Merck Serono

MERCK

Adressing Research Questions by Integration of Open PHACTS with Pipeline Pilot

Accelrys Seminar, Mainz, 14th May 2013

Christian Herhaus, Global Computational Chemistry

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

- Runs 2011-2014
- *Create a semantic integration hub ("Open Pharmacological Space")*
- Deliver services to support on-going drug discovery programs in pharma and public domain
- Leading academics in semantics, pharmacology and informatics,
- Driven by solid industry business requirements

http://www.openphacts.org/

Open PHACTS Partners

Merck Serono

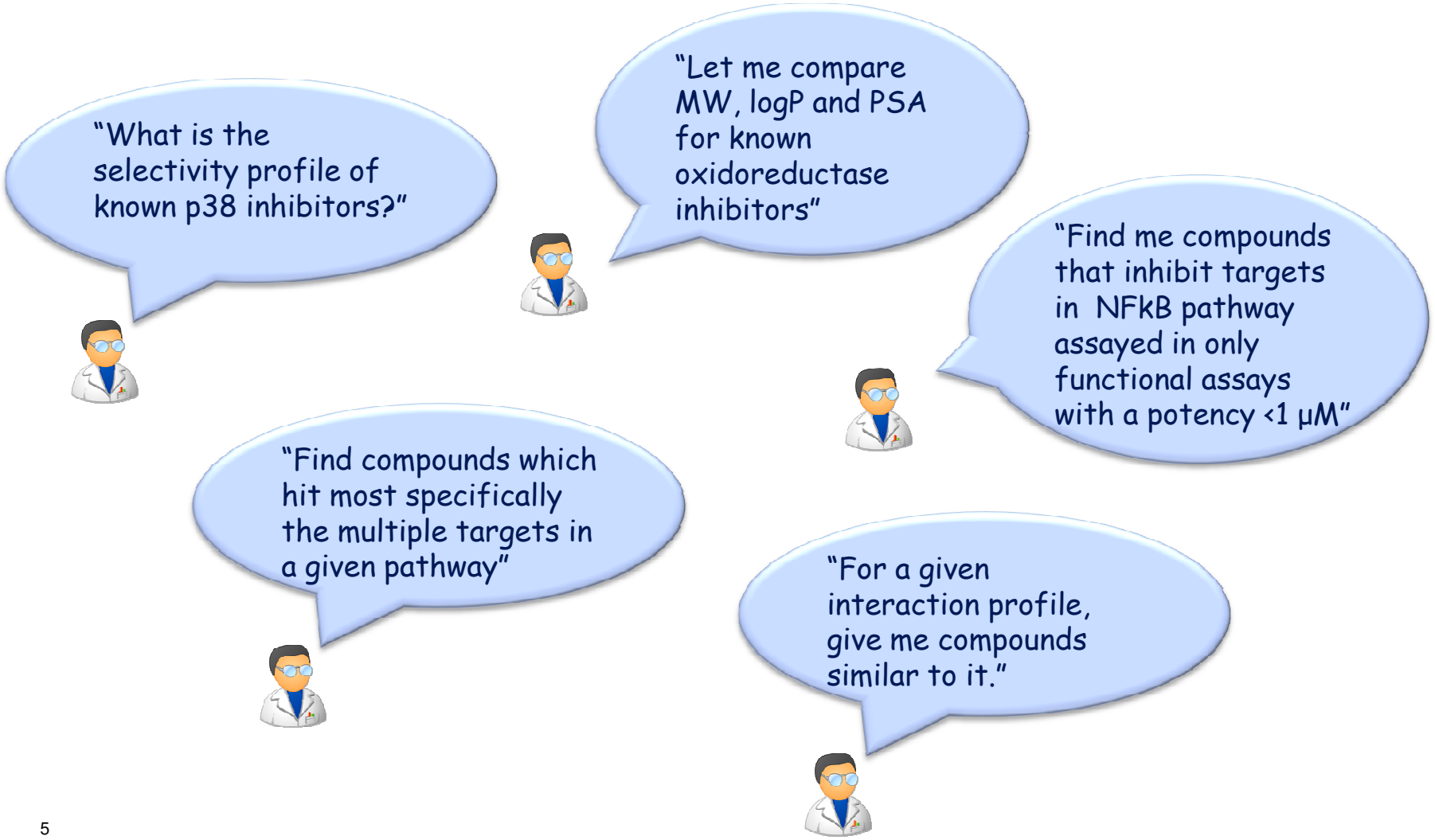
Open PHACTS

MERCK

Associate Partners

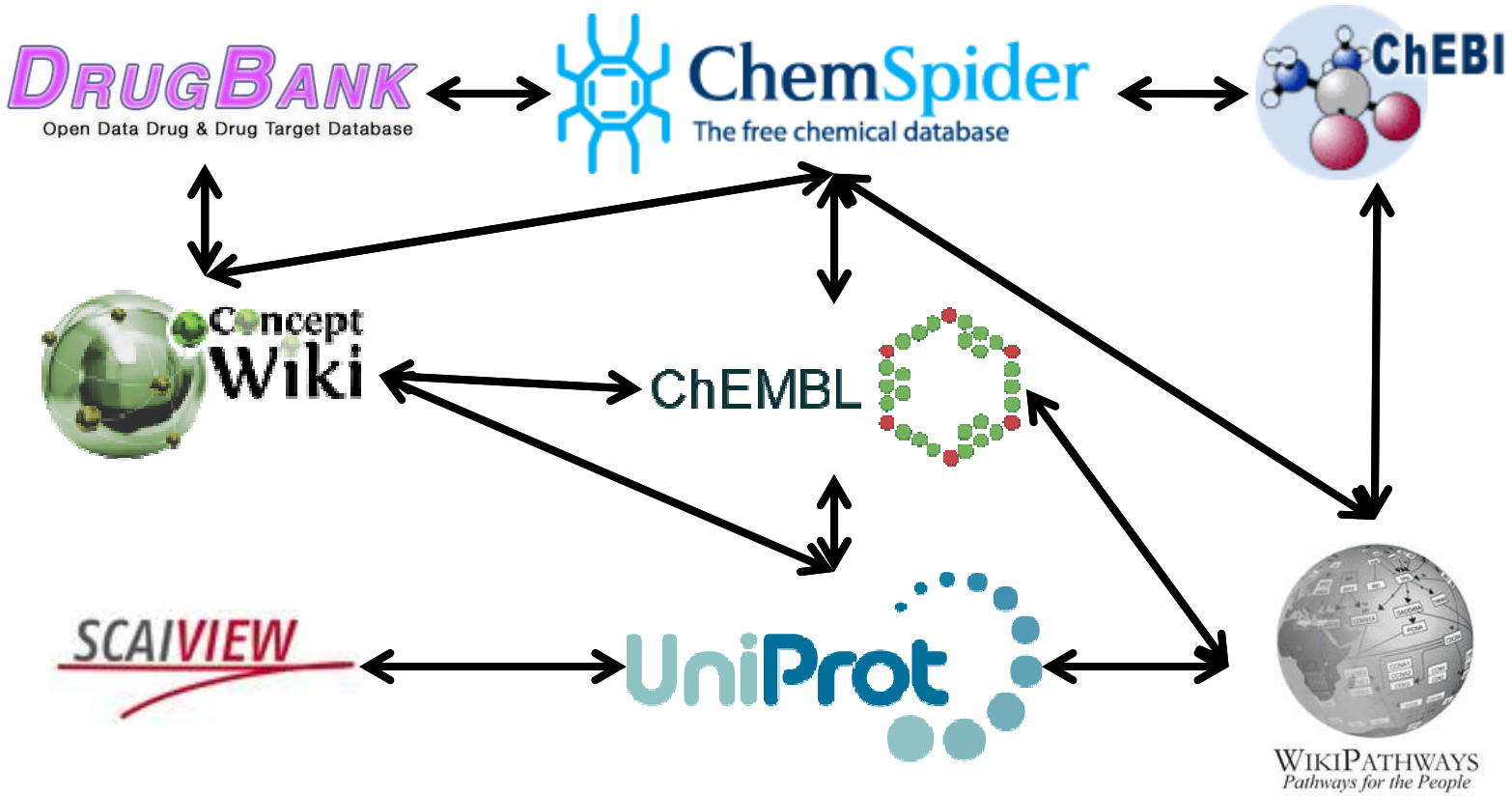
4

Focused on Answering Business Questions

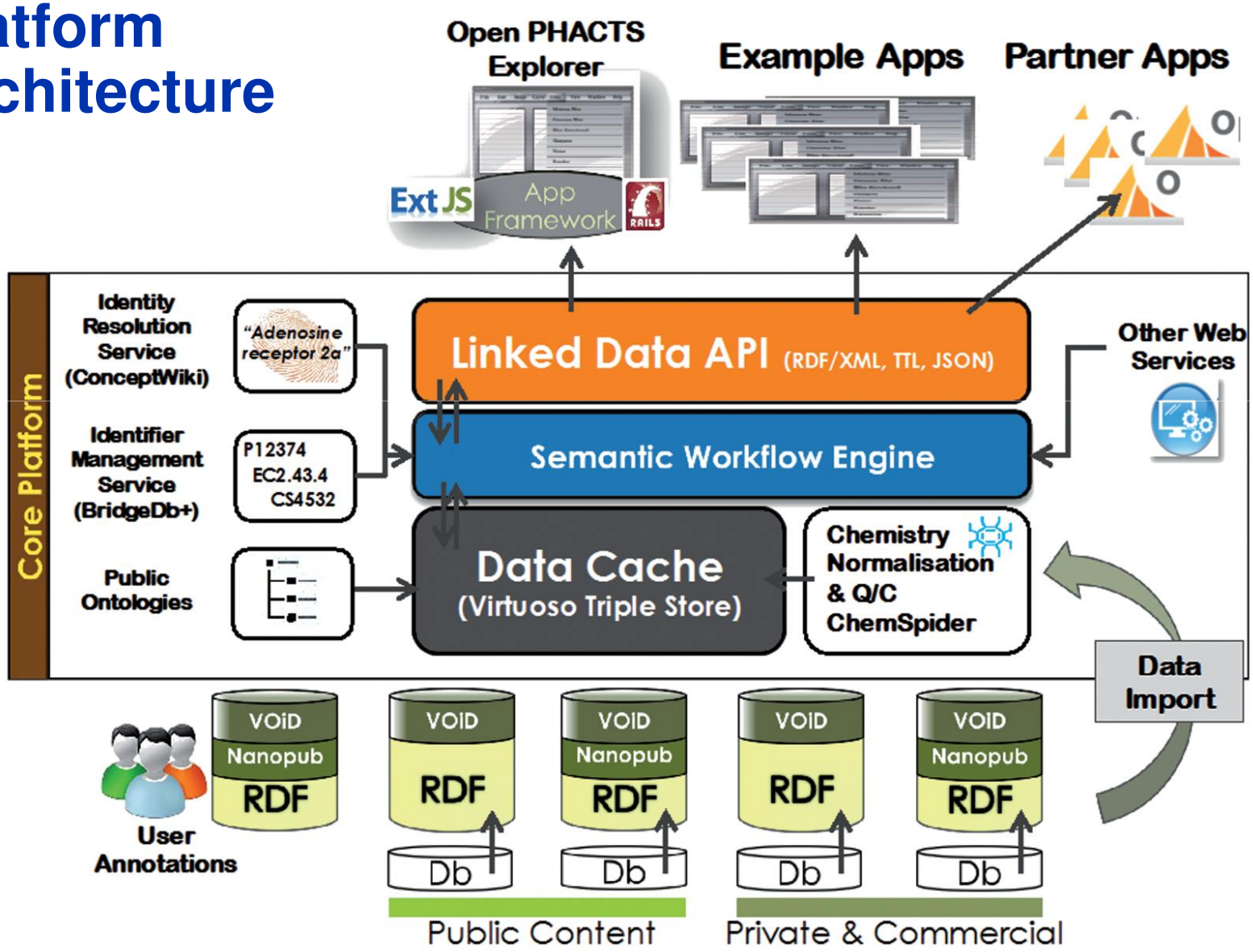


Datasets and Links

815 millions of Data triples, 18 millions of Mapping triples



Platform Architecture



Merck Serono

MERCK

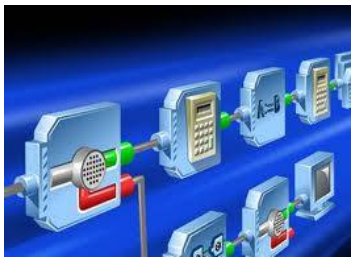
The Open PHACTS Linked Data API

Chemical Structure Exact Search	/structure/exact GET	Compound Pharmacology Complete (DEPRECATED)	/compound/pharmacology GET
InchiKey to URL	/structure GET	Compound Pharmacology Paginated	/compound/pharmacology/pages GET
Inchi to URL	/structure GET	Data Sources	/sources GET
Chemical Structure Similarity Search	/structure/similarity GET	Get Enzyme Classification Root Classes	/target/enzyme/root GET
SMILES to URL	/structure GET	Get Enzyme Classification Class	/target/enzyme/node GET
Chemical Structure Substructure Search	/structure/substructure GET	Get Enzyme Classification Class Members	/target/enzyme/members GET
Get concept description	/getConceptDescription GET	Enzyme Pharmacology Count	/target/enzyme/pharmacology/count GET
Map free text to a concept URL based on semantic tag	/search/byTag GET	Enzyme Pharmacology Complete (DEPRECATED)	/target/enzyme/pharmacology GET
Map free text to a concept URL	/search/freetext GET	Enzyme Pharmacology Paginated	/target/enzyme/pharmacology/pages GET
Map URLs	/mapURL GET	Pathway Information	/pathway GET
Get ChEBI Ontology Class Members	/compound/chebi/members GET	Activity types	/pharmacology/filters/activities GET
Get ChEBI Ontology Root Classes	/compound/chebi/root GET	Units for activity type	/pharmacology/filters/units/ GET
Get ChEBI Ontology Class	/compound/chebi/node GET	Target Information	/target GET
ChEBI Class Pharmacology Count	/compound/chebi/pharmacology/count GET	Target Pharmacology Count	/target/pharmacology/count GET
ChEBI Class Pharmacology Complete (DEPRECATED)	/compound/chebi/pharmacology GET	Target Pharmacology Complete (DEPRECATED)	/target/pharmacology GET
ChEBI Ontology Class Pharmacology Paginated	/compound/chebi/pharmacology/pages GET	Target Pharmacology Paginated	/target/pharmacology/pages GET
Compound Information	/compound GET		
Compound Pharmacology Count	/compound/pharmacology/count GET		

https://dev.openphacts.org/

Pipeline Pilot Hackathon

Cambridge, 20th Feb 2013



- Demonstrate that the integration of the Open PHACTS API and Pipeline Pilot (PP) is possible
- Create a library of PP components for the Open PHACTS API
- Enable:
  - Easy querying of the Open PHACTS API
  - “Drag and drop” creation of useful PP workflows
  - Customisation of PP components
  - PP component transparency

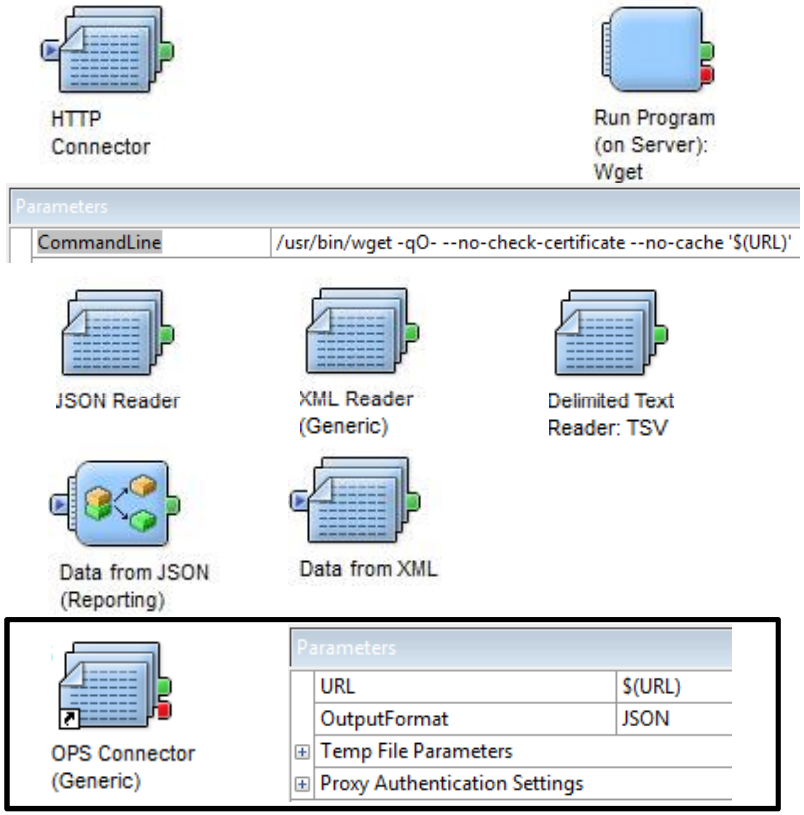
Connecting to Open PHACTS with PPilot

Chemical Structure Similarity Search `/structure/similarity` GET

Description

Returns a set of ChemSpider compound URLs, similar to the input molecule according to the specified algorithm and threshold. Driven by ChemSpider.

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
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.
_format		The desired result format.

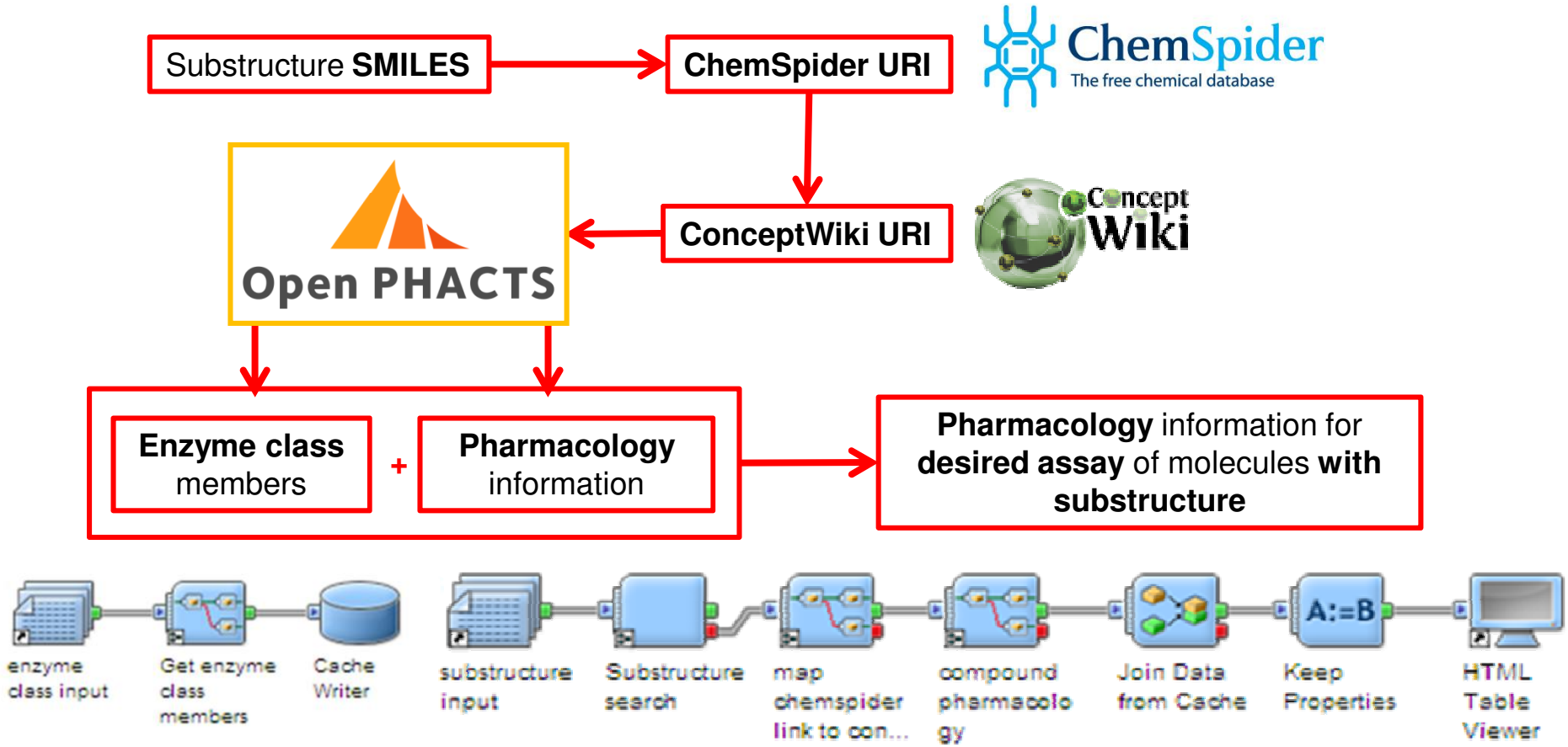


Example URL:

```
https://$(ops_server)/structure/similarity?app_id=$(app_id)&app_key=$(app_key)
&searchOptions.Molecule=c1ccccn1&searchOptions.SimilarityType=0
&searchOptions.Threshold=0.7&_format=json
```

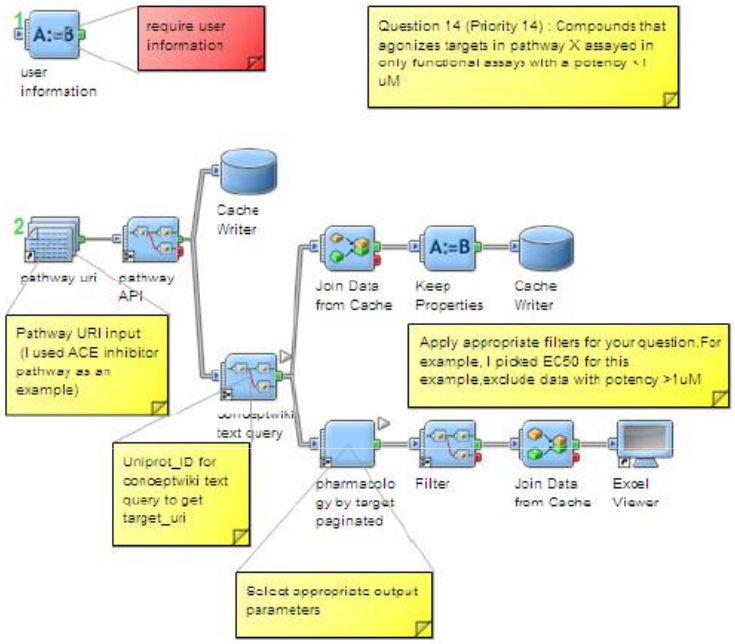
Answering Questions: Hackathon Example

*"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".*



Question 14 : Compounds that agonizes targets in pathway X assayed in only functional assays with a potency <1 uM

Pathway information	• Use pathway URI(s) as input e.g. http://rdf.wikipathways.org/WP1019_r48131 • Obtain all genes in that pathway as a result
Biomart web service	• Gene IDs provided in OPS pathway query are in form of Entrez. It needs to be converted into Uniprot IDs for further Conceptiwiki query
Conceptwiki mapping	• Text query of Uniprot IDs to get target URIs.
Target by pharmacology	• Retrieved pharmacology information for all targets in our desire pathway(s) using Conceptwiki target URIs • Option to select your output information
Filter	• Apply suitable filter(tailor made for your question.)



Filter for this specific question:

1. Select EC50 bioactivity data entry
2. Relation not equal to >
3. Convert activity unit from nM/-LogEC50 to uM
4. Assay with potency <1uM

Chris Chau (GSK)

Merck Serono

Open PHACTS

MERCK

Question 14 : Compounds that agonizes targets in pathway X assayed in only functional assays with a potency <1 uM

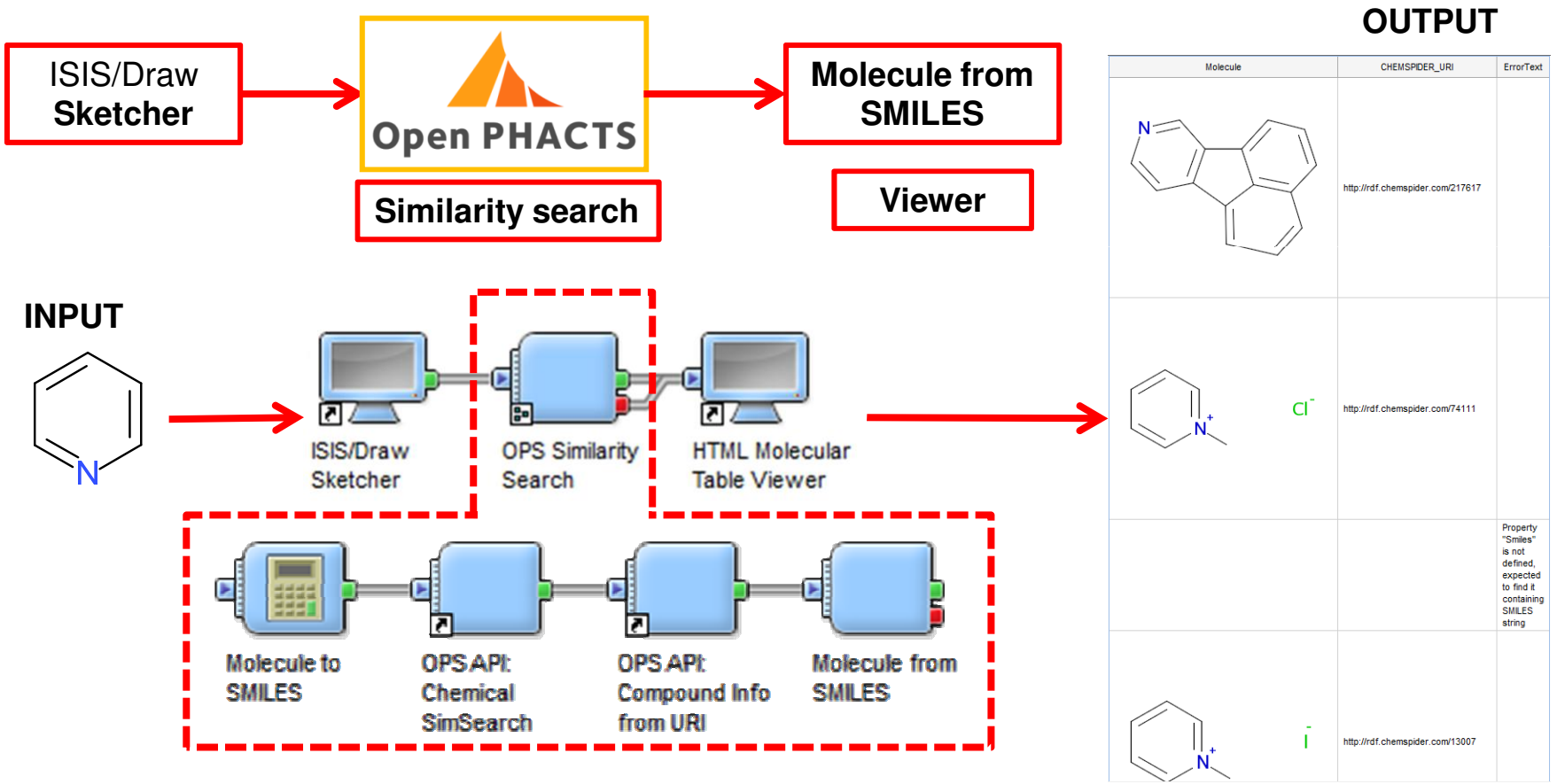
Example Results:

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	target_title	activity_ID	pmid	activity_type	relation	standardValue	standardUnits	chembl_assay_id	description	assay_organism	chembl_target_id	target_organism	chembl_compound_id	chemspider_uri	compound_uri	prefLabel	inchi	inchikey	full_mwt	smiles	ro5_violations	uM	pathway
2	Cathepsin	http://dati	2.1E+07	EC50	=	13.5	nM	http://dati	Inhibition	Homo sap	http://dati	Homo sap	http://dati	http://rdf.	http://ww	delanzomi	InChI=1S/(	SJFBTAPEP	413.275	B([C@H](C	1	1.35E-02	ACE Inh
3	Nitric oxi	http://dati	1.9E+07	EC50	=	420	nM	http://dati	Inhibition	Homo sap	http://dati	Homo sap	http://dati	http://rdf.	http://ww	N-(3-chlor	InChI=1S/(	VDKGHNO	428.819	Cn1cncc1C(=O)N(Cc2		0.42	ACE Inh
4	Mineraloc	http://dati	2E+07	EC50	=	3.8	nM	http://dati	Agonist activity at m	http://dati	Homo sap	http://dati	http://rdf.	http://ww	Aldosteroi	InChI=1S/(	PQSUYGKT	360.444	O=C(CO)[C@@H]4[C(		3.80E-03	ACE Inh
5	Mineraloc	http://dati	2.1E+07	-Log EC50	=	9.471		http://dati	Agonist ac	Homo sap	http://dati	Homo sap	http://dati	http://rdf.	http://ww	Aldosteroi	InChI=1S/(	PQSUYGKT	360.444	O=C(CO)[C@@H]4[C(		3.38E-04	ACE Inh
6	Mineraloc	http://data.kasabi.cc		EC50	=	2	nM	http://dati	Agonist ac	Homo sap	http://dati	Homo sap	http://dati	http://rdf.	http://ww	Prednisol	InChI=1S/(	OIGNJSKKI	360.444	O=C\1\C=C/[C@]4(/C		2.00E-03	ACE Inh
7	B2 bradyk	http://dati	1E+07	EC50	=	290	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	N-(4-chlor	InChI=1S/(	KTAARZWL	465.932	O=C(NCC)NCC(=O)N(c		0.29	ACE Inh
8	B2 bradyk	http://dati	1E+07	EC50	=	400	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	4-[({2-[(4-	InChI=1S/(	YAOHXRHI	523.968	Cc1ccc2cc	2	0.4	ACE Inh
9	B2 bradyk	http://dati	1E+07	EC50	=	56	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	(2E)-N-{2-[	InChI=1S/(	HRMANUL	592.995	FC(F)(F)c1	2	5.60E-02	ACE Inh
10	B2 bradyk	http://dati	1E+07	EC50	=	450	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	N-(2-cyan	InChI=1S/(	GSMWKYC	461.513	CCNC(=O)NCC(=O)N(C		0.45	ACE Inh
11	B2 bradyk	http://dati	1E+07	EC50	=	220	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	(2E)-N-{2-[	InChI=1S/(	QHVBUVLI	621.646	CCCOc1c((	1	0.22	ACE Inh
12	B2 bradyk	http://dati	1E+07	EC50	=	480	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	methyl 6-[	InChI=1S/(	DWSHMW	553.67	Cc1ccc2cc	1	0.48	ACE Inh
13	B2 bradyk	http://dati	1E+07	EC50	=	59	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	2-propena	InChI=1S/(	LQLXUORF	599.7	Cc1ccc2cc	2	5.90E-02	ACE Inh
14	B2 bradyk	http://dati	1E+07	EC50	=	54	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	2-propena	InChI=1S/(	AEFMWUU	570.487	COc1cccc(	2	5.40E-02	ACE Inh
15	B2 bradyk	http://dati	1E+07	EC50	=	350	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	N~2~-[(4-a	InChI=1S/(	ISROIASQI	509	O=C(NCCC	1	0.35	ACE Inh
16	B2 bradyk	http://dati	1E+07	EC50	=	660	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	(2E)-3-(4-r	InChI=1S/(	YMXBZVM	587.752	Cc1ccc2cc	1	0.66	ACE Inh
17	B2 bradyk	http://dati	1E+07	EC50	=	23	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	2-propena	InChI=1S/(	VUYAWMA	554.487	Cc1ccc(cc	2	2.30E-02	ACE Inh
18	B2 bradyk	http://dati	1E+07	EC50	=	260	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	2,4-penta	InChI=1S/(	RGYRXPGC	490.402	Cc1nc2c(cccc2s1)OC		0.26	ACE Inh
19	B2 bradyk	http://dati	1E+07	EC50	=	55	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	(2E)-N-{2-[	InChI=1S/(	TTZDPFYXI	598.012	FC(F)(F)c1	2	5.50E-02	ACE Inh
20	B2 bradyk	http://dati	1E+07	EC50	=	4.1	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	(2E)-N-{2-[	InChI=1S/(	LRFPQMIX	555.023	O=C(\C=C\	1	4.10E-03	ACE Inh
21	B2 bradyk	http://dati	1E+07	EC50	=	130	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	(2E)-N-{2-[	InChI=1S/(	MBVMOW	598.012	Cc1ccc2cc	1	0.13	ACE Inh
22	B2 bradyk	http://dati	1E+07	EC50	=	100	nM	http://dati	Inhibition	Cavia por	http://dati	Homo sap	http://dati	http://rdf.	http://ww	2-propena	InChI=1S/(	ACAUQEGI	608.459	FC(F)(F)c1	2	0.1	ACE Inh
23	B2 bradyk	http://dati	1.9E+07	-Log EC50	=	10.47		http://dati	Antagonis	Homo sap	http://dati	Homo sap	http://dati	http://rdf.	http://ww	L-Seryl-L-p	InChI=1S/(	WFRTYDJN	2243.59	C[C@H]([C	3	3.39E-05	ACE Inh
24	B2 bradyk	http://dati	1.9E+07	-Log EC50	=	10.31		http://dati	Antagonis	Homo sap	http://dati	Homo sap	http://dati	http://rdf.	http://ww	Bradykini	InChI=1S/(	QXZGBUJJ	1060.21	O=C(N[C@	3	4.90E-05	ACE Inh

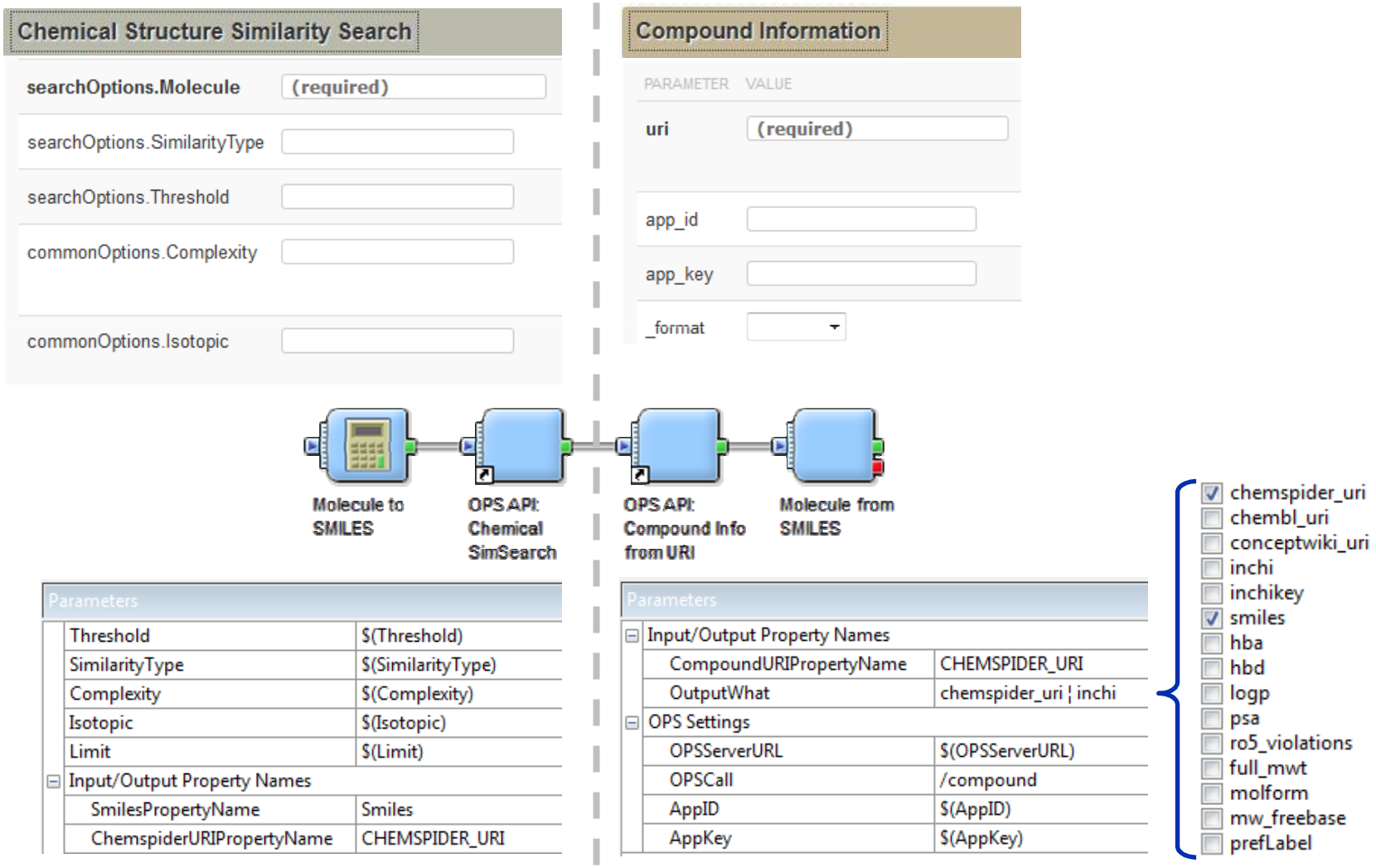
By Chris Chau (GSK)

From Tailored Prototyping to a Toolbox

Modularisation of partial solutions along Open PHACTS API calls



Mapping API Calls to Components

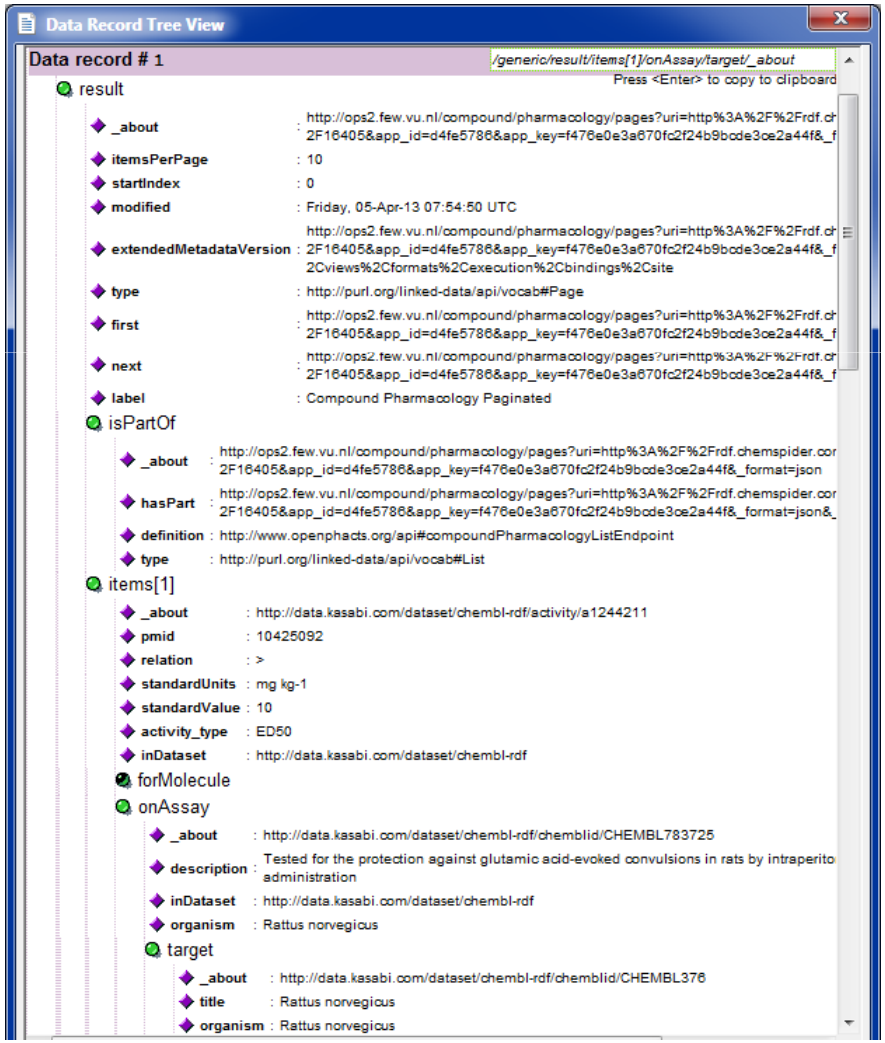


Merck Serono

MERCK

Open PHACTS Hierarchical Data Structure

A Challenge for Pipeline Pilot Data Tree Functionality

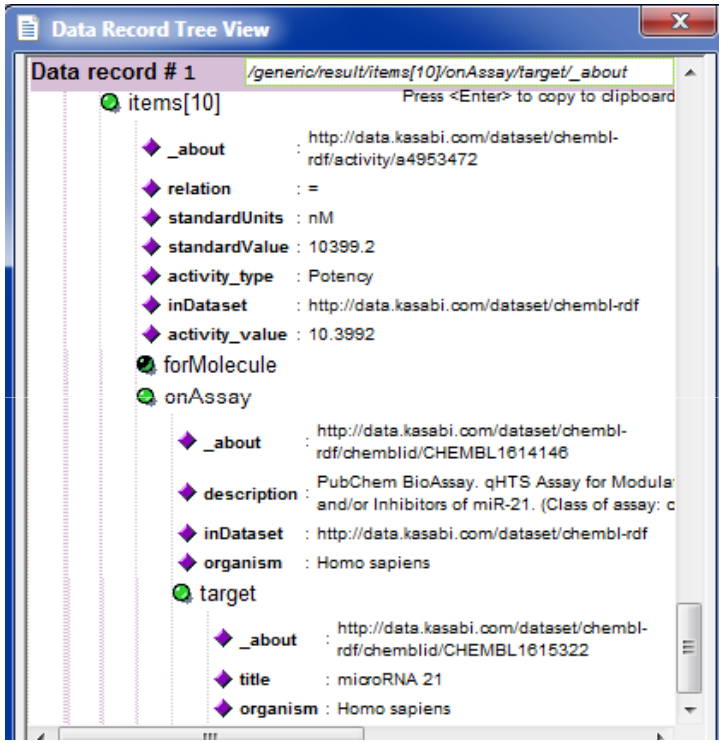


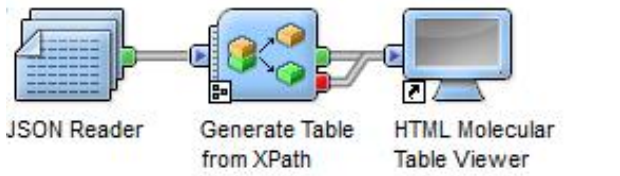
- Data Tree Manipulators
  - Copy Hierarchical Properties
  - Copy Nodes
  - Delete Hierarchical Properties
  - Delete Nodes
  - Detach Nodes
  - Flatten Hierarchy
  - Insert Nodes
  - Leaf Nodes to Properties
  - Move Hierarchical Properties
  - Move Nodes
  - Pad Properties
  - Properties to Leaf Nodes
  - Query Nodes
  - Reattach Nodes
  - Unflatten Hierarchy

Hierarchical Functions Index

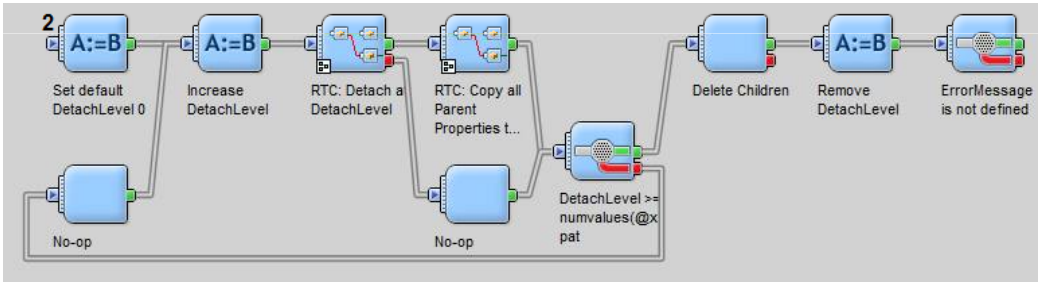
- AddChild
- DataRoot
- DataSetRoot
- GlobalRoot
- IsValidLocation
- NodeAppendChild
- NodeDeleteAllChildren
- NodeDeleteChild
- NodeGetFirstChild
- NodeHasChild
- NodeIsNull
- NodeIthChild
- NodeIthProperty
- NodeIthPropertyName
- NodeMetaDataProperty
- NodeNumChildren
- NodeNumProperties
- NodeParent
- NodePathLocation
- NodeProperty
- NodeQuery
- NodeRemoveMetaData
- NodeSerialize
- NodeSetType
- NodeType
- NodesEqual
- NumChildren

Unfolding the Open PHACTS Data Hierarchy





Parameters	
RootXPath	/generic/result
ExtractionXPath	/items/onAssay/target



_about	title	organism	description	inDataset	relation	standardUnits	standardValue	activity_type	activity_value
http://data.kasabi.com/dataset/chembl-rdf/chemblid/CHEMBL1615322 http://data.kasabi.com/dataset/chembl-rdf/chemblid/CHEMBL1614146 http://data.kasabi.com/dataset/chembl-rdf/activity/a4953472	microRNA 21	Homo sapiens Homo sapiens	PubChem BioAssay. qHTS Assay for Modulators of miRNAs and/or Inhibitors of miR-21. (Class of assay: confirmatory)	http://data.kasabi.com/dataset/chembl-rdf http://data.kasabi.com/dataset/chembl-rdf	=	nM	10399	Potency	10.399

Thank you!

Slides:

- Lee Harland (Connected Discovery Ltd.)
- Jack Gibb (Royal Society of Chemistry)
- Chris Chau (GlaxoSmithKline)
- … and many other authors from EFPIA and Open PHACTS

Administration:

- Stefan Senger (GlaxoSmithKline)
- Noj Malcolm (Accelrys)

Merck Serono:

- Thomas Grombacher
- Oliver Karch
- Friedrich Rippmann