

# Open Pharmacological Space

## The Open PHACTS Discovery Platform

Collaborations in Semantic Data Integration for Life Sciences Nick Lynch





## Open PHACTS Mission: Integrate Multiple Research Biomedical Data Resources Into A Single Open, Sustainable & Free Access Point





## **Pre-competitive Informatics:**

Pharma are all accessing, processing, storing & re-processing external research data



Lowering industry firewalls: pre-competitive informatics in drug discovery Nature Reviews Drug Discovery (2009) 8, 701-708 doi:10.1038/nrd2944

















### **Business Question Driven Approach**

Numb er	sum	Nr of 1	Question				
15	12	9	All oxidoreductase inhibitors active <100nM in both human and mouse				
18	14	8	Given comp off,target sat evidence (jo	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 tara Determine	Given a target find me all actives against that target. Find/predict polypharmacology of actives			
32	13	8	For a give	ELSEVIER	Volume 18 Jasues 17–18 September 2013 Pages 843–852		
37	13	8	The curre data in se		Volume 10, issues 17–10, September 2013, Pages 043–032		
38	13	8	Retrieve a	Review			
41	13	8	A project i compound modulate least at th literature.	oject i pound dulate t at th ature. Scientific competency questions as the basis for semantically enriched open pharmacological space development Kamal Azzaoui <sup>1</sup> , Edgar Jacoby <sup>14</sup> , Stefan Senger <sup>2</sup> , Emiliano Cuadrado Rodríguez <sup>3</sup> , Mabel Loza <sup>3</sup> , Barbara Zdrazil <sup>4</sup> , Marta Pinto <sup>4</sup> , Antony J. Williams <sup>5</sup> , Victor de la Torre <sup>6</sup> , Jordi Mestres <sup>7</sup> , Manuel Pastor <sup>7</sup> , Olivier Taboureau <sup>8</sup> , Matthias Rarey <sup>9</sup> , Christine Chichester <sup>10</sup> , Steve Pettifer <sup>11</sup> , Niklas Blomberg <sup>12, a</sup> , Lee Harland <sup>13</sup> , Bryn Williams-Jones <sup>13</sup> , Gerhard F. Ecker <sup>4</sup> , A. Science Pettifer <sup>11</sup> , Niklas Blomberg <sup>12, a</sup> , Lee me the compound(s) which hit most specifically the multiple targets in a given pathway ease)			
44	13	8	Give me a T				
46	13	8	Give me trie (disease)				





## The Open PHACTS Discovery Platform

•Cloud-Based "Production" Level System. Secure & Private

### •Guided By Business Questions

•Uses Semantic Web Technology <u>But</u> provides a simple REST-ful API for everyone else



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

### http://dx.doi.org/10.1016/j.drudis.2013.05.008



Web Semantics: Science, Services and Agents on the World Wide Web

Available online 8 April 2014



In Press, Accepted Manuscript --- Note to users

API-centric Linked data integration: The open PHACTS discovery platform case study

Paul Groth<sup>a</sup>, Antonis Loizou<sup>a</sup>, Alasdair J.G. Gray<sup>d</sup>, Carole Goble<sup>b</sup>, Lee Harland<sup>c</sup>, Steve Pettifer<sup>b</sup> http://dx.doi.org/10.1016/j.websem.2014.03.003











### Are These Two Molecules The Same(\*)









\*Really: Is it sensible to combine data associated with these two molecules?









Sources							
ChEMBL							
DrugBank	bille						
UniProt							
ENZYME							
ChEBI	Regular updates						
GO/GOA	Future new sources planned based on						
ChemSpider/ACD	•Options for hosting commercial data ( e g						
ConceptWiki	GVKBio)						
Disgenet	<ul> <li>Options for hosting private data</li> </ul>						
WikiPathways	Potential for Data update pipelines						
SureChEMBL	♦ VM platform						





### Explorer

### Explorer2

A Open PHA	CTSEppo		alwa Jaet				tige:3	
New LOON	Crows P	rist				here	b.	3
:Hi	Cafe or	190 (50 171	of 2071 results loaded) <u>rectant</u>	licth				
ja.	0	(NO;	Interfactorial cases, access in service and access and a supervised case. Collect a neuroscience and Collection and access and acce access and access					
30			ing .					
ten .	Option	Ogene	lauste	Į0	Nile	84.0	Ata	-
Admass All magna	ion spin	Cena posital	licity alto gan alwansi i saportin yanag basarinkansi anglis Dispologialasina albigat	K,		00000 1	1 1 22	
Alexand II	kta mepa		lindig after system annow 4 mages for a main measure, ang 4 (1)-é solara T amanda solatigné ): A solaran af 2 d'Assignet admonts	. 6		27220 el	( 1923	
Agreem At	Ma.	Ma .	meter (7/m2/com/app) addressing t answer superior scatter	· 6		44000 18	( 100	

Collector

### MOE



Col	lector						
285							
10 100 100	Safeta a		Data				
14	lage	Tempreci		Sec	2m	10100	1
m03.84	100	UPUTOr IBBO -1	1		Patrola		
Passiente patternites.	Paulo dap patterist.	Under Man-Aris					
Name and part (see \$12)		Interview O					
Associate producements	Passe dap permetal.	UPUTAL BOOK - TIME					
Annie imp pet twelfel	President stage gest there is all	Universities - K					
Annual segmentations:	Annual memory design.	United and they					
MILLION PUTCH	NEX X NUMBER OF COLOR	UPUTAL BOOK - TR					
			ale	es son nas fully			

### ChemBioNavigator Target Dossier



### Cytophacts

Acytos	cape 3 plugin for chemical educture search and pathway Inite, based on the OpenPhadis API
Ins	tallation
Step 1:	git clone the project according to standard github procedure.
Step 2	streve dependencies with an intrinsive. They will be placed in the "hypothe" subdirectory
Step 3:	From your cyloscope 3 installation, gather all api jam. The following command will do it in one step:
gisse	as also 1.2.4 SWHITS fiel transmitsysteming times "tab" (or ' table $\alpha$ ' () ' , ()
Step 4:	Orease a build properties file and place it in the same directory as build only. This build properties file
should	define the location of all the API jars, for example:
otes	and direction with the international and a second

### **KNIME**



Mol. Data Sheets	PipelinePi	lot
	Use Case Pipeline Pilot Protocol	Ph
Carrier 7 2:05 PM	Search for Target Sequence in OPS	
New New Reactions Templates WebService     Scratch Sheet	Same of a second	
Example Reactions	Search composeds surgisting ACL in OPS	
and Place	Same Area Sector Area Sector Area	
	Send conservation Constrained	
	Prop Math not	
	Janssen J	1 1

### scinav.it

hase 2

CREATE

Marchaelan	() Tapic List 4 Die munsenter teent Daten <sup>me</sup>
Controls the provide set of the s	<ul> <li>Antonio Receiption of the second secon</li></ul>
Constitution of Constitution o	Oriel Al series

### Pharmatrek



### Garfield



### Q99572 2 3

Sci

Change

Type -Helix A1

Job Control

### **SciBite**

Bite. Pharma: Right Now. Now Mathematic Pro-	lach Corpany Q,South MySolikian
atest News / CAFFEINE NewsStream Antonion	
ration for <b>GATERE</b> , as $(T \cup V ter 2)^{2}$ Note Exact Equal (spatiant 4) and one the "Bady Styleboch too on the sites that is a new minimized tables in the data set parts. The set of target these attains, and target the synch spatial (spatial data set), and the set of the set of target target the spatial set of the set of the spatial set of the set of the set of the set of the set of target targe	Topic Activity [ant 14 days]
d Plante See Contemported Team Process Propriet Performance in Vale Volephal Projects PERE-Contemp Drawy Date Enteries Propriet Performance in Vale Volephal Projects and Vol. III. 2004 Team Proceedings of Palate	Sin Report Speen Sp: Enal or HSS     Sinci No. 2001     Bend This Tapic
alan Eugeni (Mengilin Andreig Den Kinde) Simhar Orang (Ja-Pala Ellen News Bin Anart (Fangel King O New Jac Types) (New Yok Winkel(Den Stelland)	
Herr Erber Taute Bude May die Die The Fritz Inn Janochily (trakena) <b>Han D</b>	Adminy     Removement     Adminy     Removement     Administration     Key Researces:
Carls SELINO 212 - Pele Ban News In April (Hayne) Tatte D . Joint D . Thermit: D Inter La Carl (Har D . Mathematics Works)	Pharmacelogy (2 Open PAACH)     A     Ovend (Resetting)     Ovendyder (Resetting)     Repetition     Repetition

12	- Britchyot	let -
C	-1997	Jaci Nal
	File	let .
	0.405	Un
-	( <b>R</b>	
5	CR	
	(m.	
	Trend	
	1.	
40		
	- Inputlip/	hemanitata kon kema
	law -	068338 068366 0666831 Be
	1010	Delan den set berteren berteren
	aste s	anne koro kansk koro bis bis den se

### Taverna



http://www.openphactsfoundation.org/apps.html



### Utopia





## Sustaining Impact

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

...and more resourcefor future development







HOME

### 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.



## Acknowledgements





GlaxoSmithKline – 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 Novartis Merck Serono H. Lundbeck A/S Eli Lilly Netherlands Bioinformatics Centre Swiss Institute of Bioinformatics ConnectedDiscovery **EMBL-European Bioinformatics Institute** Janssen Esteve Almirall OpenLink Scibite The Open PHACTS Foundation Spanish National Cancer Research Centre University of Manchester Maastricht University Aqnowledge University of Santiago de Compostela Rheinische Friedrich-Wilhelms-Universität Bonn AstraZeneca Pfizer

## **Collaborations around Key Questions**





## Platform, Data & Experience Tacking Big Data challenges

























nick.lynch@curlewresearch.com bryn@connecteddiscovery.com