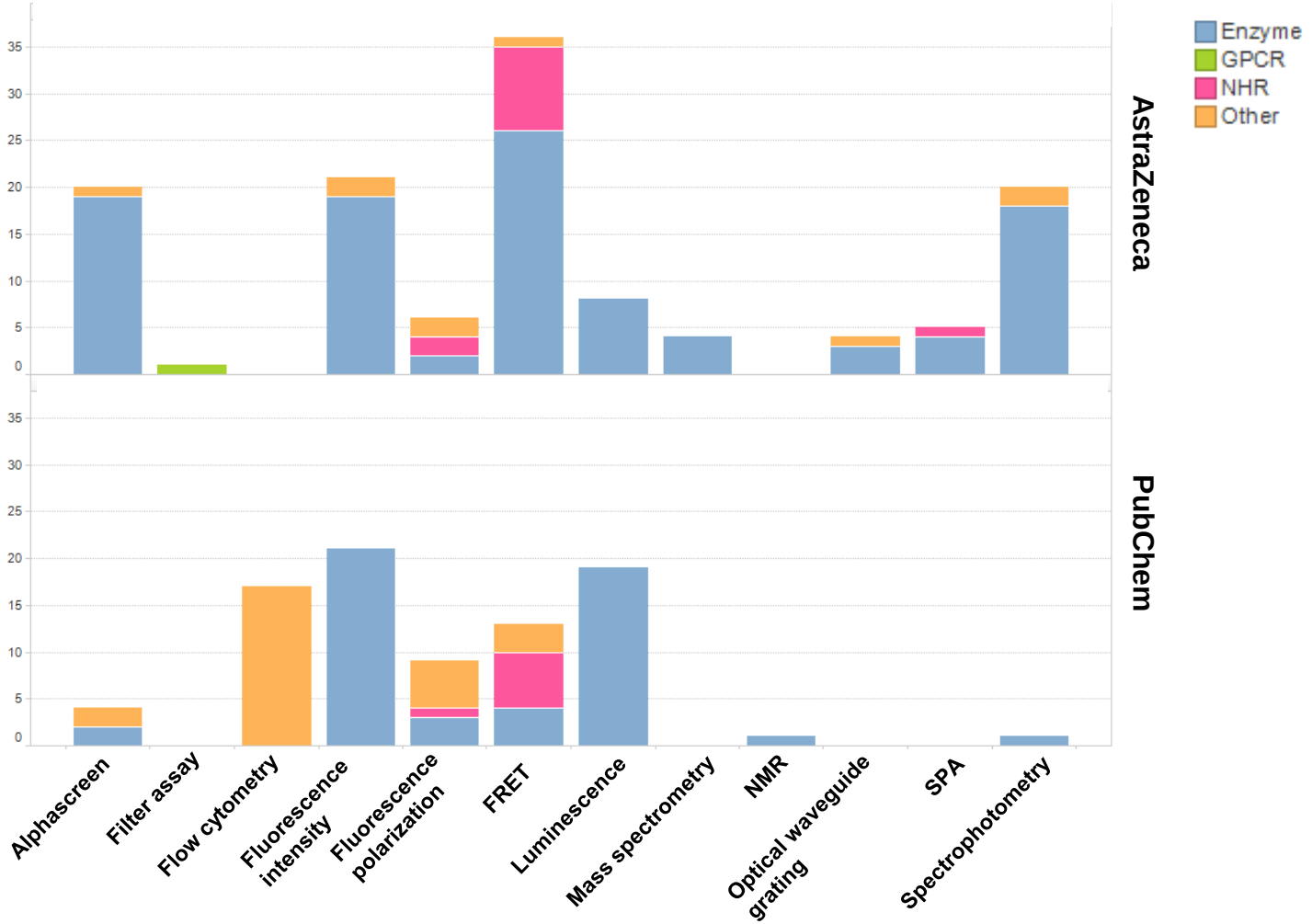
From the annotated assays, 515 assays were using human targets and combined 311 different human targets were represented in the study.

15 of the in-house targets were also screened in at least one PubChem assay. Eight of these were GPCR targets.



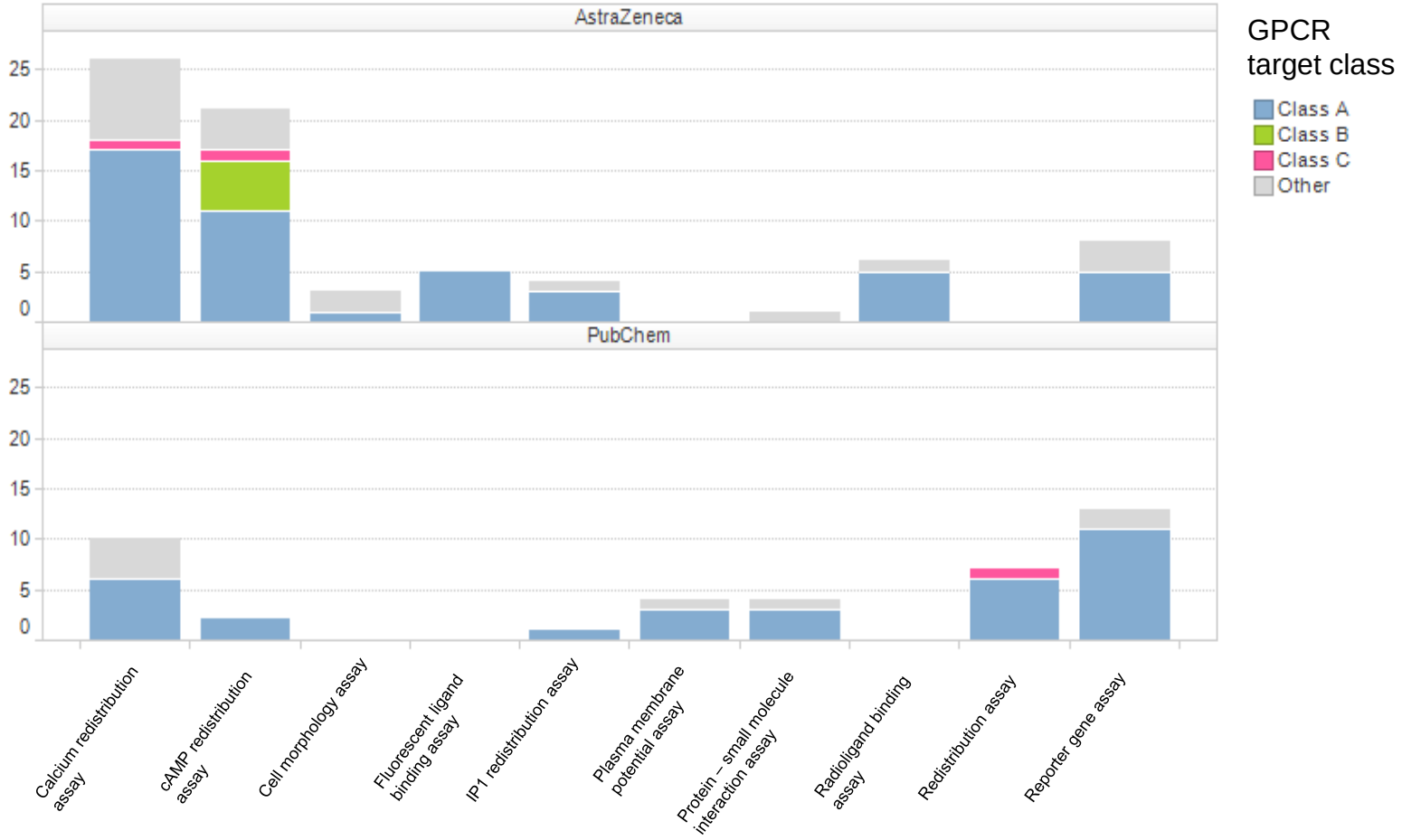
Assay Development Support

Detection Technology of AZ and PubChem Biochemical Assays



Assay Development Support

Assay design of in-house and PubChem GPCR HTS



One explanation for the low usage of cAMP redistribution method among the annotated PubChem assays could be that no class B GPCRs have been screened

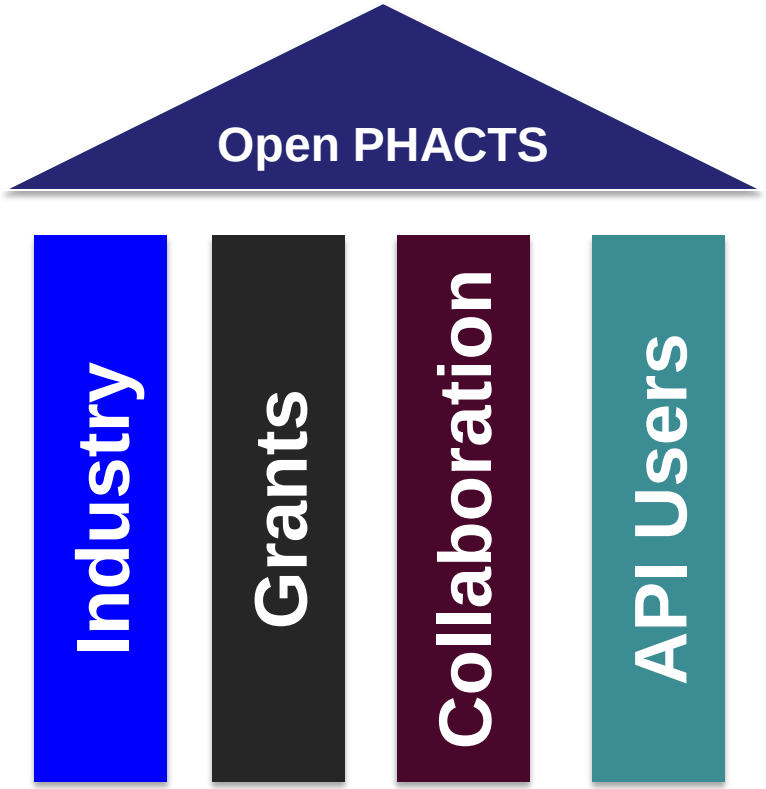
Open PHACTS
Open Pharmacological Space

Sustaining The Project

The Open PHACTS Foundation

Open PHACTS
Open Pharmacological Space

Kick-Starting Sustainability

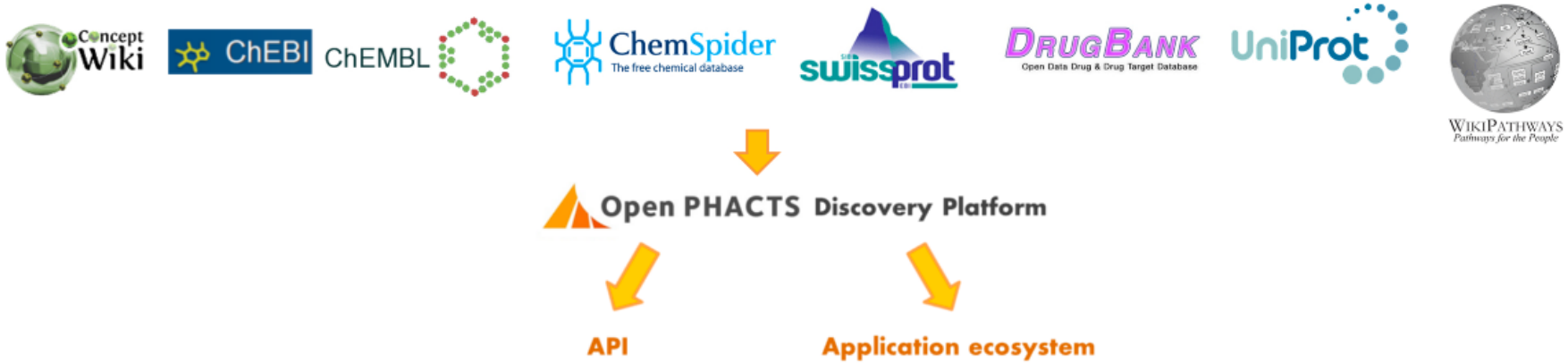


The Open PHACTS Foundation

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

pmu@openphacts.org

@Open_PHACTS

Open PHACTS