



Faculty of Health, Medicine and Life Sciences

Exposing WikiPathways as Linked Open Data



Please scan me

or visit grate.me/23819

Andra Waagmeester^A, Helena Deus^B, Chris Evelo^A

- Department of Bioinformatics BiGCaT, Maastricht University, the Netherlands А
- Digital Enterprise Research Institute, National University of Ireland, Galway В



Introduction

Biology has become a data intensive science. Discovery of new biological facts increasingly relies on the ability to find and match appropriate biological data

Exposing WikiPathways as Linked Open Data.

Exposing data to the Linked Open Data Cloud is a 4-step process: 1. Understand your data

elements.



Figure 1. Modern Biologists are required to process and integrate a wide variety of knowledge formats.

Linked Open Data Cloud

Semantic web and Linked Data technologies eliminate the barriers between database silos by relying on a set of standards and best practices for representing and describing data. The architecture of the semantic web relies on the architecture of the web itself for integrating and mapping universal resource identifiers (URI), coupled with basic inference mechanisms to enable matching concepts and properties across data sources. Semantic Web and Linked Data technologies are increasingly being successfully applied as integration engines for linking biological elements.

- 2. Select ontologies and controlled vocabularies
- 3. Map the data and the ontologies/vocabularies (e.g. Jena: http://jena.sourceforge.net/ and (tripliser: http://daverog.github.com/tripliser/).
- 4. Host the data on a triplestore (http://www.w3.org/wiki/LargeTripleStores)



Figure 4: This illustrates a selected mapping of WikiPathways classes to three ontologies/ Controlled vocabularies





WP: <http://www.wikipathways.org/index.php/Pathway:> DCTerms: <http://purl.org/dc/terms/> pubmed: <http://www.ncbi.nlm.nih.gov/pubmed/>

Figure 2: (A) "Linking Open Data cloud diagram (LOD), by Richard Cyganiak and Anja Jentzsch. http://lod-cloud.net/", (B) Linking to the LOD requires exposing your data in so called triples.

WikiPathways (http://www.wikipathways.org)

WikiPathways is an online repository of Pathways, where the content is curated by the online community. WikiPathways enables automatic querying with 24 predefined functions (ie. Listorganisms, listpathways, findinteractions, etc) (see: http://wikipathways.org/index.php/Help:WikiPathways_Webservice/API)



Tangible Results

With WikiPathways contact available as Linked Open Data we can:

- 1. Do SPARQL queries on WikiPathways content at http://semantics.bigcat.unimaas.nl:8000/sparql/
- 2.Integrate our content with other resources automatically (see Figure 5)
- 3. provide much more complex API calls through so calles SPARQL queries (see: semantics.bigcat.unimaas.nl/sparqlBookmarker.html)

→ C ⑤ ops.few.vu.nl:30	00			公 ()		
Open PHACTS G	IUI			Log in Create ac		
.60	Pathway from compound					
Navigation -	Compound name: L-Glutar	mate		Search		
OPS	Pathways including compou	nd - Records found: 25				
Compound Target	View with an 2 Mild Dather to work (Courts Tourist, 25					
Pharmacology	view pathway in whiteathways applict (Google Chrome only)					
a 🔁 Pathways	Download to Excel					
Pathway from compound Pathway from compound	Compound name	Protein or compound name Pathway description	Pathway id 🔺			
E Pathway from protein	1 L-Glutamate	Protein-tyrosine kinases The system involved i	n the reduction o WP100			
	2 L-Glutamate	Alcohol dehydrogen Pathway from WikiPathway	5	×		
	3 L-Glutamate	Aldehyde reductase				
	4 L-Glutamate		OPLAH			
	5 L-Glutamate	2,3-butanediol dehy	L-Clutamate	- Glutamate metabolism		
	6 L-Glutamate	NAD-linked glycerol	Cysteine metabolism			
	7 L-Glutamate	L-alpha-glycerol ph				
	8 L-Glutamate	L-alpha-glyceropho:	sm L-Cysteine			
	g L-Glutamate	L-giyceroi-3-phosph				
	10 L-Glutamate	NAD-dependent giv				
	11 L-Glutamate	NAD-linked giycerol				
	12 L-Glutamate	Xylitol denydrogena	Gyere			
	13 L-Glutamate	L-xyluiose reductasi	i-alycine			
	14 L-Giutamate	Carabinitol 4-denye				
	15 L-Giutamate		Cycle acid metabolism			
	16 L-Giutamate		rtat			
	1/ L-Giutamate	Aldenyde reductase	inducer) GSS g-L-Cutarry-L-systeme	1.8.7.13		
	18 L-Giutamate		hian			
	19 L-Giutamate	Aidose reductase	petion			
	20 L-Giutamate	mistration denyarog		Close		

Figure 5: Here WikiPathways LOD is integrated with chemical dataresources which used the same vocabularies and ontologies in a prototype of chemical compound browser. The pathways where a selected set of compounds are active were selected using LOD.

Acknowledgements

	Search	Browse		
= Home		Browse Pathways		
= Help		Distance patients from spanners [microson] in samples (Metabolic, Process]]	Ϋ́Υ	
athway		Austalization Jointines A Faily And Design Character BDCs7 Advances and sequences memory RDCs3 And And Particles SECC37	lang, res	
Create	Search	Knapers Annue Apriliaise Consensus Management Management (200) Consensus Responses Consensus Responses	LEGIND	
Browse	You can search by:	Powership Systems		
Wich Liet	Dillion and the second	Fely Not Rev Orderer 3 800047	Promis-postale interaction	
= Download	Pathway name (Apoptosis)	Fely har tere familier thes total Feloren Moscolen Beloren watchen	(0001) PRUPA PRUPA PRUPA PRUPA	
 Download 	 Gene or protein name (p53) 		Distriction Distriction Anti-collegion Anti-collegion Anti-collegion	
VVeb service API	Any page content (cancer)	Browse by species and estanony	(man) Bassipation Property Line	
overview		blowse by species and category	INNTS	
Recent Changes	Contributing New Pathwave		STATU STATU STATU	
Most Viewed	Contributing New Fatiways		MARKE Demithylition	
Most Edited	Create	Suggest		
New Pathways		ouggost	Transferrer	
Statistics		Pathway wahilat	INTERNAL STATUS	
	to a general fill	Name and Party and Sprant and Sprant	E Entracellation	
community		Ethiopen exception from the section of the section	03 Golgi superstant 11. Mittacheditice	
About us	Bream + Add Star day information	Party Learning trans. March 19 March 19 August		
Contact us	Transis Billionation Information - Academ	Management in Management in a second se	IL-9 signaling pathway	
How to cite		2007 Annual Statement Printers		
10 PT	Church dide	Territoria de la constante a constante de la const	Latest discussions	

We acknowledge financial support provided by the IMI-JU, grant agreement No 115191.





Correspondence to: Andra Waagmeester	Dept of Bioinformatics	Maastricht University	
Andra.Waagmeester@maastrichtuniversity.	nl		
www.bigcat.unimaas.nl	T +3143 388 2913	P.O. Box 616	
	F +3143 388 4149	6200 MD Maastricht, The Netherlands	