

Visualizing metabolomics data in directed biological networks

Denise Slenter, Chris Evelo, Egon Willighagen

Twitter: @SMaLLCaT4Sci and @BiGCaT_UM
Blog: <http://smallcats4science.blogspot.nl>
ORCID: 0000-0001-8449-1318

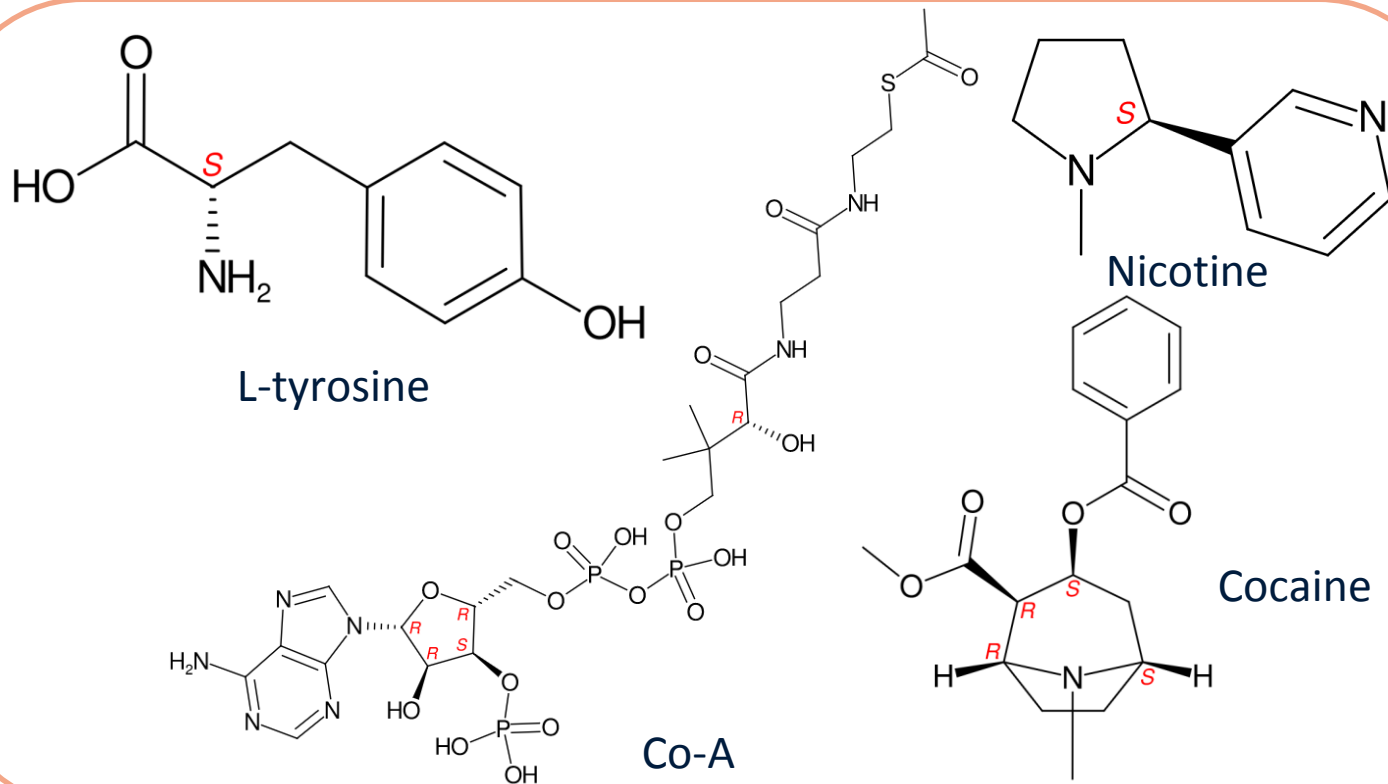
2019-06-25 Metabolomics 2019, The Hague



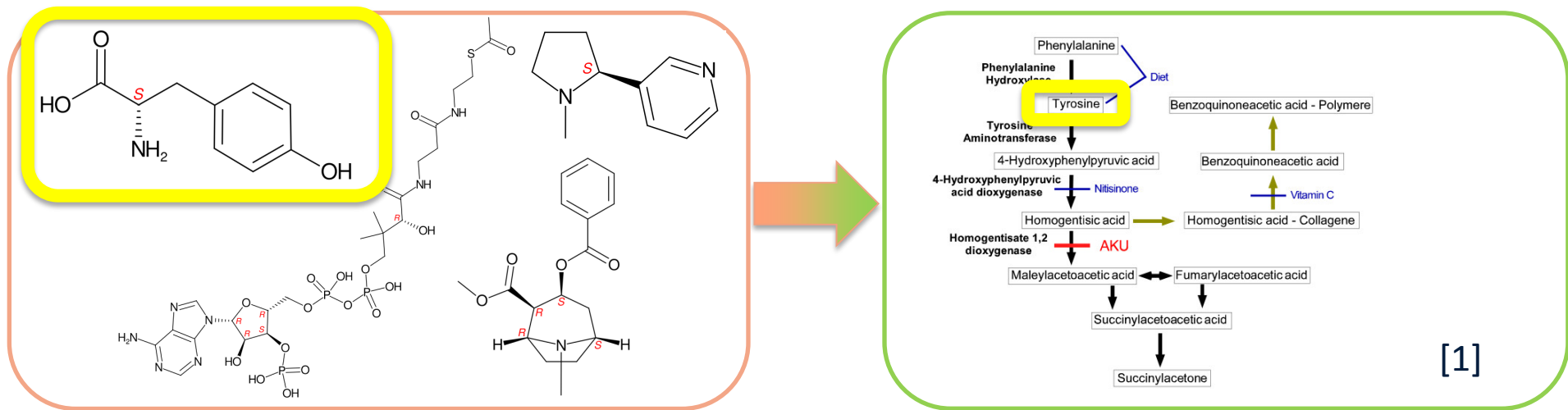
Maastricht University



Linking metabolomics data to pathways...



Linking metabolomics data to pathways...

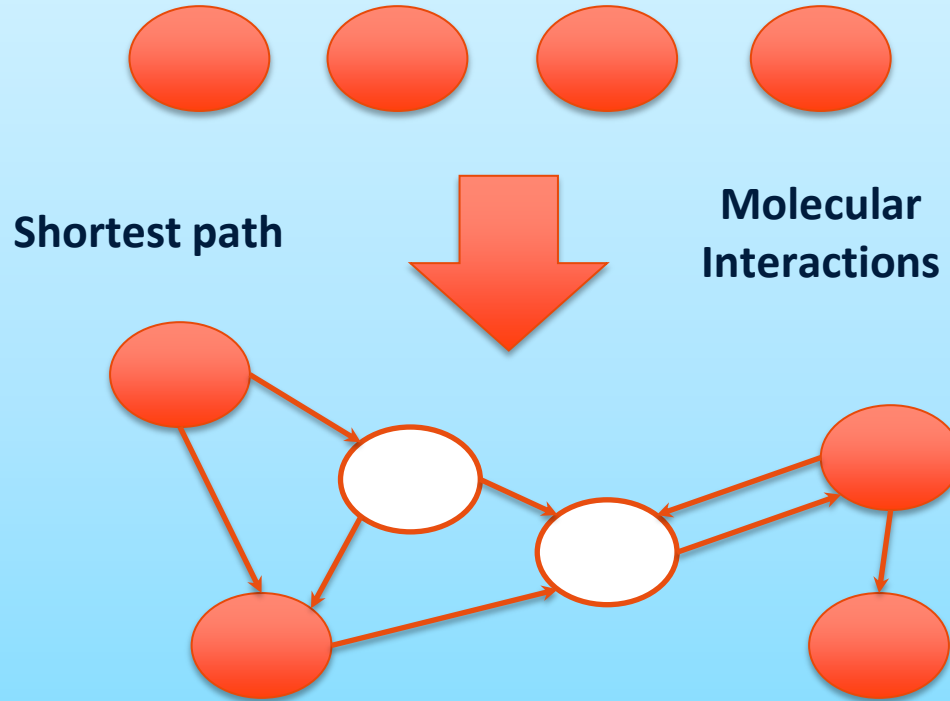


Sparseness of Data

Amount of data

Identifier mapping

Network approach [1]



Aim of Network Approach

- Directed network of metabolites from pathway knowledge bases

Aim of Network Approach

- Directed network of metabolites from pathway knowledge bases
- Calculate sub-network between active metabolites

Aim of Network Approach

- Directed network of metabolites from pathway knowledge bases
- Calculate sub-network between active metabolites
- Visualise directed paths

Aim of Network Approach

- Directed network of metabolites from pathway knowledge bases
- Calculate sub-network between active metabolites
- Visualise directed paths
- Interpret metabolomics datasets

WORKFLOW



Metabolic interactions

Directed metabolic conversions

3 pathway models

Homo sapiens (Human)

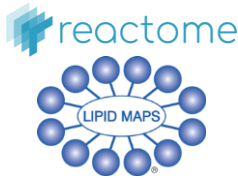
WikiPathways [1] RDF [2]

[1] WikiPathways: <https://www.wikipathways.org>, Slenter *et al.* (2018), DOI: [10.1093/nar/gkx1064](https://doi.org/10.1093/nar/gkx1064)

[2] SPARQL endpoint: <http://sparql.wikipathways.org/>, Waagmeester *et al.* (2016) DOI: [10.1371/journal.pcbi.1004989](https://doi.org/10.1371/journal.pcbi.1004989)

[3] Neo4j: <https://neo4j.com/> [4] Cytoscape: <http://cytoscape.org/>, Shannon *et al.* (2003) DOI: [10.1101/gr.1239303](https://doi.org/10.1101/gr.1239303)

WORKFLOW



Metabolic
interactions



Graph
Database

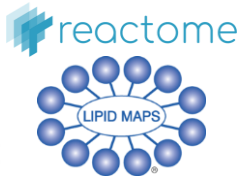
Directed metabolic conversions
3 pathway models
Homo sapiens (Human)

WikiPathways [1] RDF [2]

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]

WORKFLOW



Metabolic
interactions



Graph
Database



Visualisation

Directed metabolic conversions
3 pathway models
Homo sapiens (Human)

WikiPathways [1] RDF [2]

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]

Omics data visualisation
Network extendable
Automatisable (REST-API)

Cytoscape [4]

WORKFLOW



Metabolomics
data

Directed metabolic conversions

3 pathway models

Homo sapiens (Human)

WikiPathways [1] RDF [2]

WikiPathways RDF

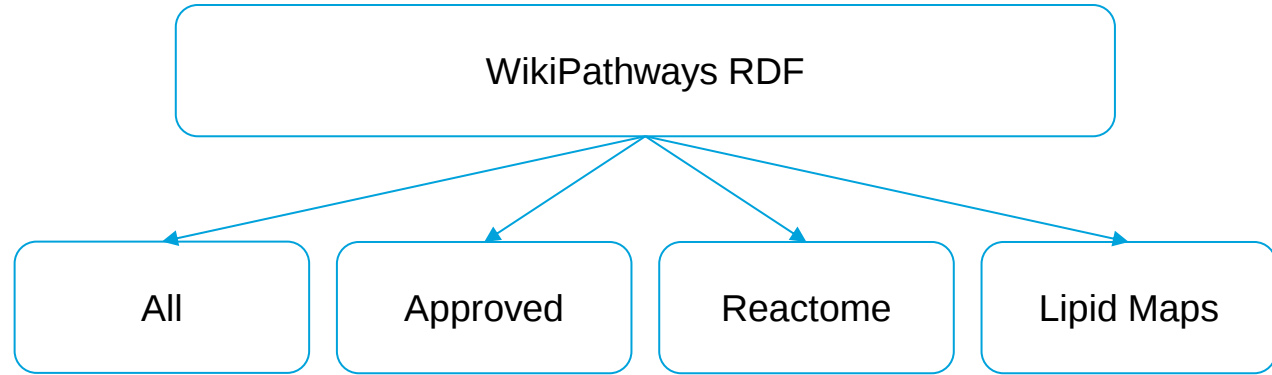
WORKFLOW



Metabolomics
data

Directed metabolic conversions
3 pathway models
Homo sapiens (Human)

WikiPathways [1] RDF [2]



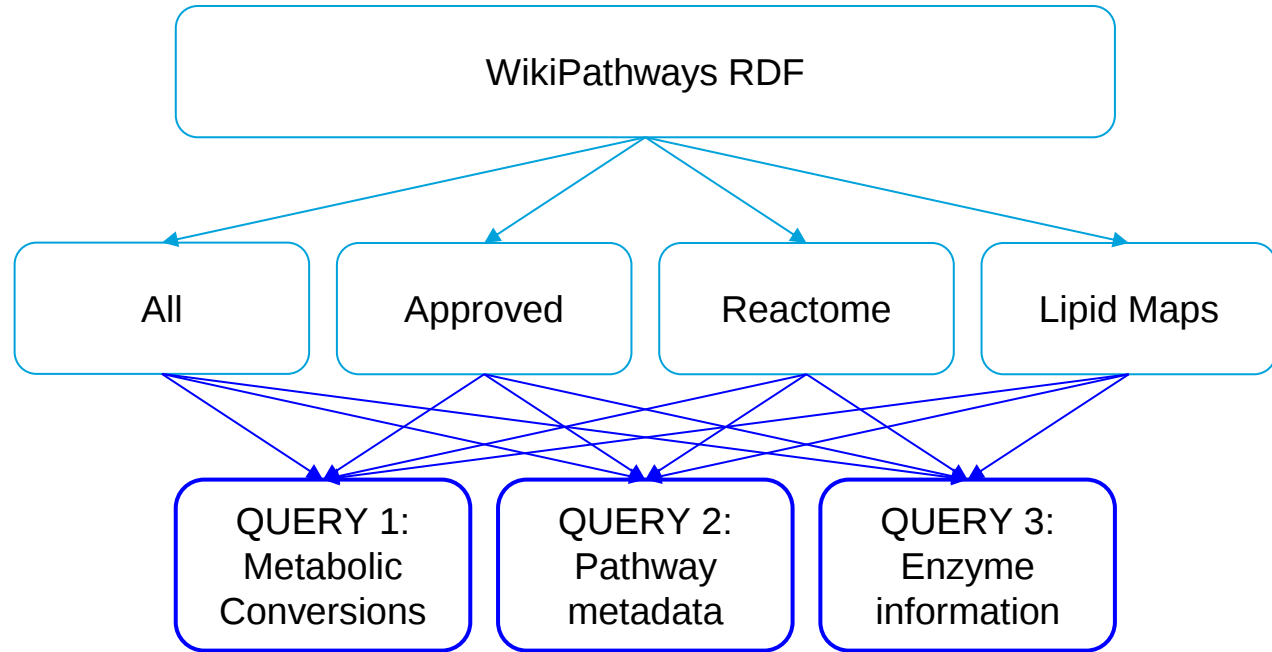
WORKFLOW



Metabolomics
data

Directed metabolic conversions
3 pathway models
Homo sapiens (Human)

WikiPathways [1] RDF [2]



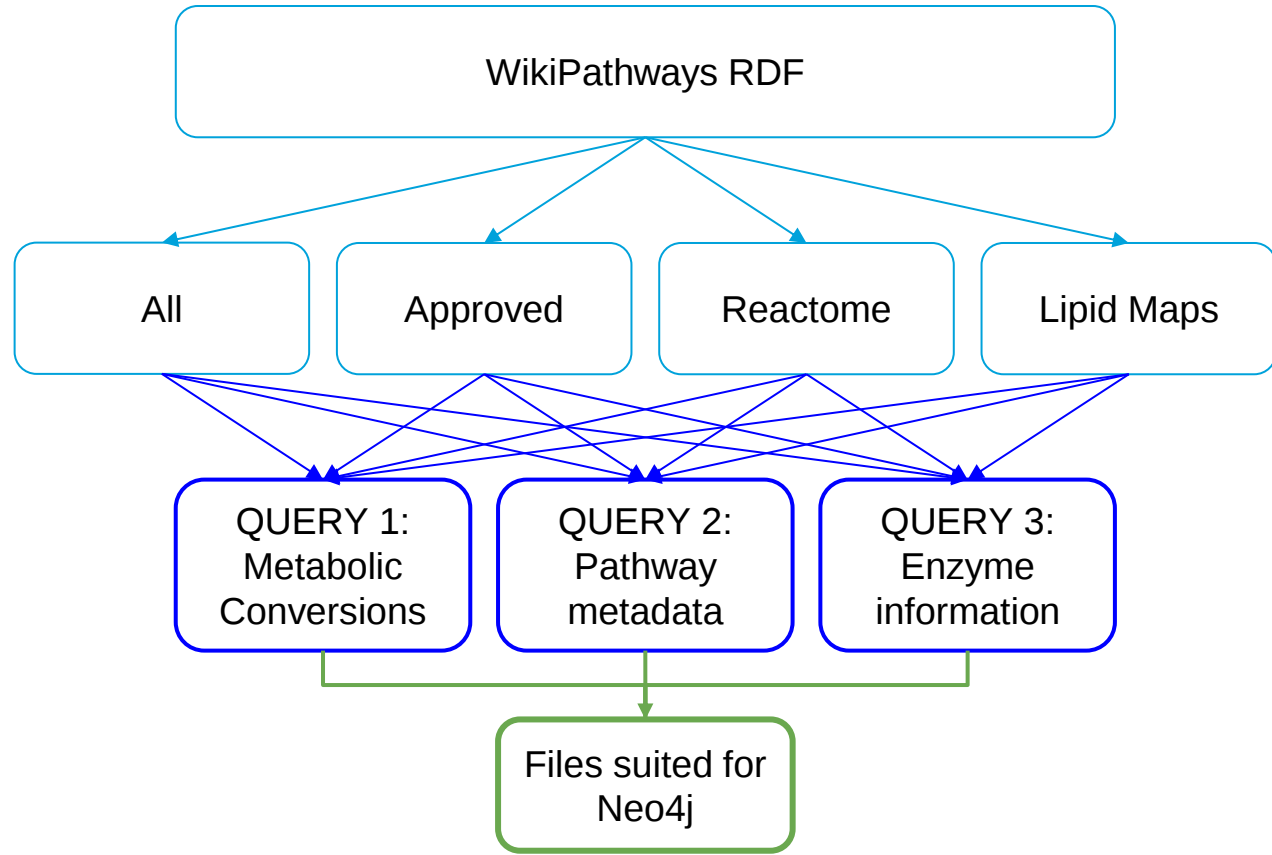
WORKFLOW



Metabolomics
data

Directed metabolic conversions
3 pathway models
Homo sapiens (Human)

WikiPathways [1] RDF [2]



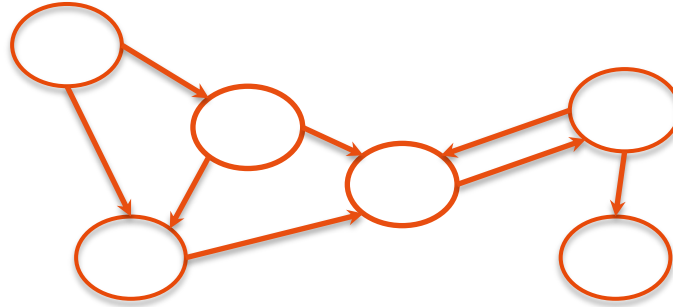
WORKFLOW



Graph
Database

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



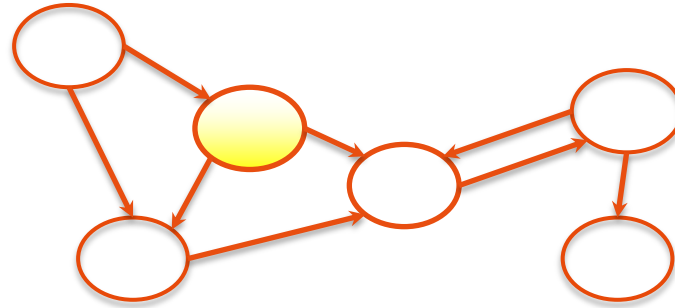
WORKFLOW



Graph Database

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



Property	Value	Identifier
Source	L-Tyrosine	Q188017 (Wikidata)
Target	L-Dopa	Q300989 (Wikidata)
Occurrence	2	WP4220, WP4156
Enzyme	Tyrosine 3-mono oxygenase	P07101 (Uniprot-TrEMBL)

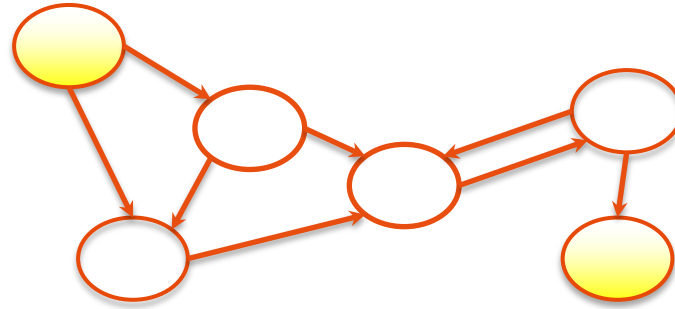
WORKFLOW



Graph
Database

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



Shortest Path

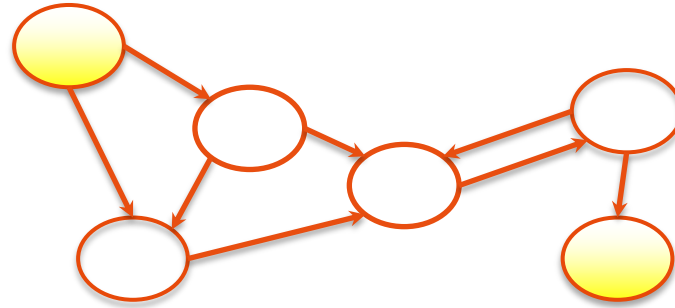
WORKFLOW



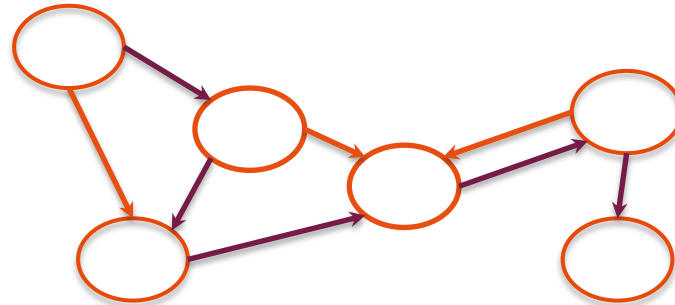
Graph
Database

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



Shortest Path



Steps:

5

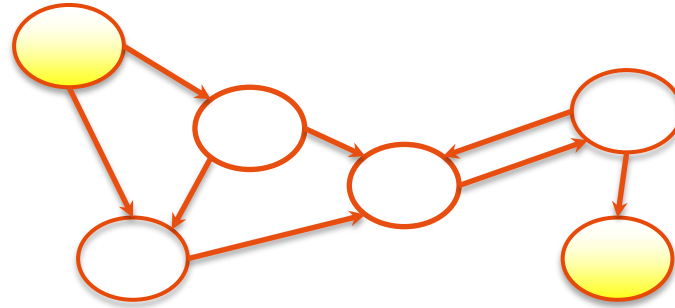
WORKFLOW



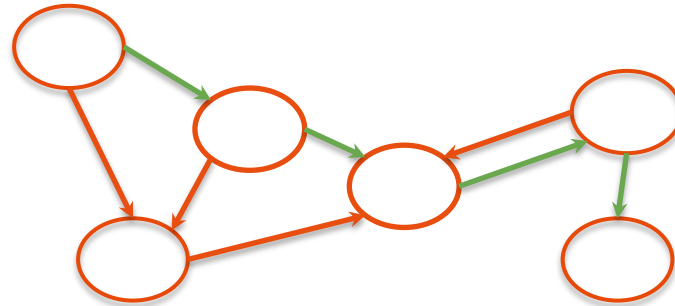
Graph
Database

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



Shortest Path



Steps:

4

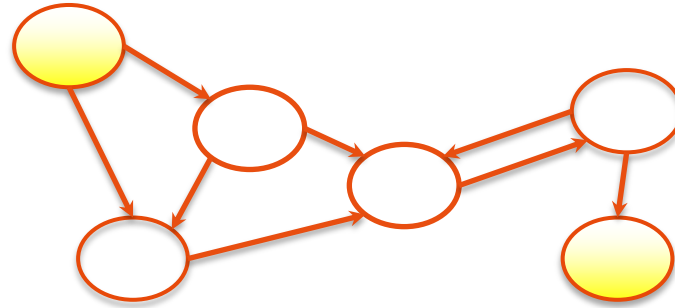
WORKFLOW



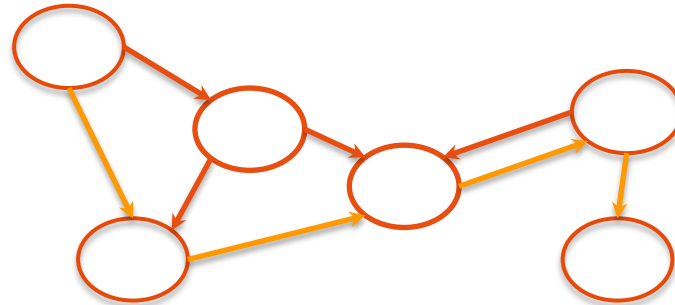
Graph
Database

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



Shortest Path



Steps:

4

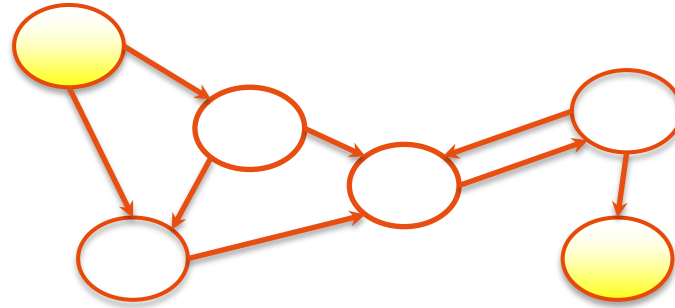
WORKFLOW



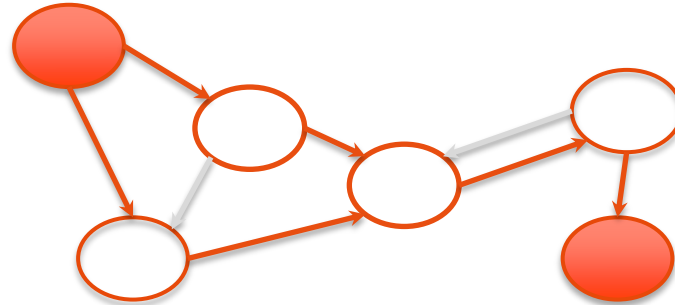
Graph
Database

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



Shortest Path



WORKFLOW



Visualisation

Omics data visualisation
Network extendable
Automatisable (REST-API)

Cytoscape [4]

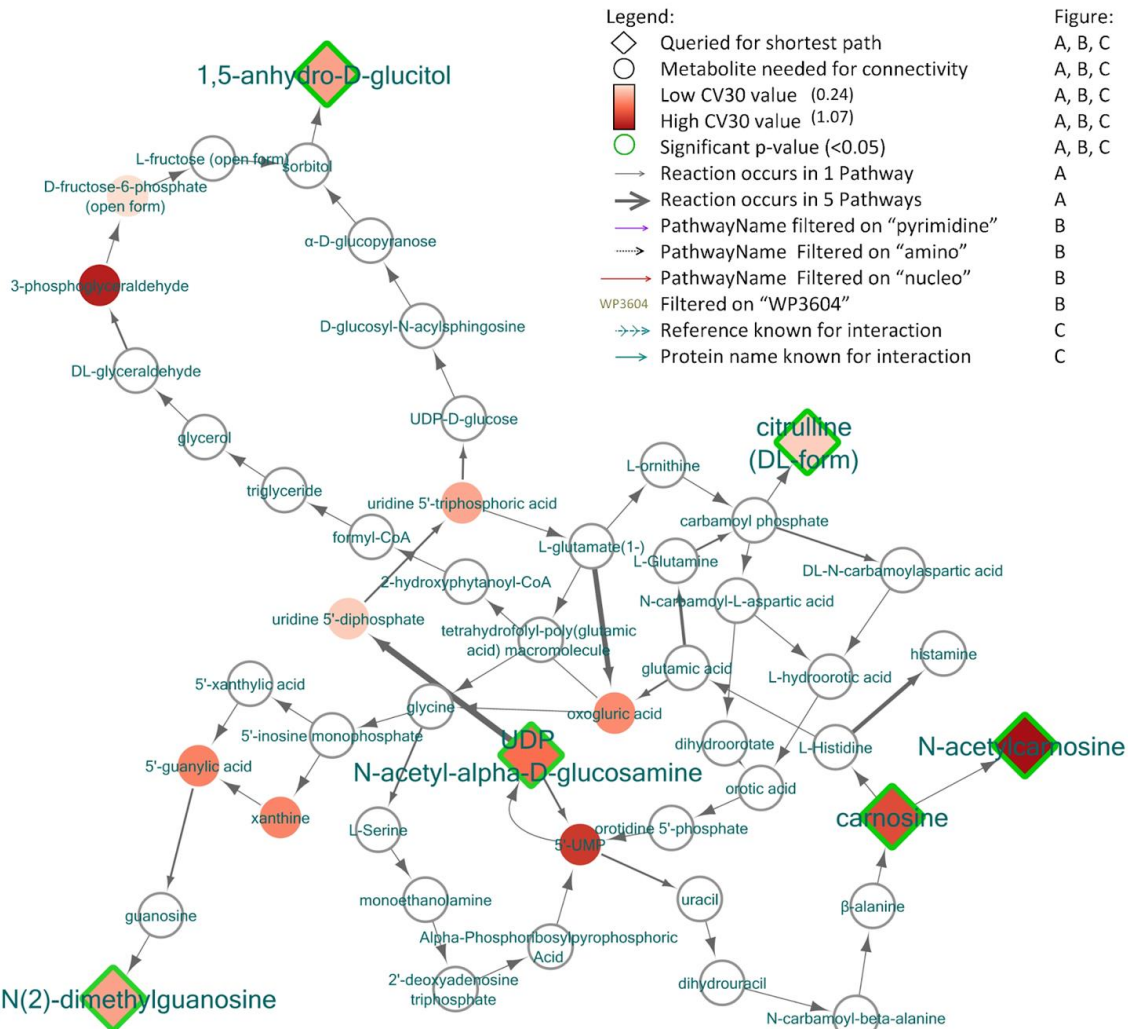
1. Used in biological network community
1. Free to use, source code available online
1. Connectable through CyNeo4j app
1. Automatisable through REST-API, from R and Python
1. Possibility to integrate OMICS data
1. Extendable with CyTargetLinker app (TF, drug-target, SNPs)

Datasets used to test visualisation

Name	Technique	Matrix	Metabolites of interest detected with:
MTBLS265 [1]	LC-MS	Blood	CV-30, p-value
MTBLS404 [2]	LC-HRMS	Urine	Spearman rank correlation test, Orthogonal partial least-squares (OPLS)
Rist <i>et al.</i> [3]	(GCxGC)-MS, targeted GC-MS and LC-MS/MS, 1H-NMR	Blood & Urine	Support Vector Machines (SVM, linear kernel), generalized linear model net (glmnet), Partial least squares (PLS)

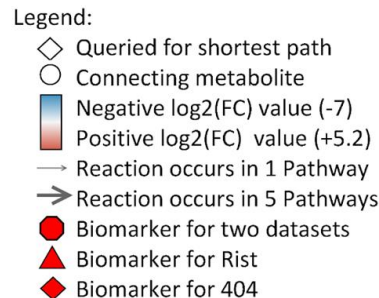
Name	Male/Female	Age range (y)	# Identified metabolites	# age related	Compounds linked through shortest path
MTBLS265	14/16	Young: 29 ± 4 Old: 81 ± 7	126	14	6
MTBLS404	100/83	40.9 ± 10.3	120	30	14
Rist <i>et al.</i>	172/129	47.5 ± 17.1	400 (plasma), >500 (urine)	8 + 6	6 + 4

A

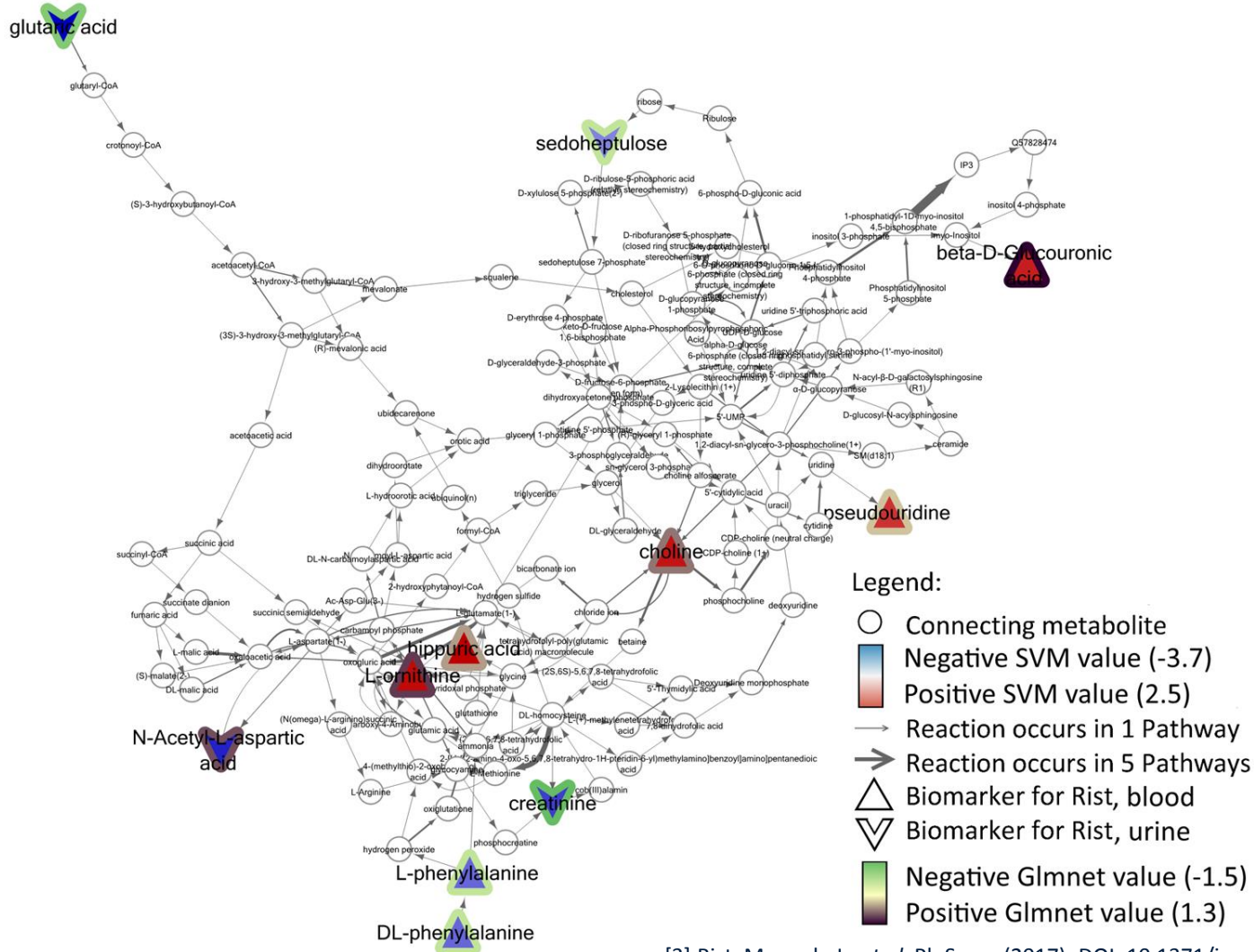




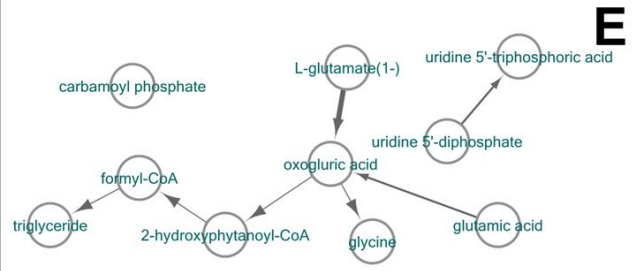
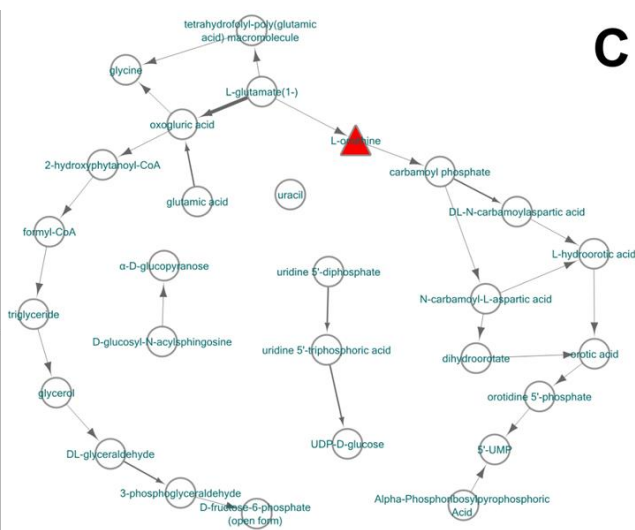
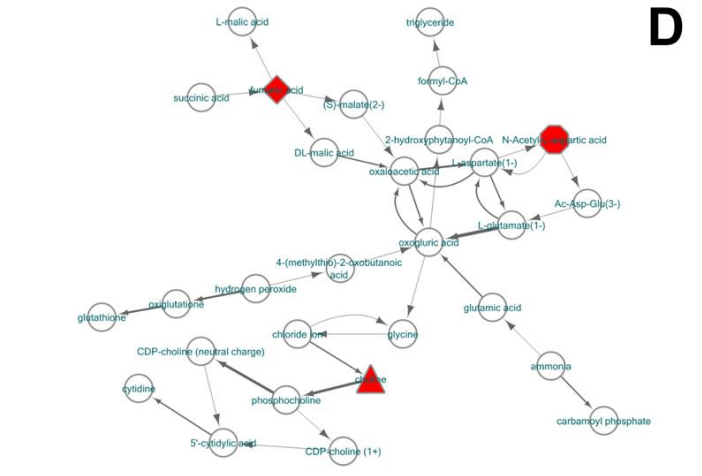
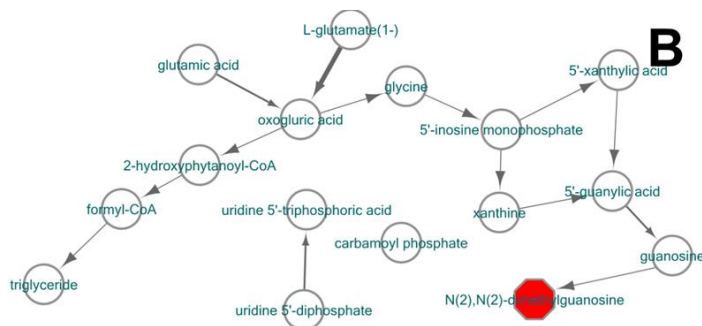
C



A
A - E
A
A
A - E
A - E
B, C
C, D
D



Comparing all



◇ Queried for shortest path	A
○ Connecting metabolite	A - E
■ Negative log2(FC) value (-7)	A
■ Positive log2(FC) value (+5.2)	A
→ Reaction occurs in 1 Pathway	A - E
➔ Reaction occurs in 5 Pathways	A - E
● Biomarker for two datasets	B, C
▲ Biomarker for Rist	C, D
◆ Biomarker for 404	D

B. MTBLS265(blood) vs MTBLS404(urine);

C. MTBLS265(blood) vs Rist(blood and urine); Only overlaps on 1 blood metabolite

D. MTBLS404(urine) vs Rist(blood and urine); Only overlaps on 1 urine metabolite

E. Comparison of all three datasets (MTBLS265, MTBLS404 and Rist).

Take home messages

Calculation of directed subnetwork connecting active metabolites is possible with presented workflow

Pathway analysis of (untargeted) metabolomics data becomes easier, if you use network approaches and graph databases

Workflow is easy for novel users, and adaptable for skilled users, integratable with the data you have and customizable for the analysis you need

Work ahead

Add more pathway knowledge bases → larger coverage of metabolic reactions

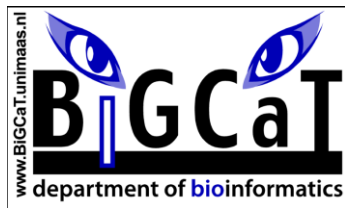
Test app in larger audience → first tests have been run

Create complete R-Markdown script for bulk approach

Allow easy integration of other omics data (proteomics and transcriptomics)

Acknowledgements, questions, discussion

- Martina Kutmon
- Jonathan Mélius
- Ryan Miller
- Georg Summer
- Chris T Evelo
- Egon L Willighagen



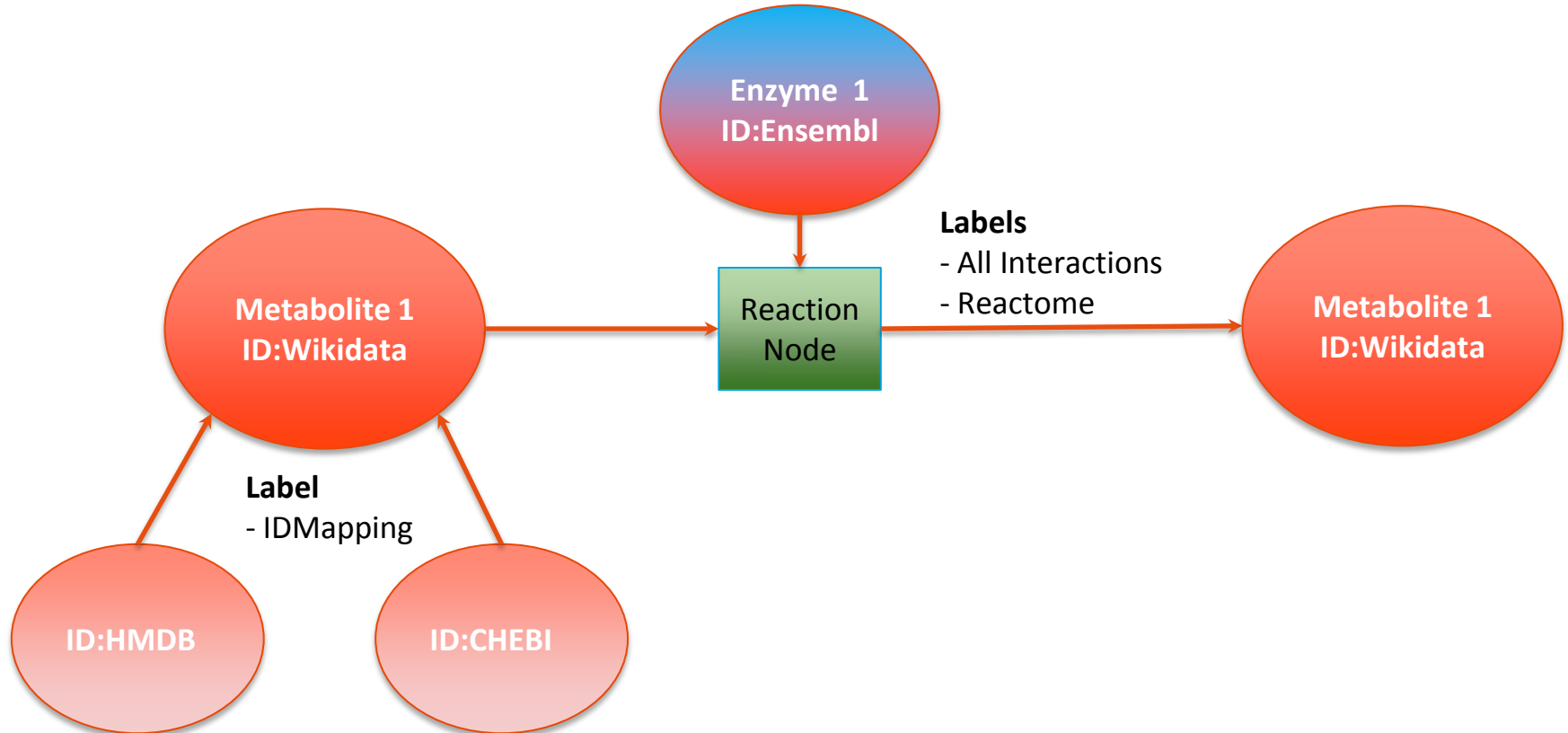
- WikiPathways: team and curators
- Reactome
- Lipid Maps
- Metabolights
- Elixir



- Twitter: @SMaLLCaT4Sci and @BiGCaT_UM
- Blog: <http://smallcats4science.blogspot.nl>
- ORCID: 0000-0001-8449-1318

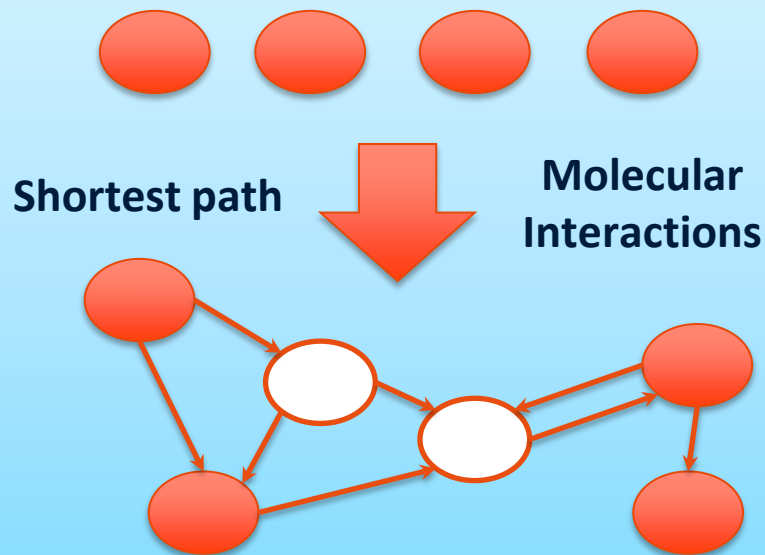
Find the original poster @ DOI: [10.6084/m9.figshare.5234851.v1](https://doi.org/10.6084/m9.figshare.5234851.v1)

Neo4j database details

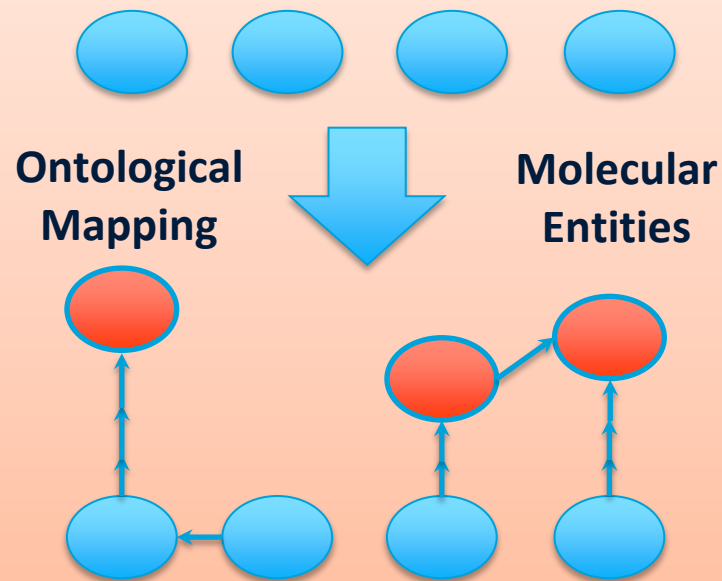


Two approaches

Network approach [1]

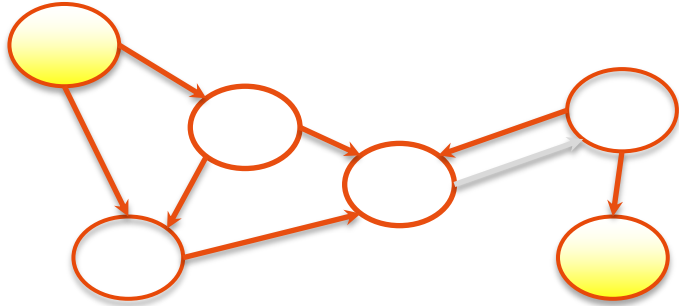


Ontological approach [2]



Biological role					
Electron donor/receiver.		Energy donor/receiver.		Miscellaneous, relevant for various metabolic reactions.	
Identifier	Name	Identifier	Name	Identifier	Name
Q5203615	O ₂	Q80863	ATP	Q307434	S-adenosyl-L-homocysteine
Q506710	H ⁺	Q185253	ADP	Q201312	S-adenosyl-L-methionine
Q20856948	Na ⁺ (redirected to Q3154110)	Q318369	AMP	Q407635	Coenzyme-A
Q3154110	Na ⁺	Q422582	GDP	Q715317	Acetyl coenzyme a
Q283	H ₂ O	Q392227	GTP	Side metabolites	
Q1997	CO ₂	Q26987754	NADP ⁺		
Q177811	PO ₄ 3-	Q26841327	NADPH		
Q411092	Pyrophosphoric acid	Q26987253	NAD ⁺		
Q190901	ammonium cation	Q26987453	NADH		
		Q27102690	FADH ₂		

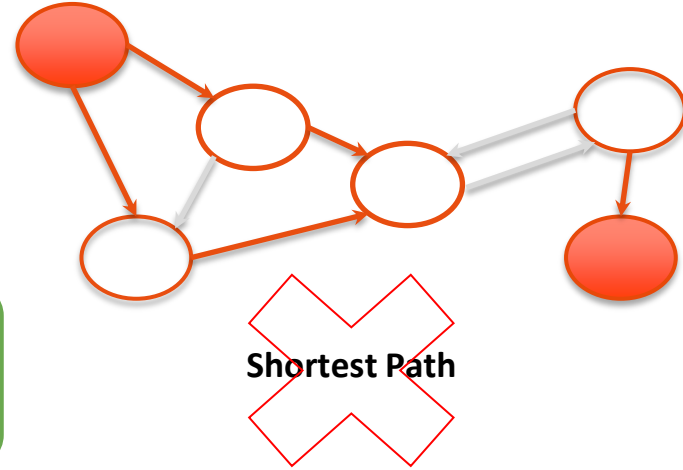
WORKFLOW



Graph
Database

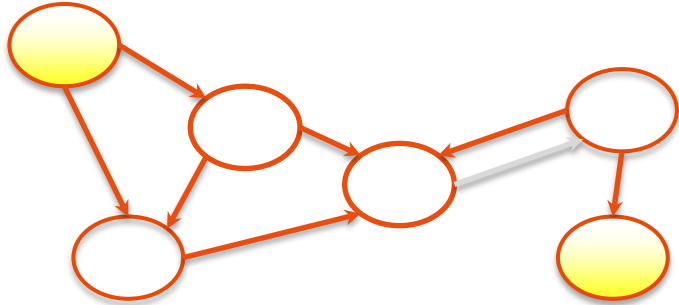
Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



Nearest Neighbour

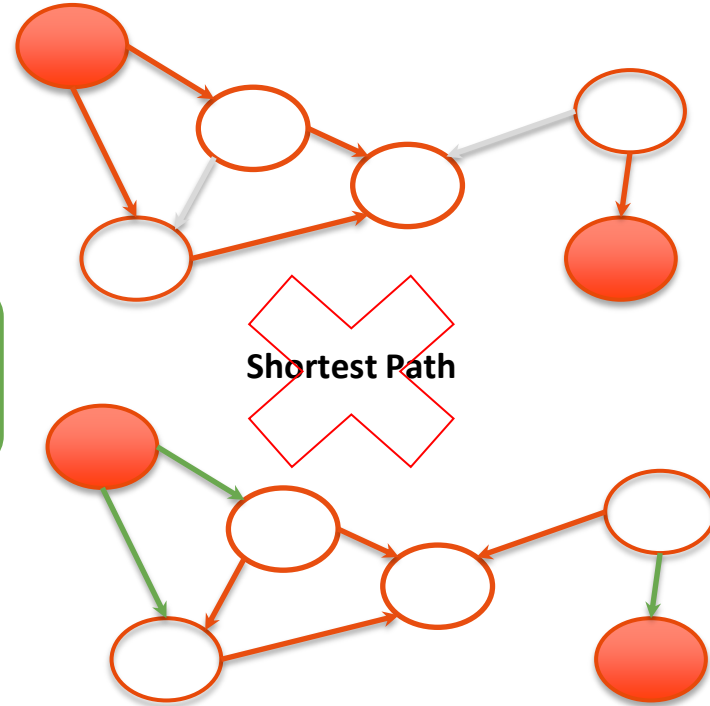
WORKFLOW



Graph
Database

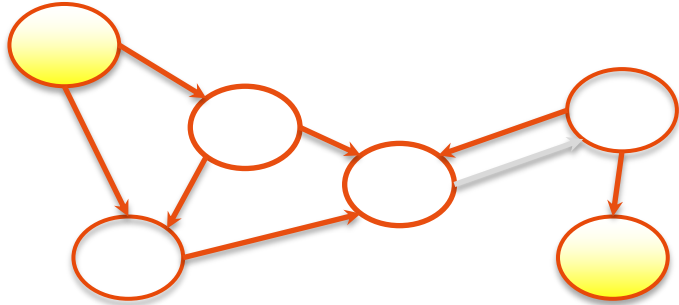
Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



Nearest Neighbour

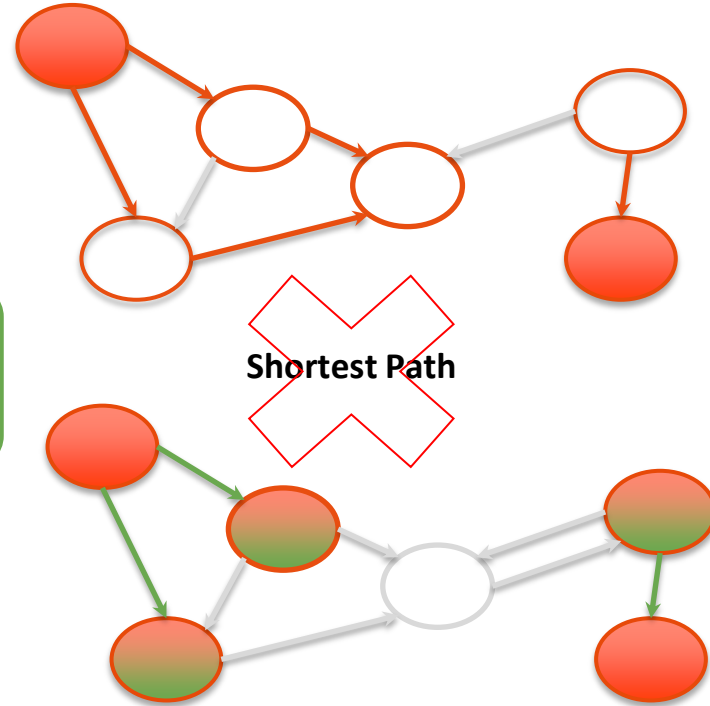
WORKFLOW



Graph
Database

Store conversions
Cypher queries
Several available
algorithms

Neo4j [3]



Nearest Neighbour



Maastricht University