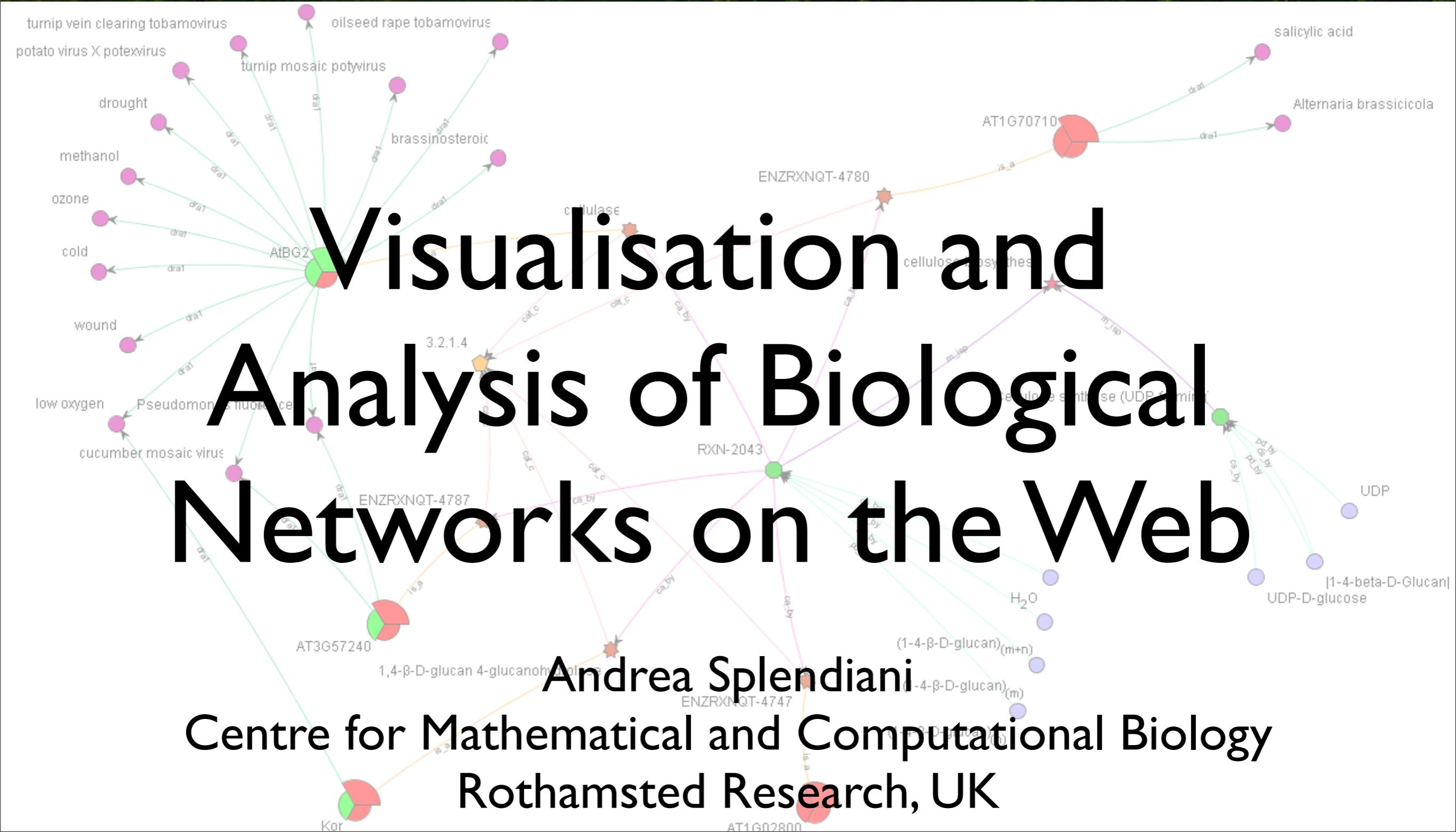




Visualisation and Analysis of Biological Networks on the Web

Andrea Splendiani

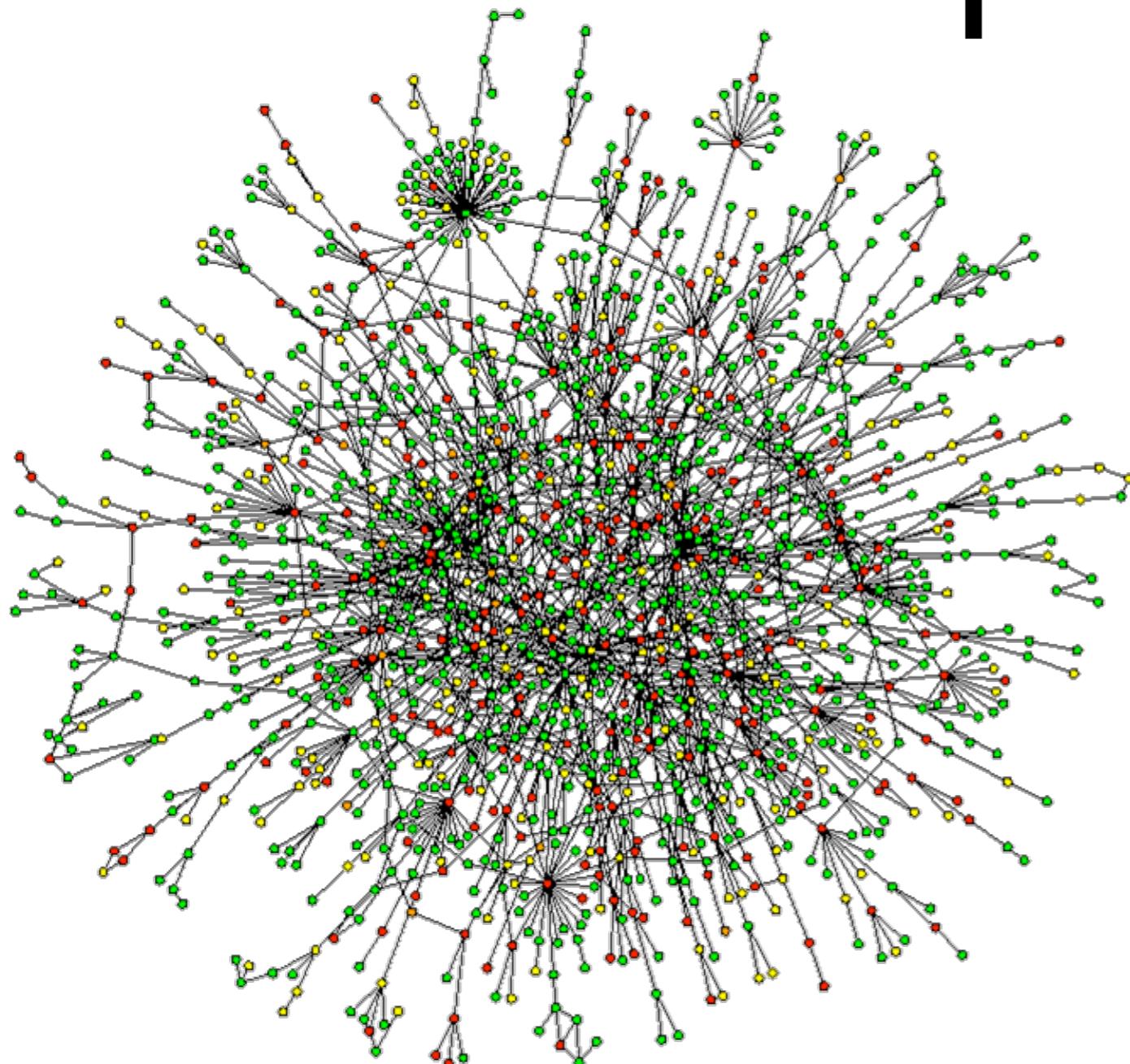
Centre for Mathematical and Computational Biology
Rothamsted Research, UK



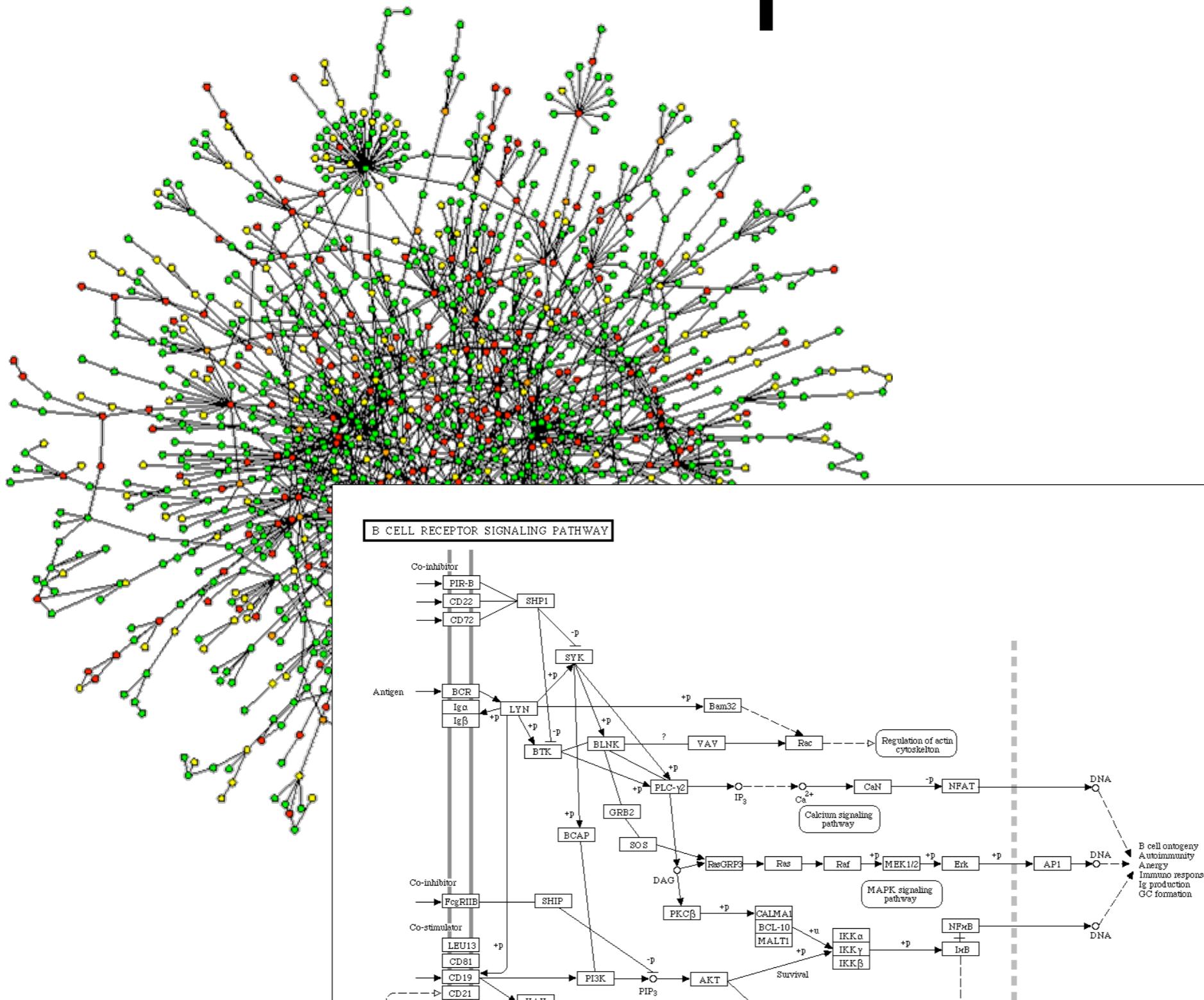
The parts

Andrea Splendiani, Visualisation and Analysis of Biological Networks on the Web
BioHackathon3 Symposium

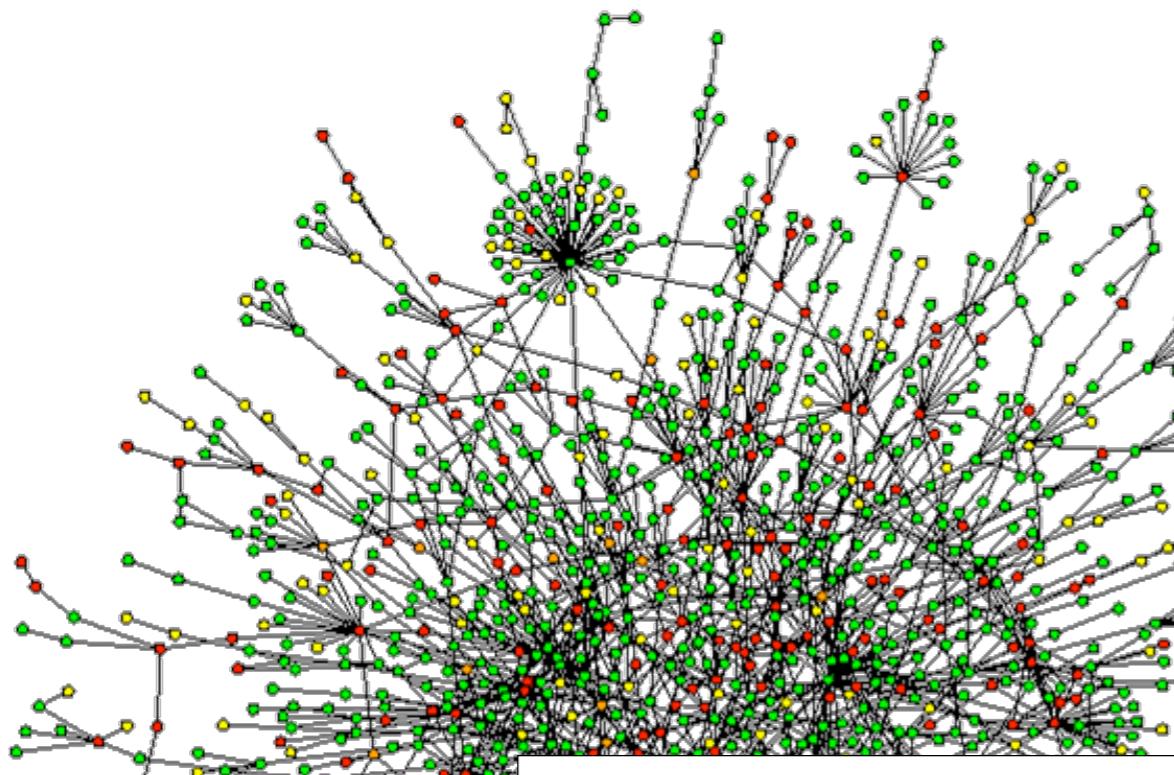
The parts



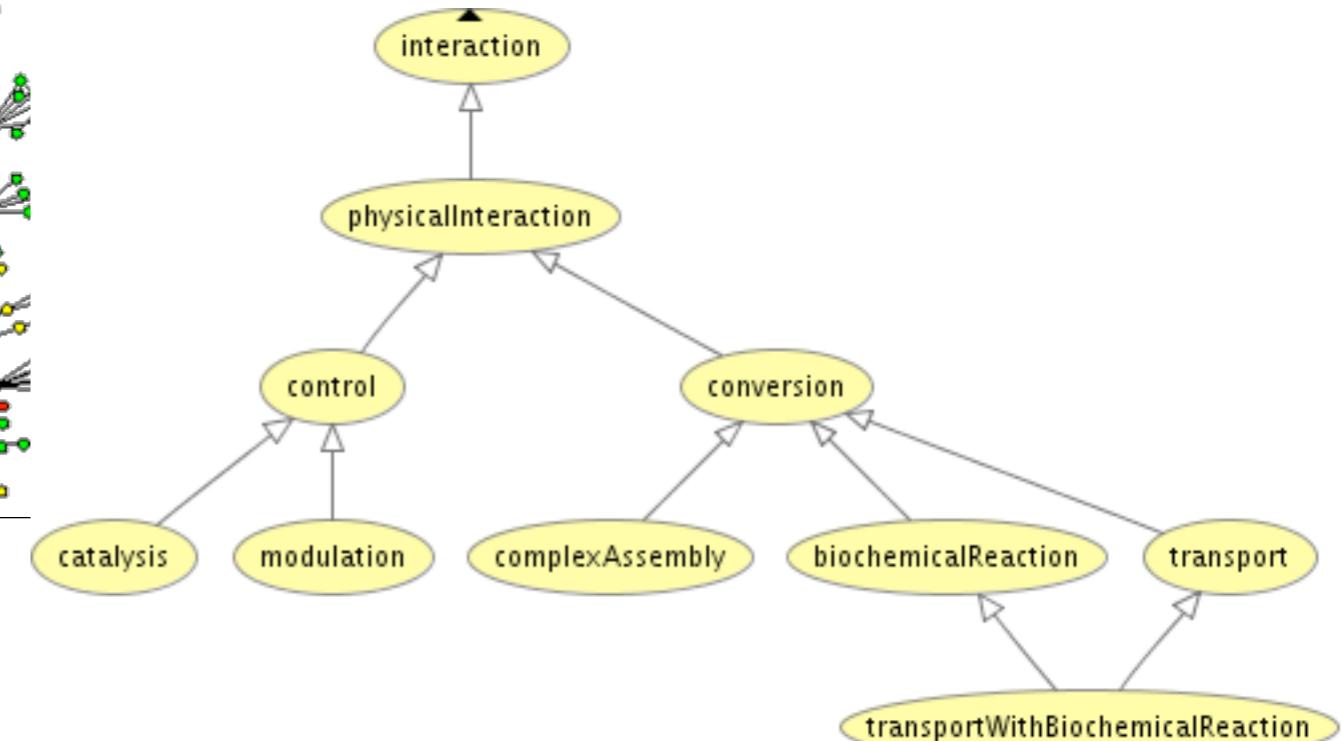
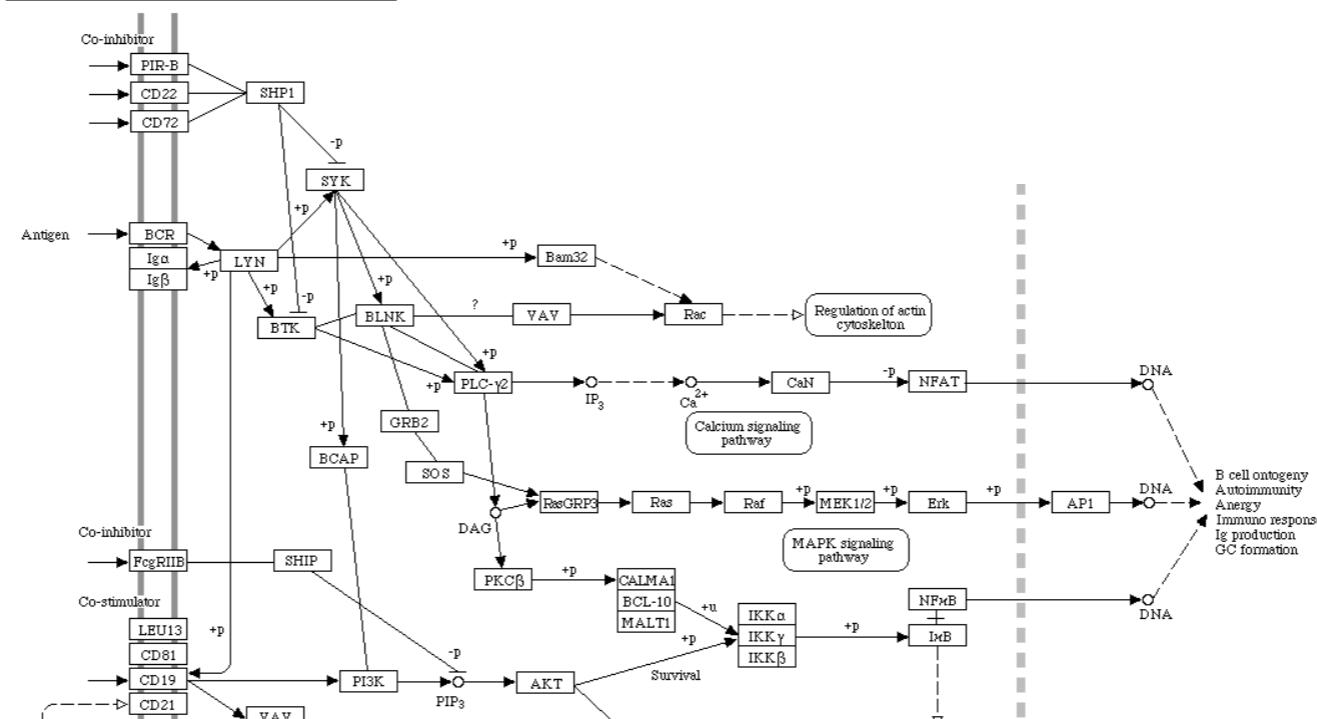
The parts



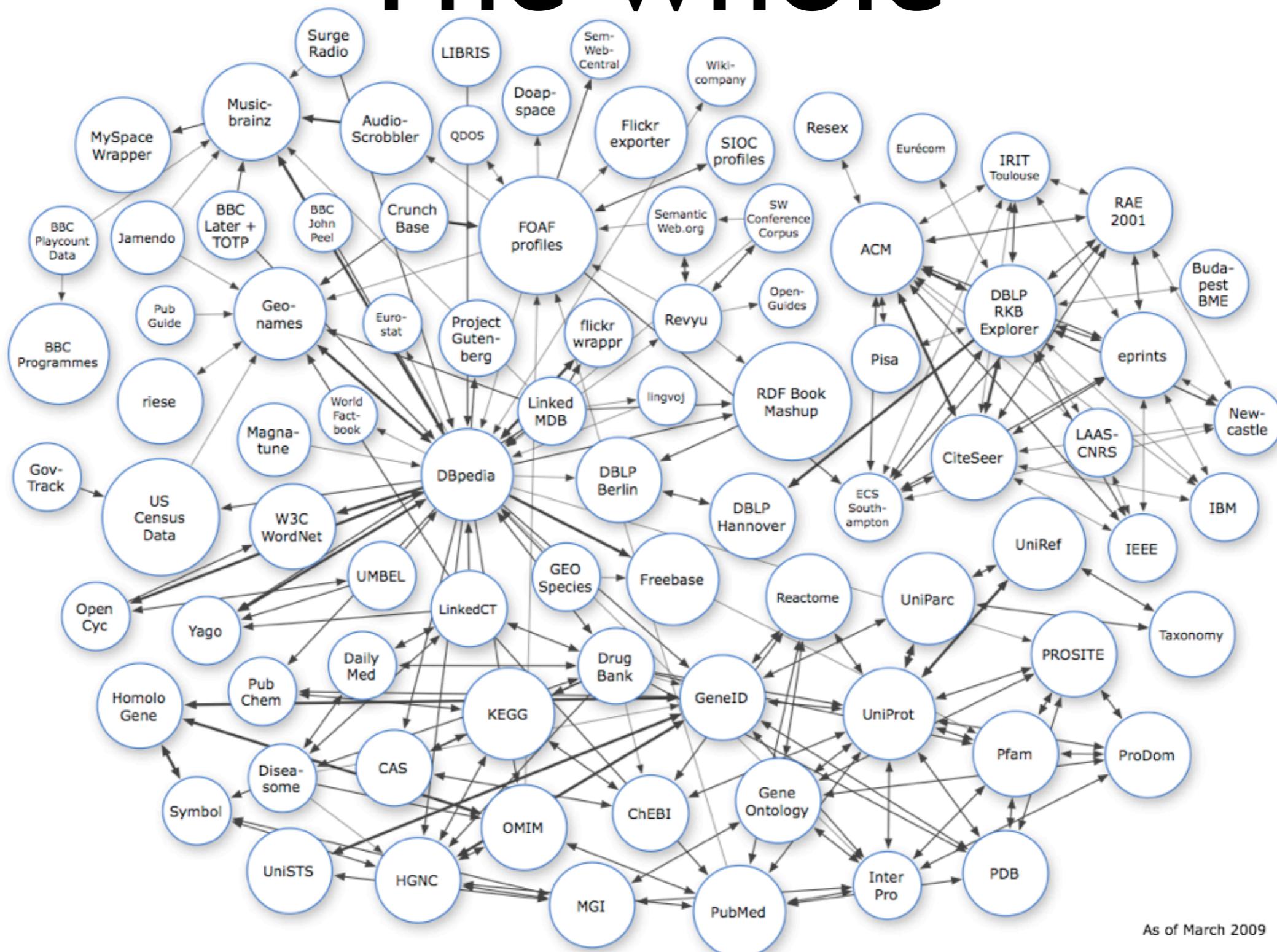
The parts



B CELL RECEPTOR SIGNALING PATHWAY



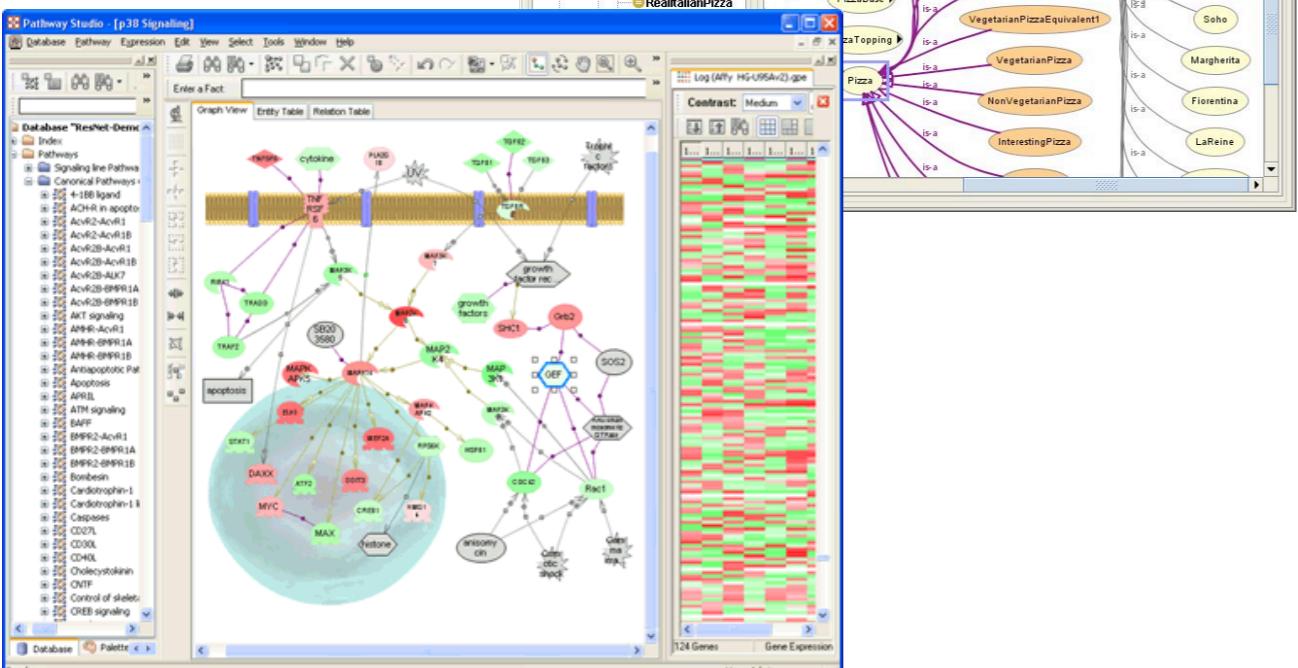
The whole



As of March 2009

Interfaces

Parts



Query

Default Graph URI

Security restrictions of this server do not allow you to retrieve remote RDF data. DBA may wish to grant privilege to "SPARQL" account to remove the restriction. In order to do this, please perform the following steps:

1. Go to the Virtuoso Administration Conductor i.e. <http://localhost:8892/conductor>
2. Login as dba user
3. Go to System Admin->User Accounts->Roles
4. Click the link "Edit" for "SPARQL_SPONGE"
5. Select from the list of available user/groups "SPARQL" and click the ">>" button so to add it to the role
6. Click the button "Update"
7. Access again the sparql endpoint in order to be able to retrieve remote data.

Query text

```
SELECT COUNT (distinct ?graph) WHERE {GRAPH ?graph {?s ?p ?o}}
```

Display Results As: Rigorous check of the query

Whole



Interfaces

DBpedia search powered by

RelFinder

URL

between examples

(1) Albert Einstein
(2) Kurt Gödel

add **clear** **Find Relations**

Filter by:

- length
- class
- link
- connection

number of objects	num	vi
1	4/4	
2	8/8	

item type

start typing...
Species (62)
Eukaryote (57)
Fungus (51)

born in

start typing...
United States (2)
Los Angeles (1)
Missouri (1)

released in year

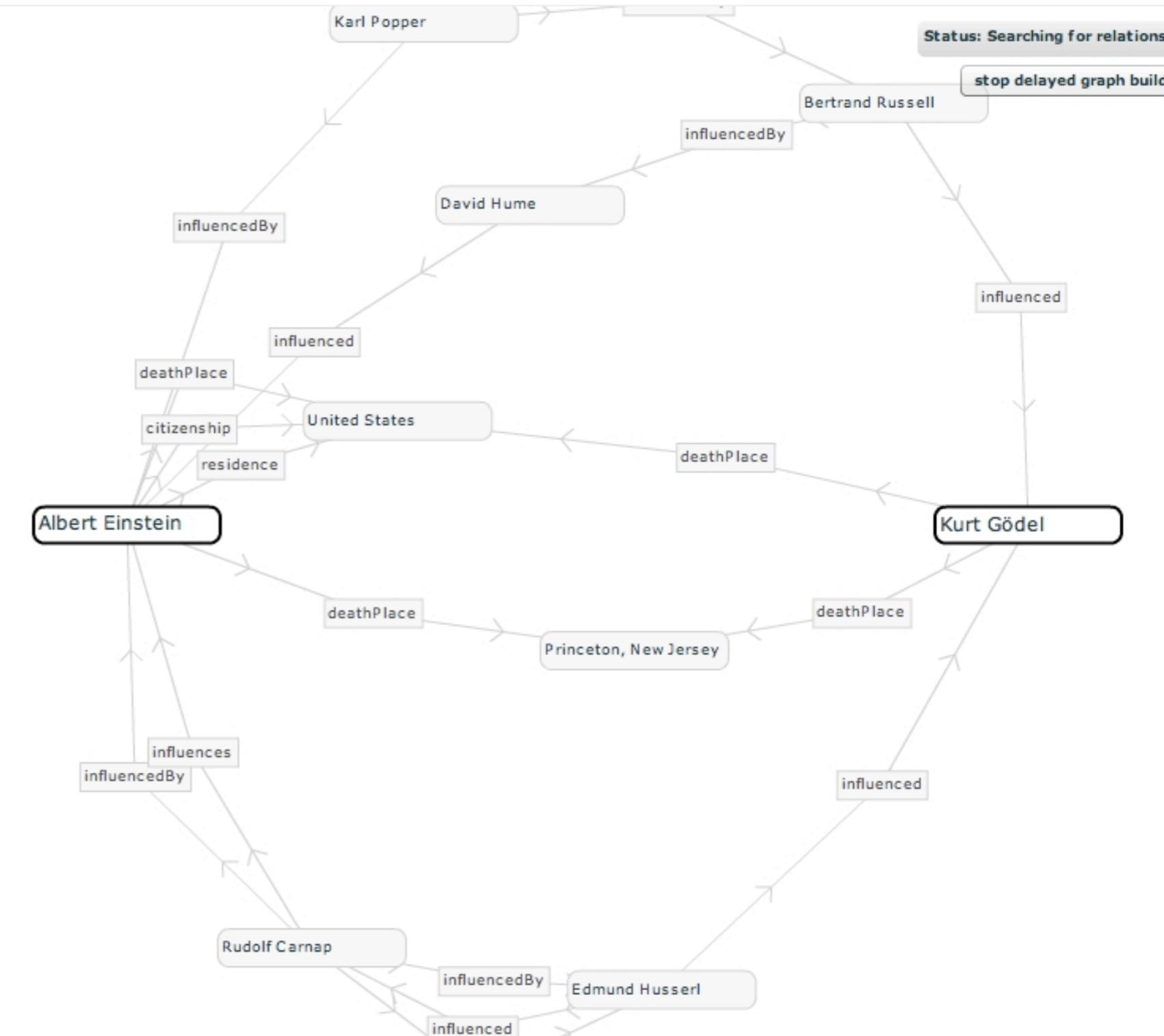
start typing...
from... to...
1997 (1)
1995 (1)
2005 (1)

born in year

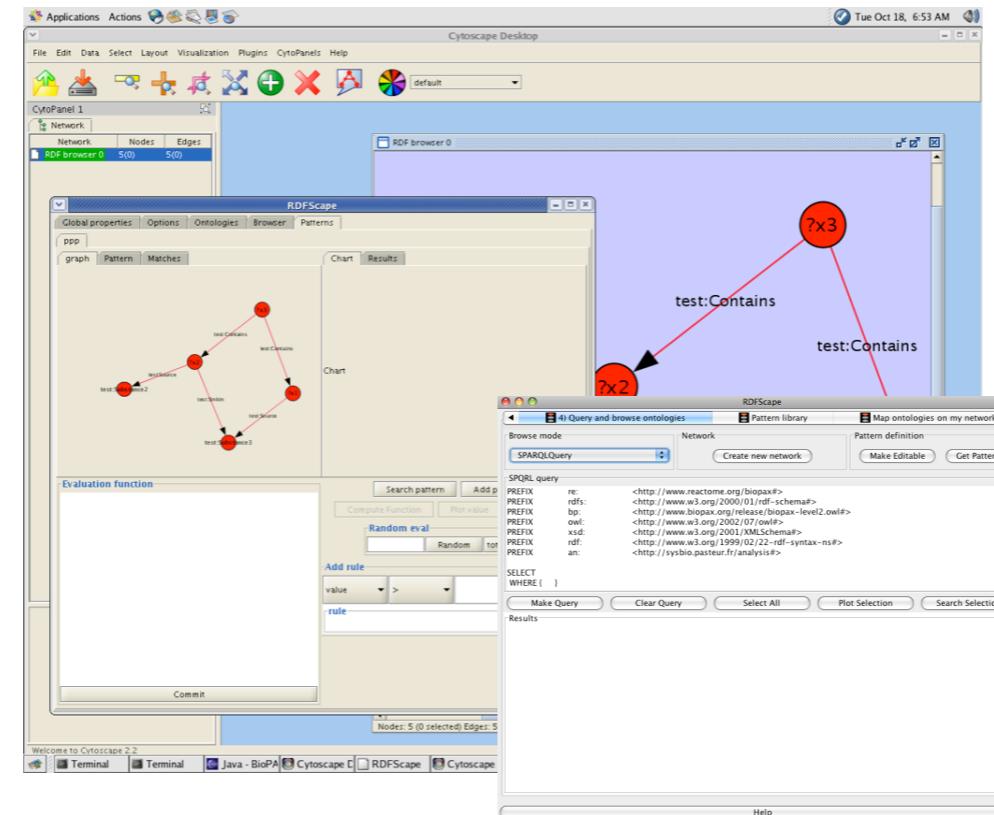
start typing...
from... to...
1891 (1)

Source:

<http://relnet.dbpedia.org/>



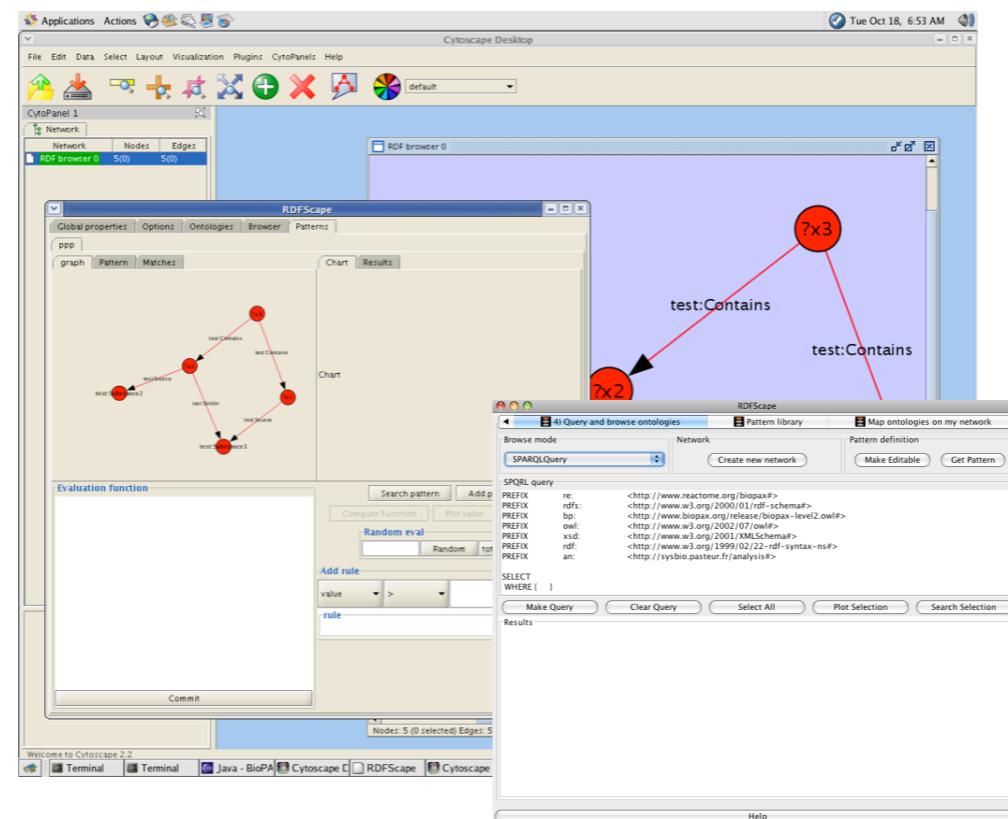
RDFScape



Andrea Splendiani, Visualisation and Analysis of Biological Networks on the Web
BioHackathon3 Symposium

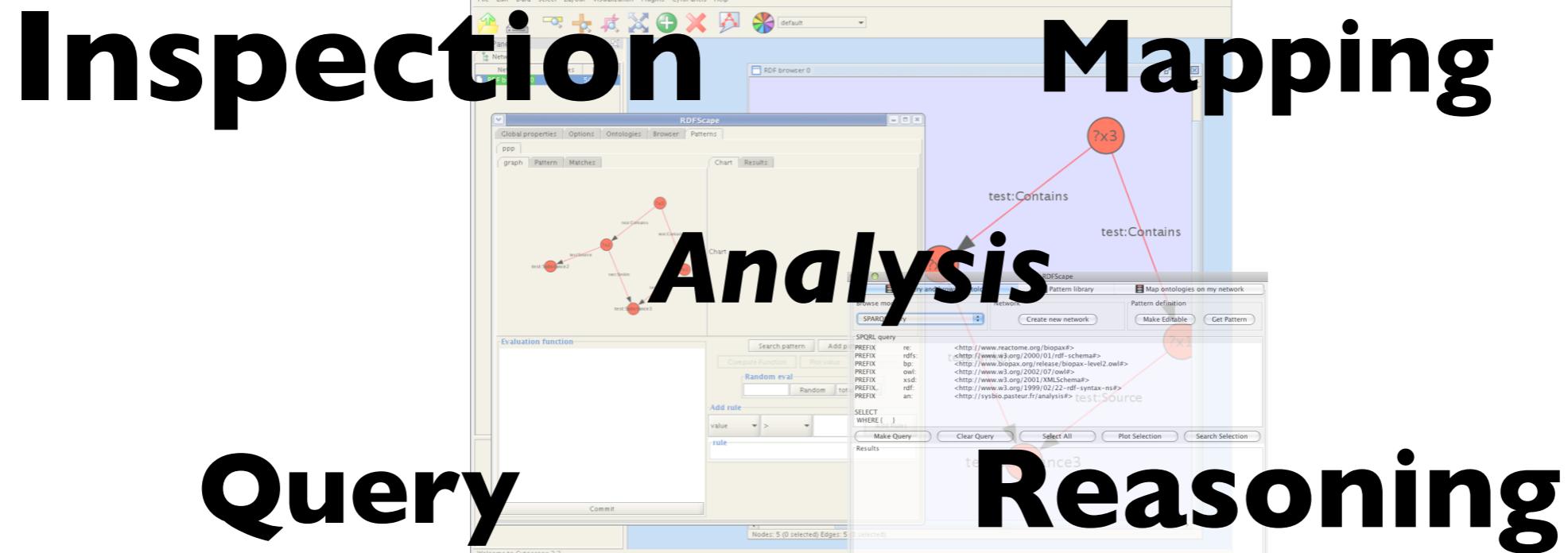
RDFscape = Interface

Using Cytoscape
for interaction
on the Semantic
Web



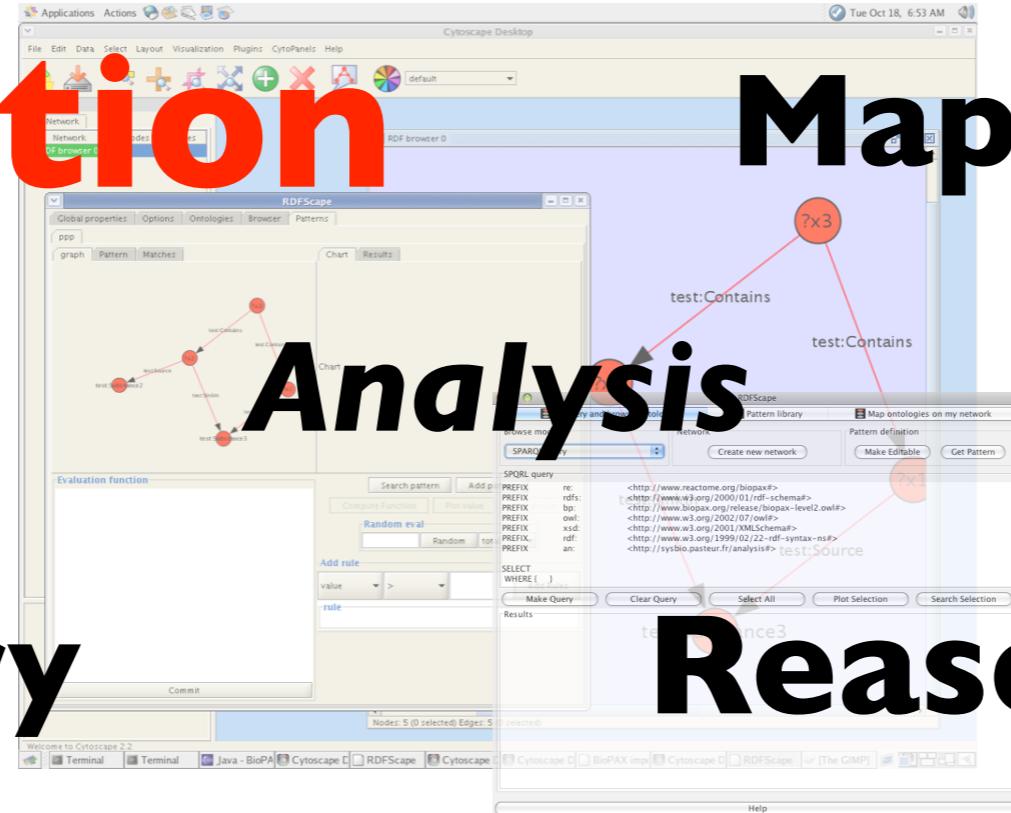
Using the
Semantic Web
for the analysis
of networks in
Cytoscape

RDFscape = Interface



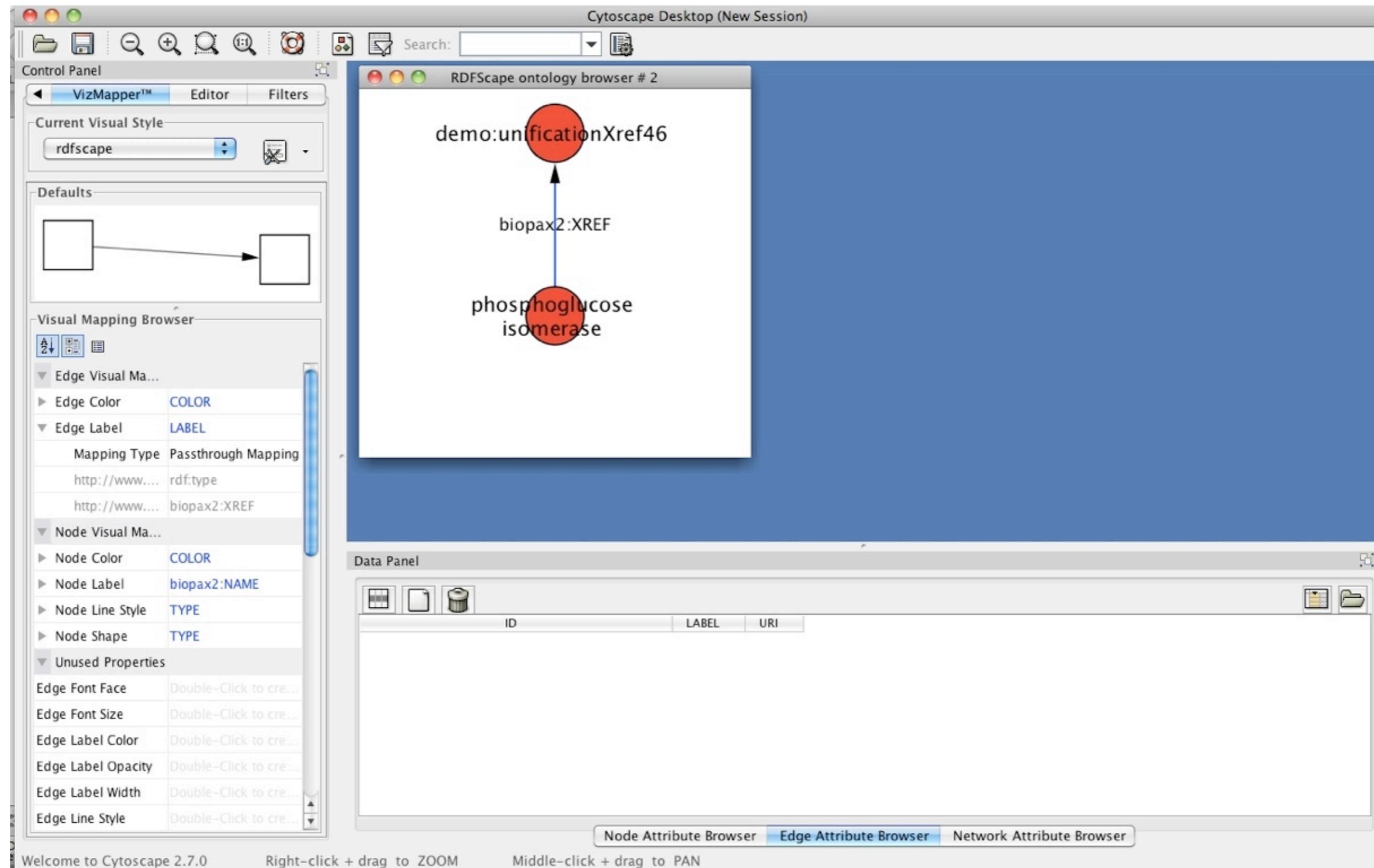
RDFscape = Interface

Inspection Mapping
Query Analysis Reasoning

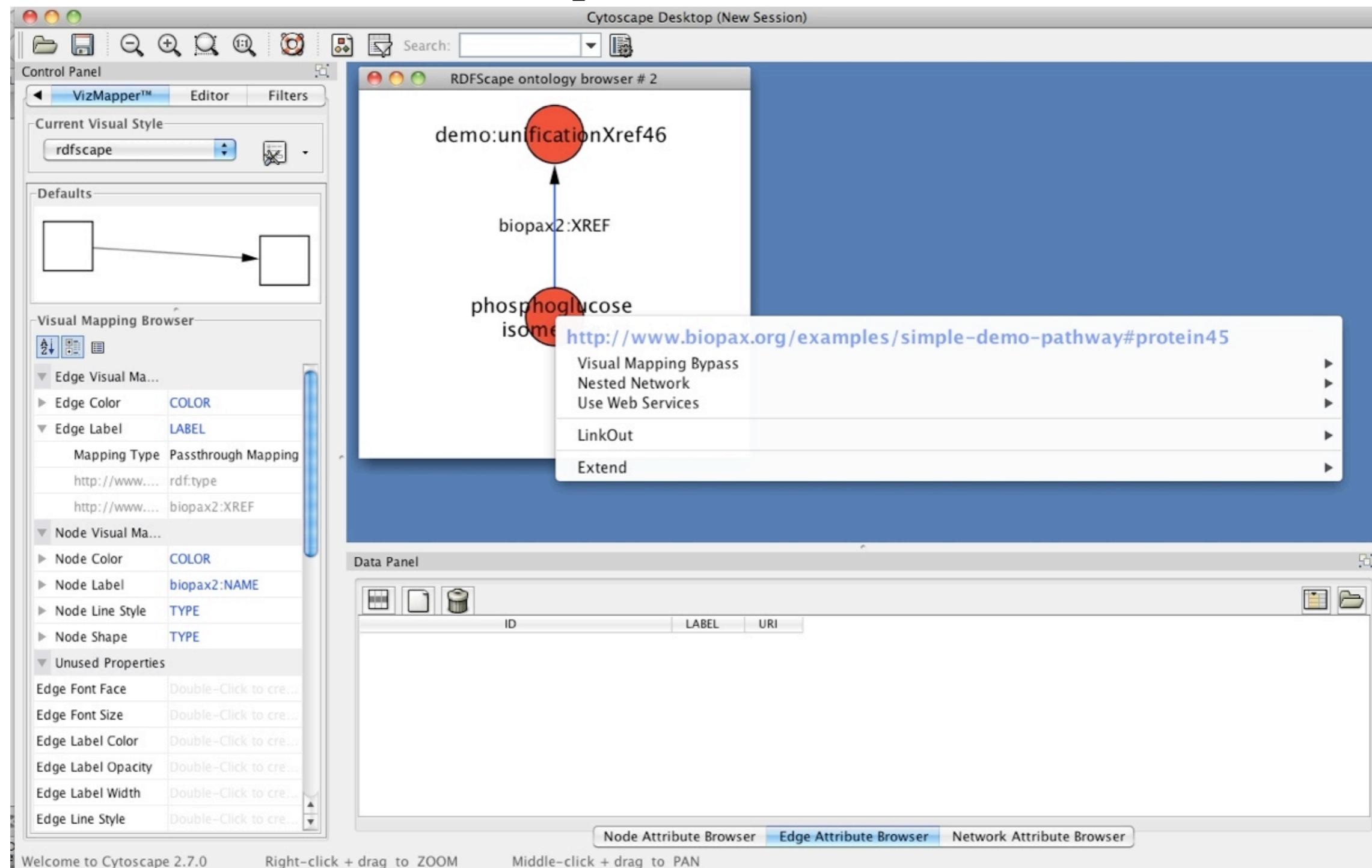


The screenshot shows the RDFscape interface integrated into the Cytoscape desktop environment. The main window displays a network graph with nodes and edges. Overlaid on the graph are several large, bold text labels: 'Inspection' (in red) on the left, 'Mapping' on the right, 'Query' at the bottom left, 'Analysis' in the center, and 'Reasoning' at the bottom right. The background shows the Cytoscape interface with various panels like 'RDFscape', 'Evaluation function', 'SPARQL query', and 'Results'.

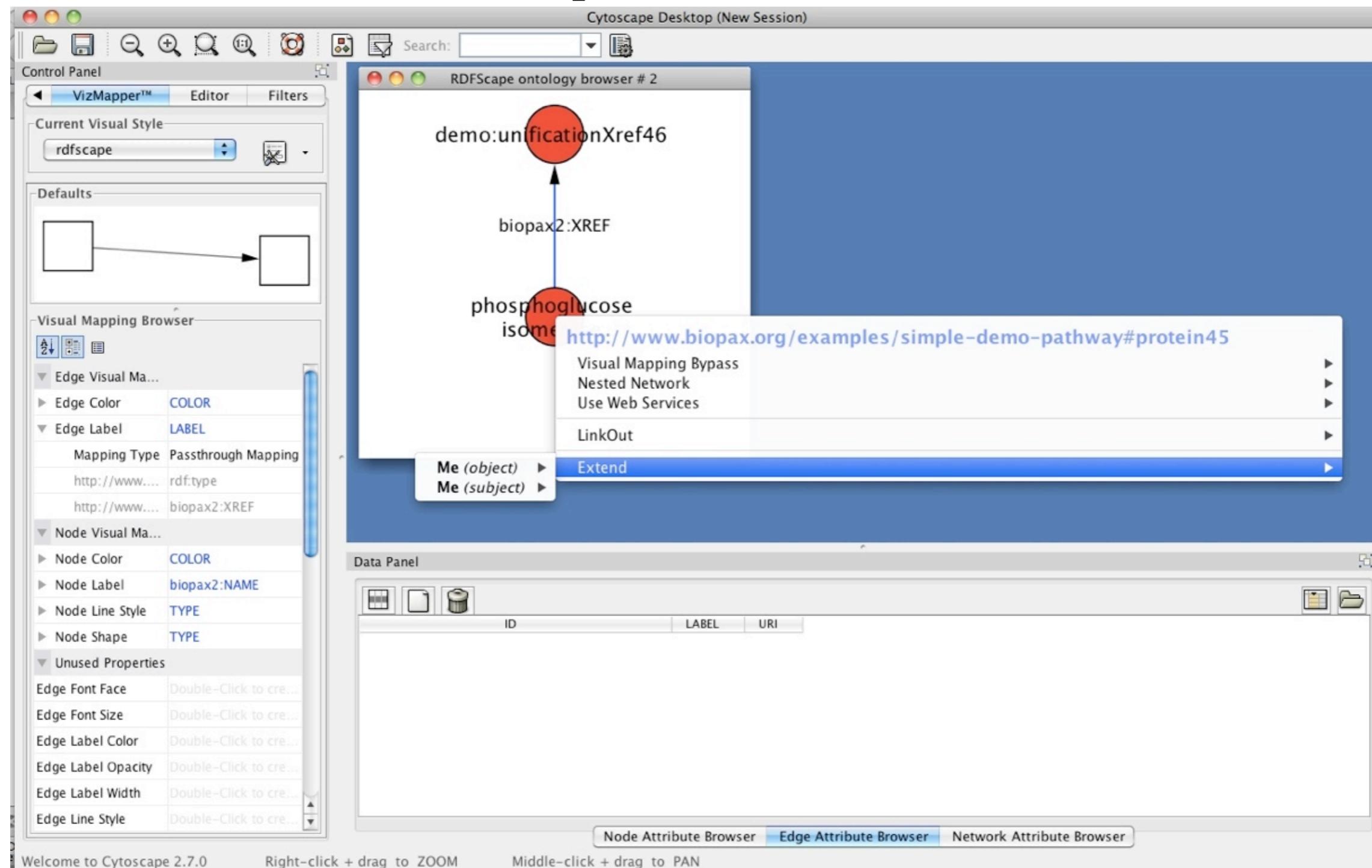
Inspection



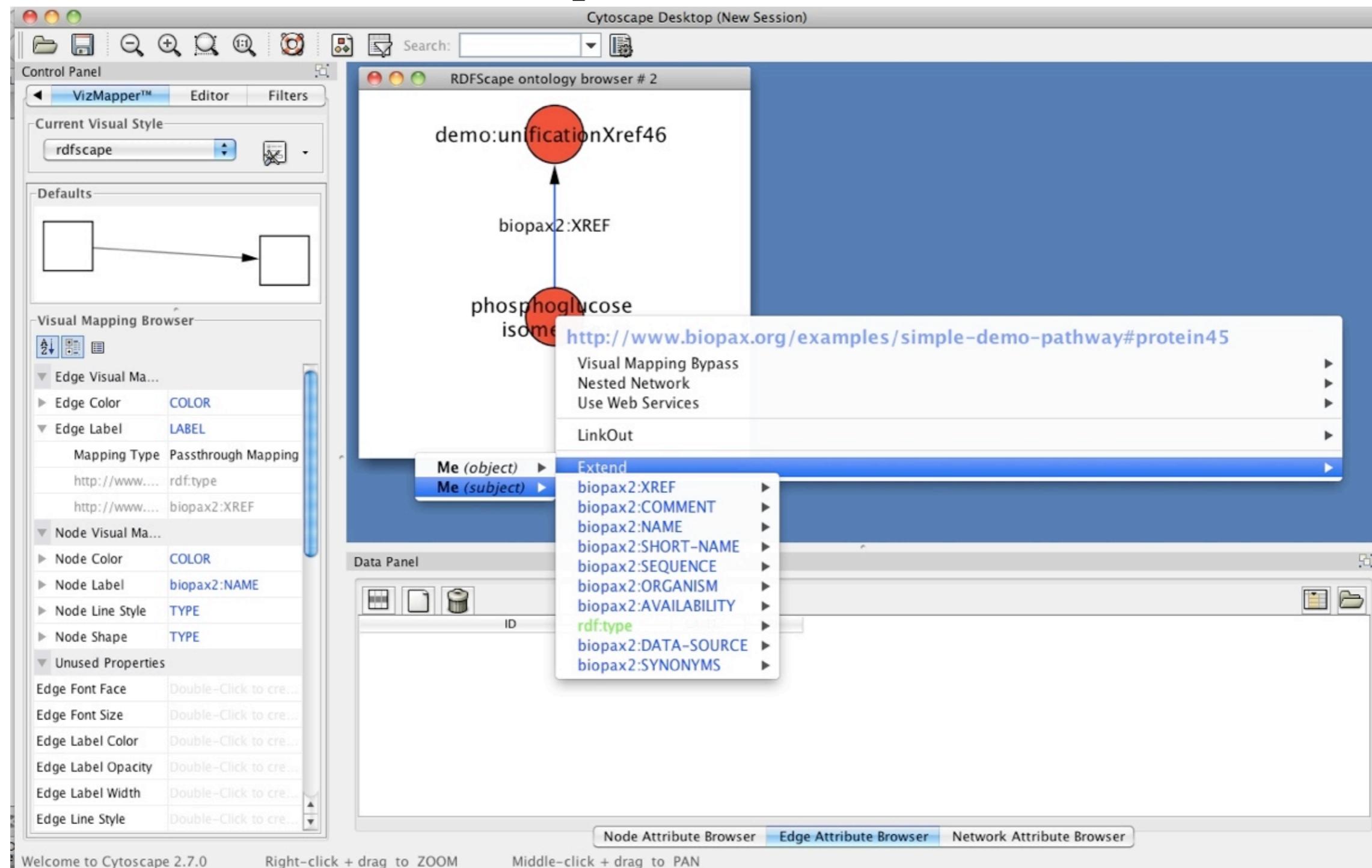
Inspection



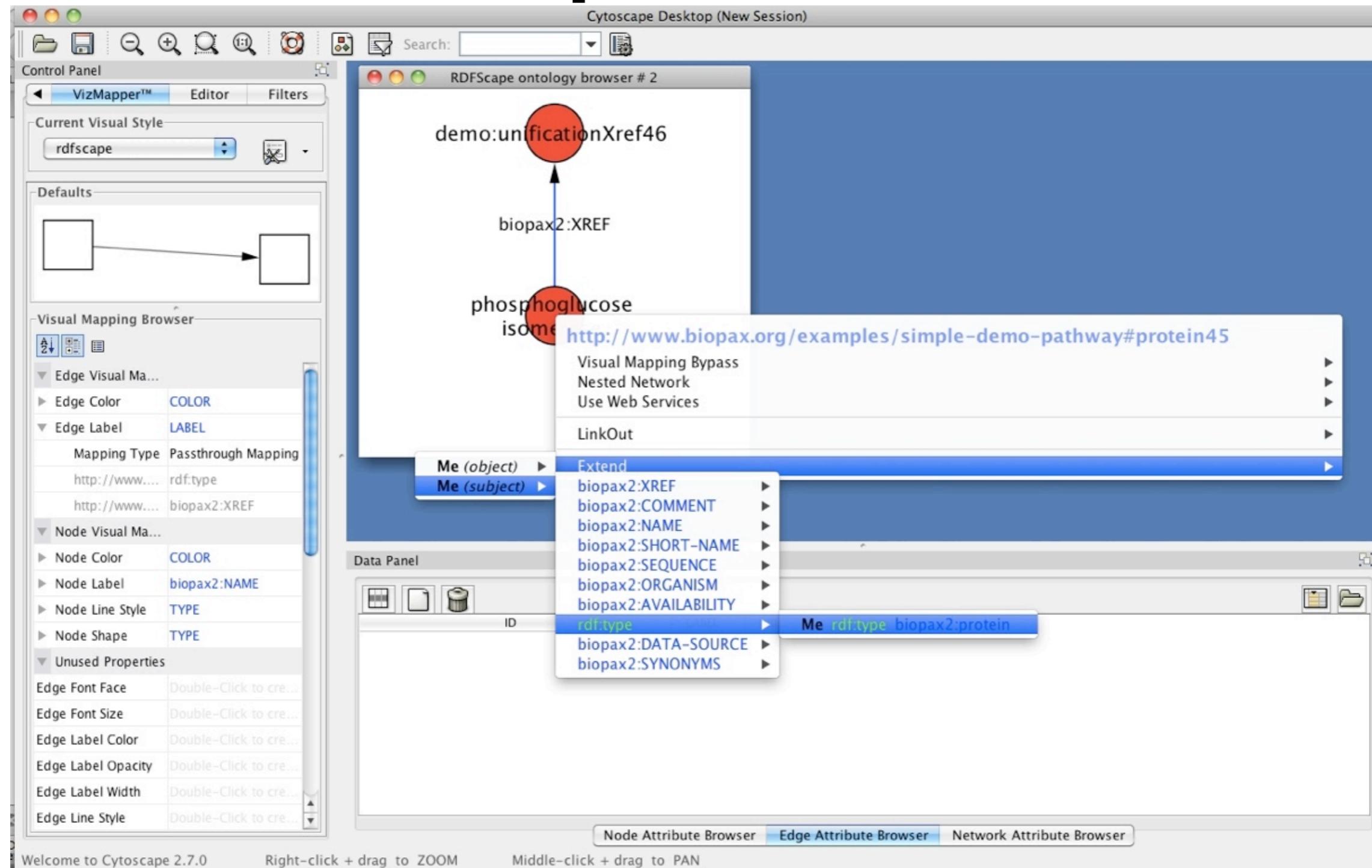
Inspection



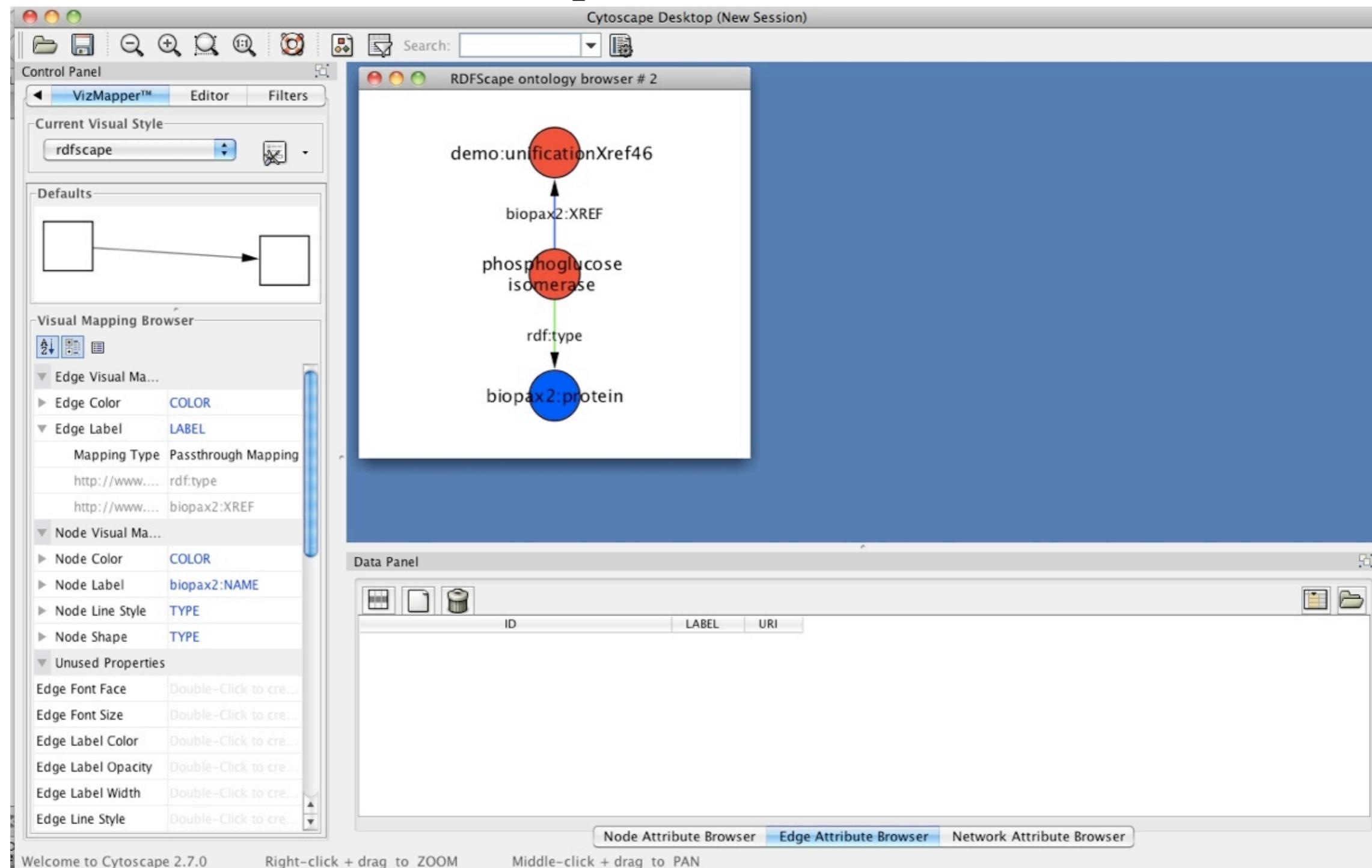
Inspection



Inspection

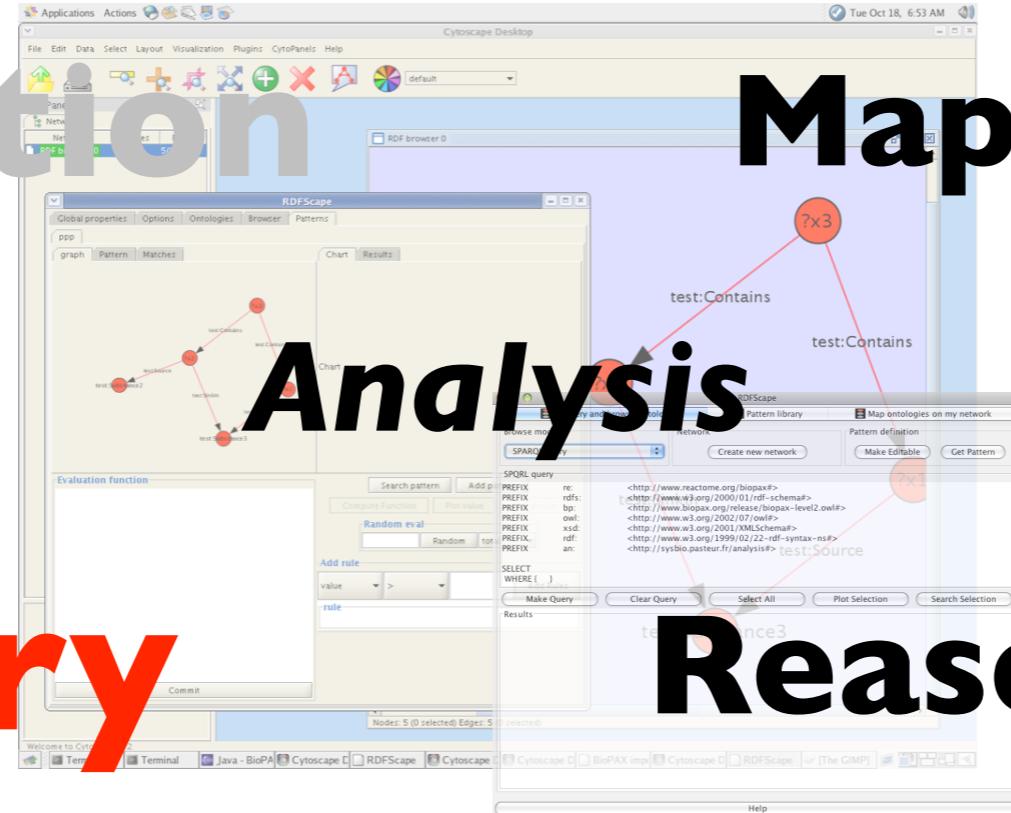


Inspection

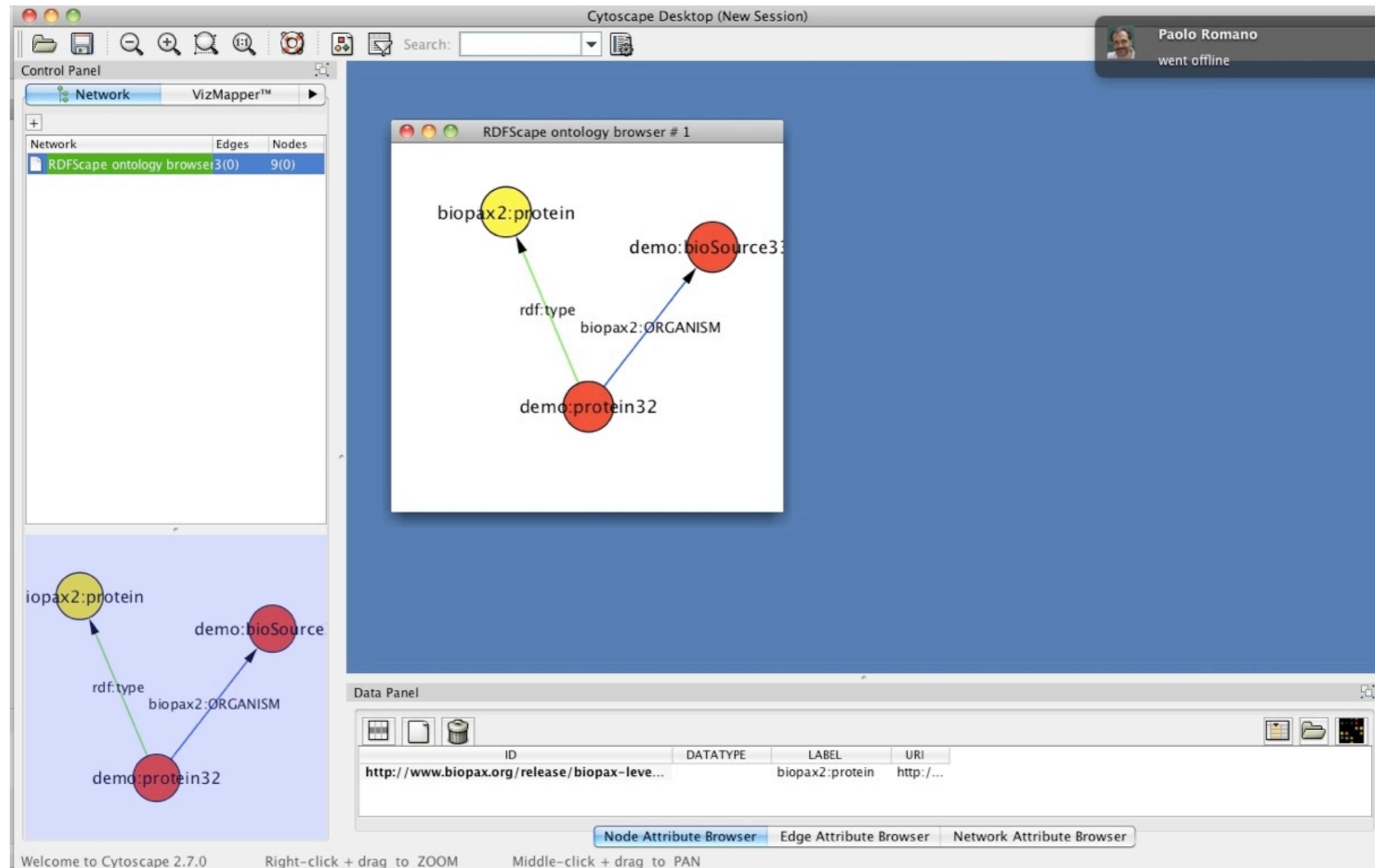


RDFScape = Interface

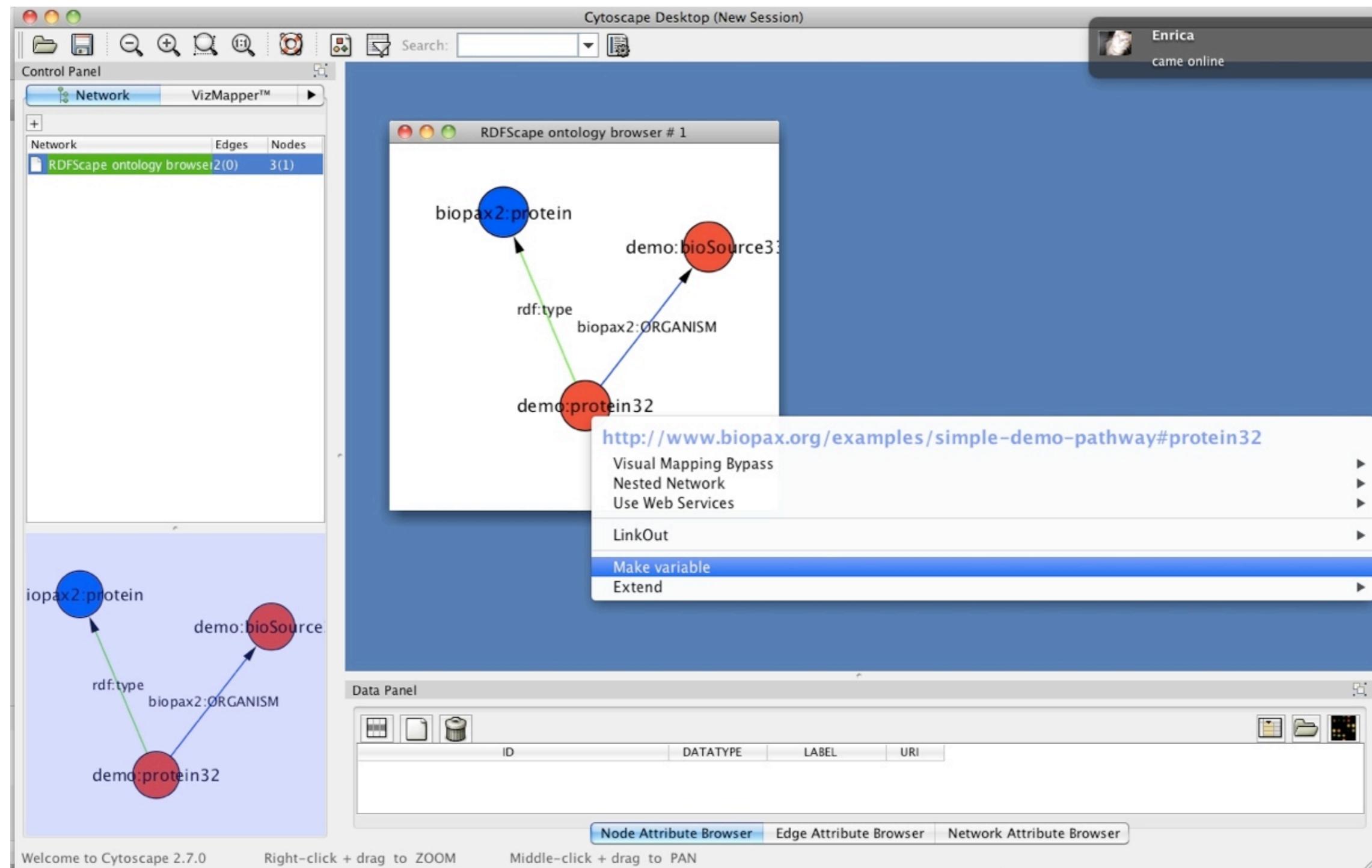
Inspection Mapping
Query Analysis Reasoning



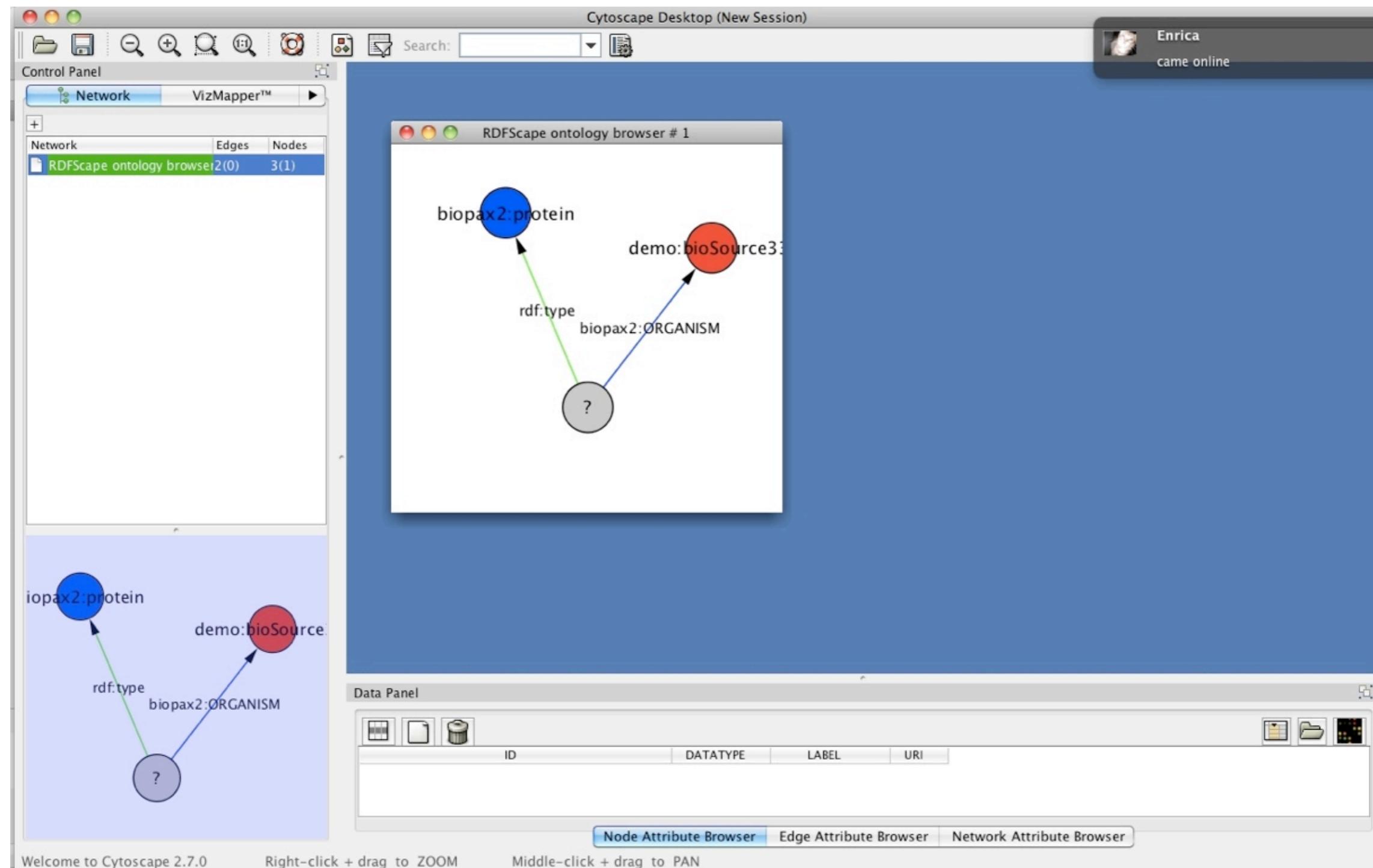
Query



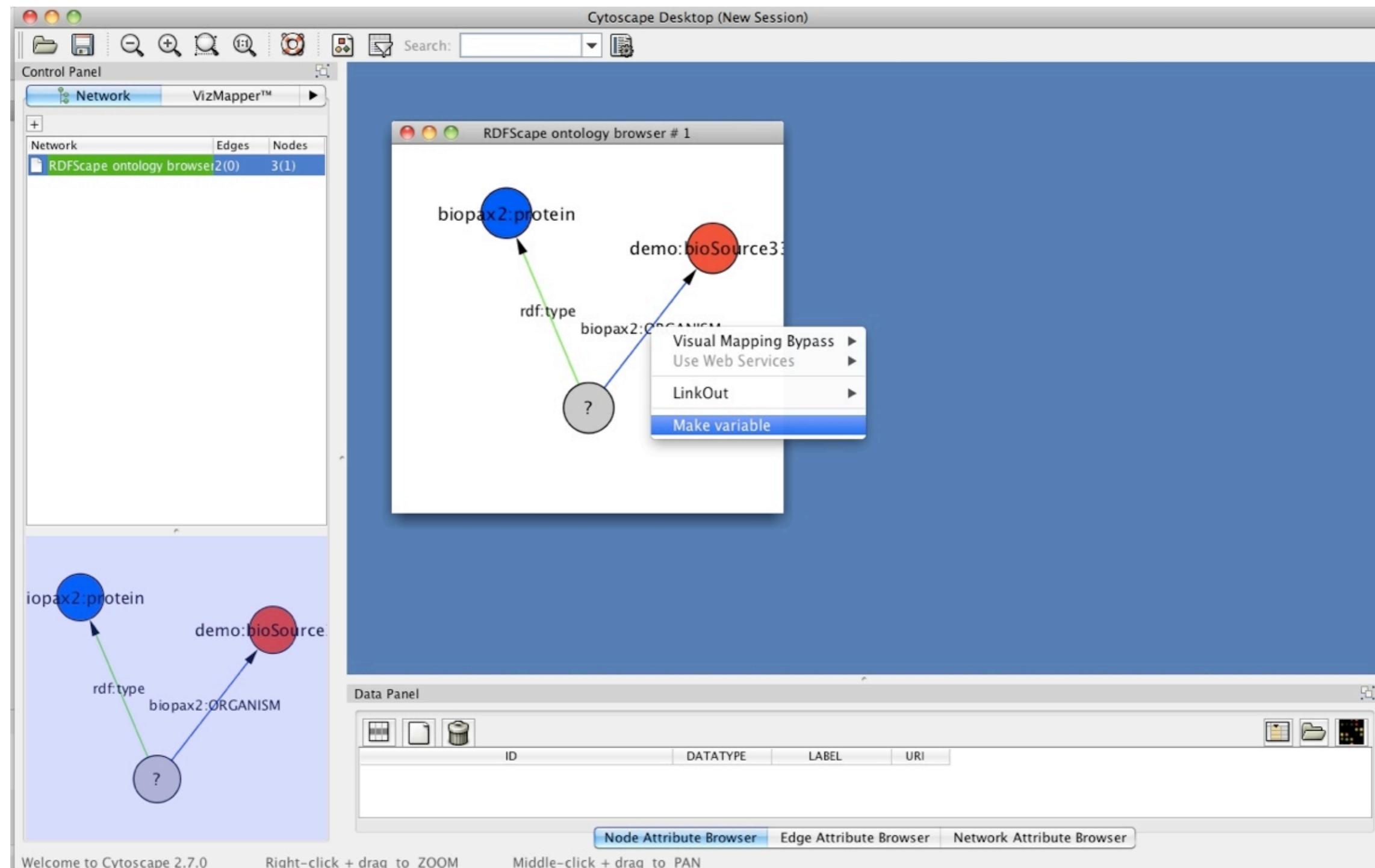
Query



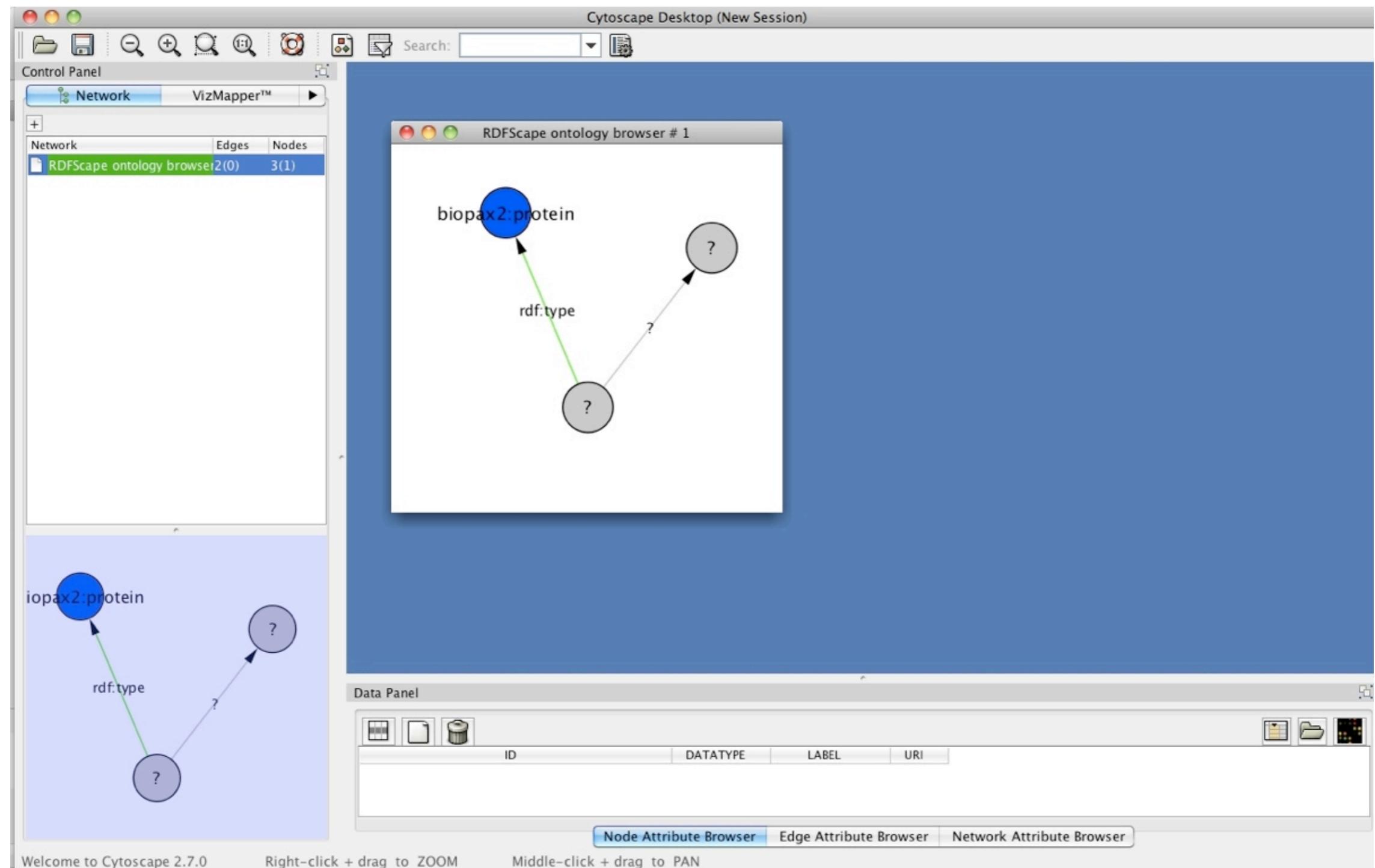
Query



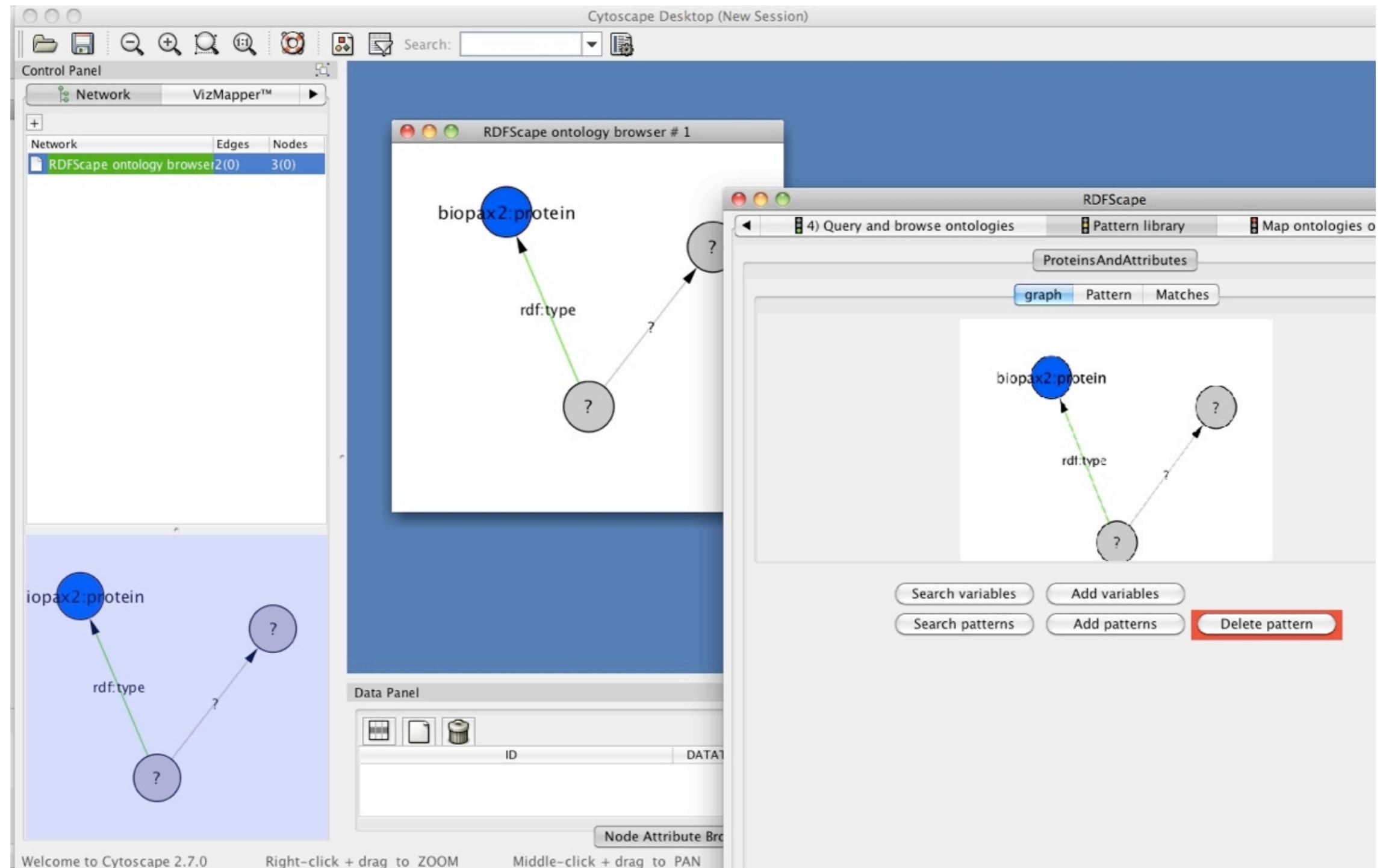
Query



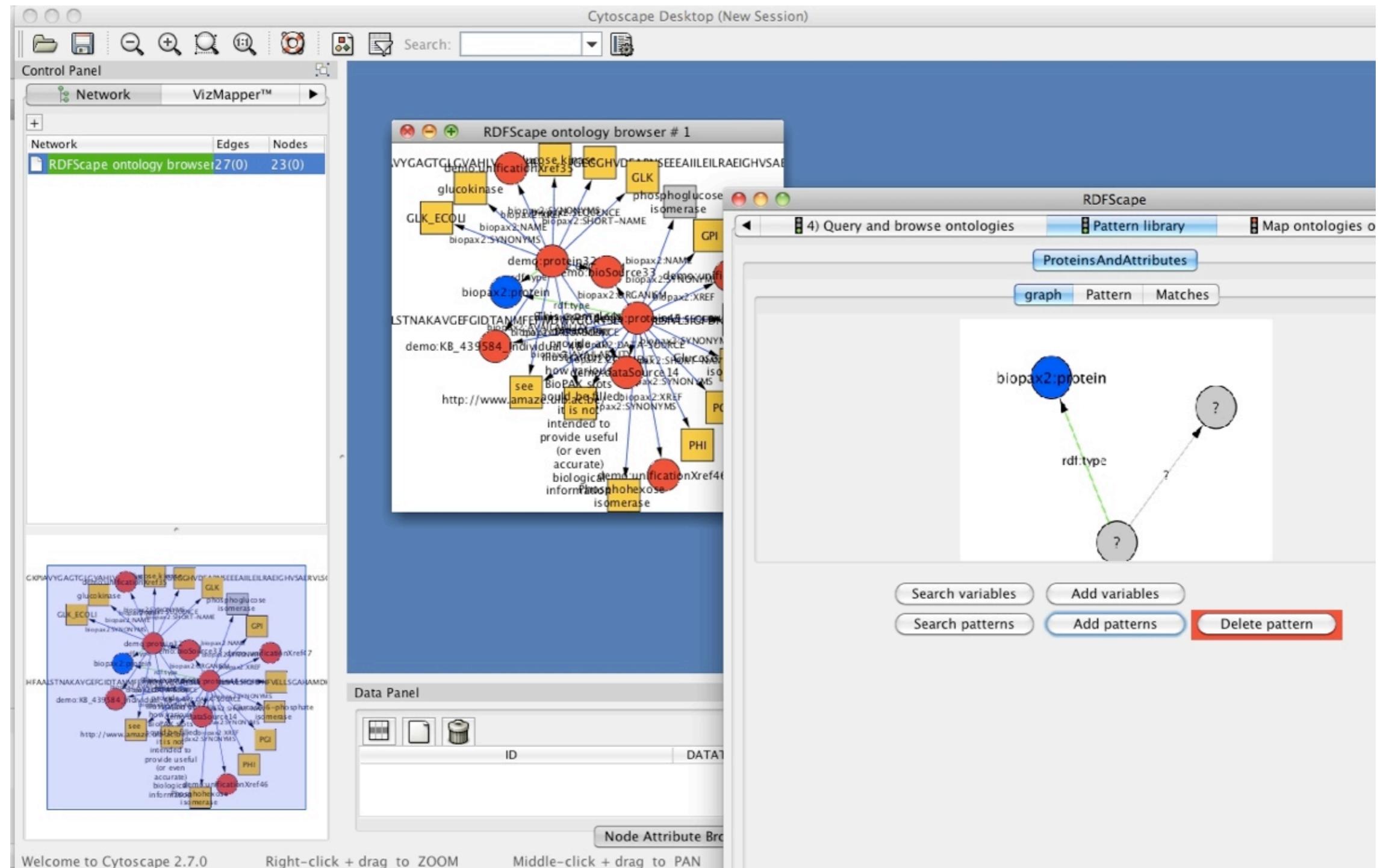
Query



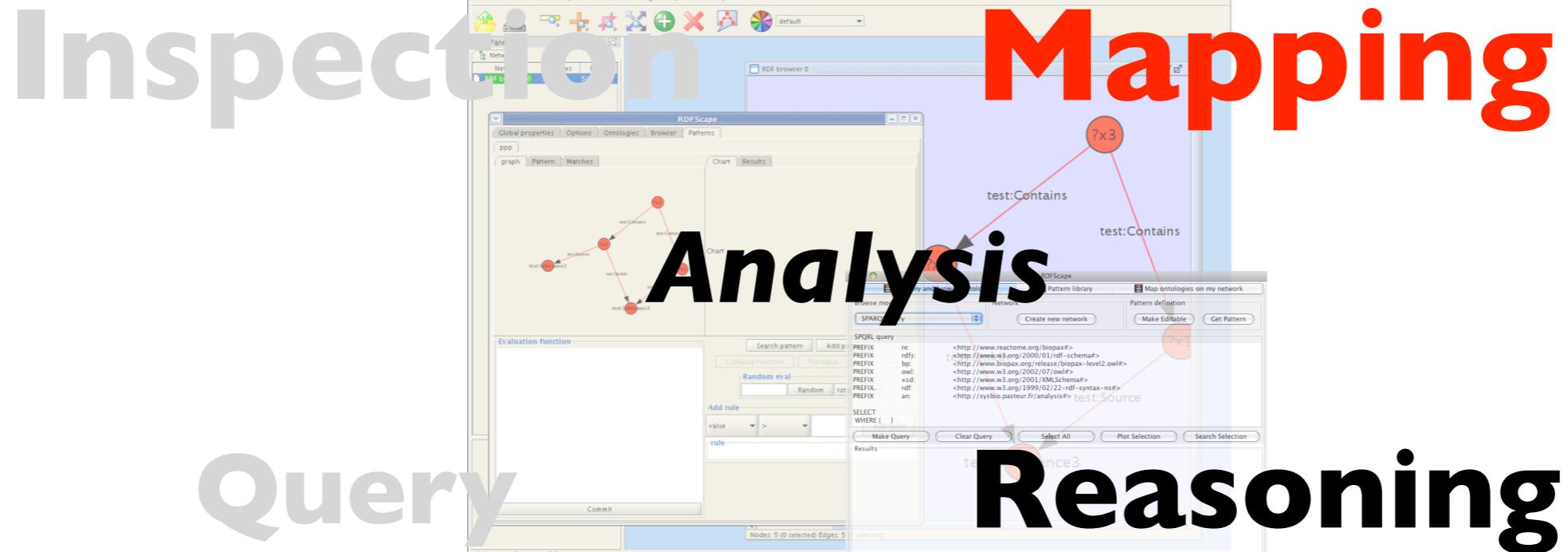
Query



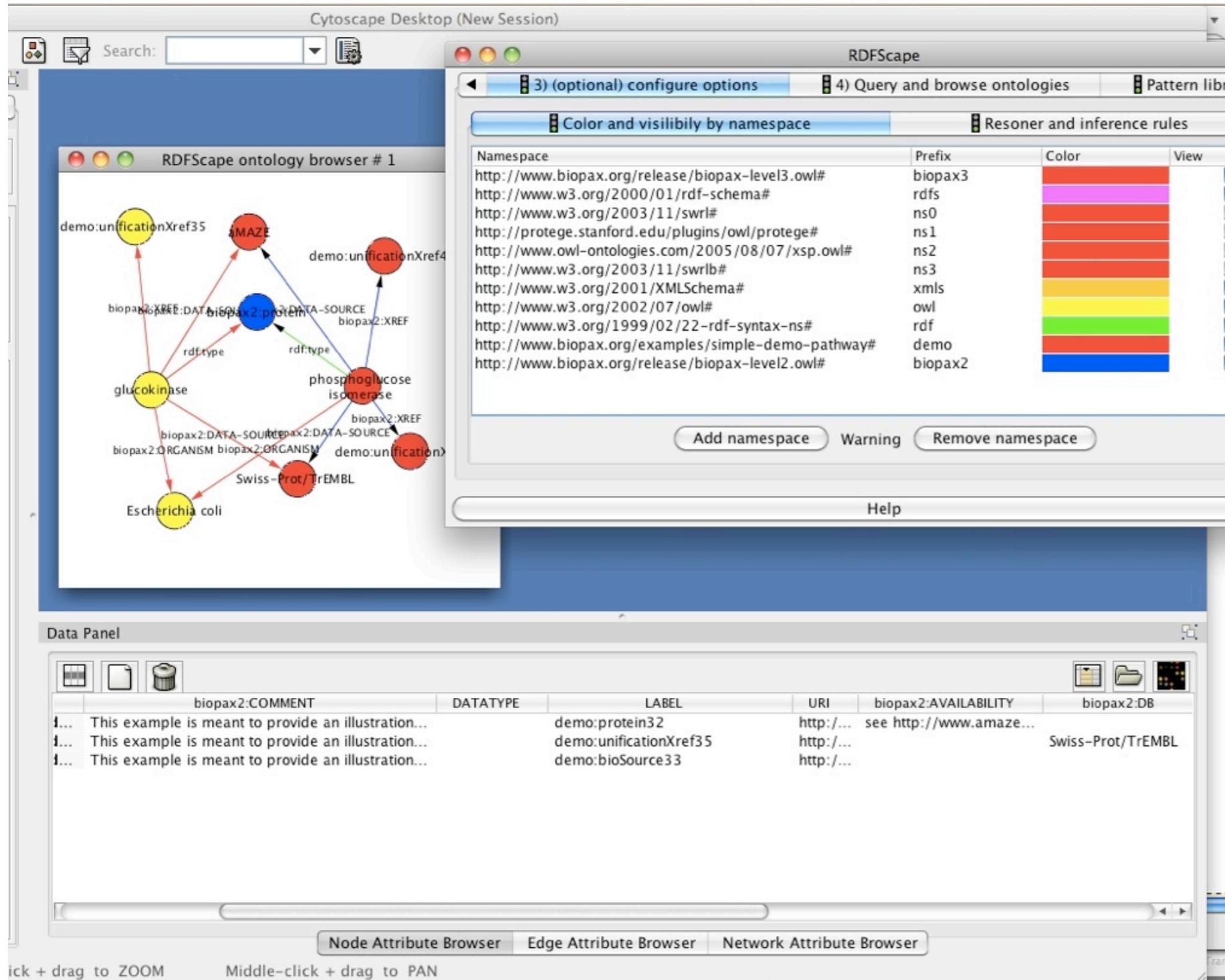
Query



RDFscape = Interface

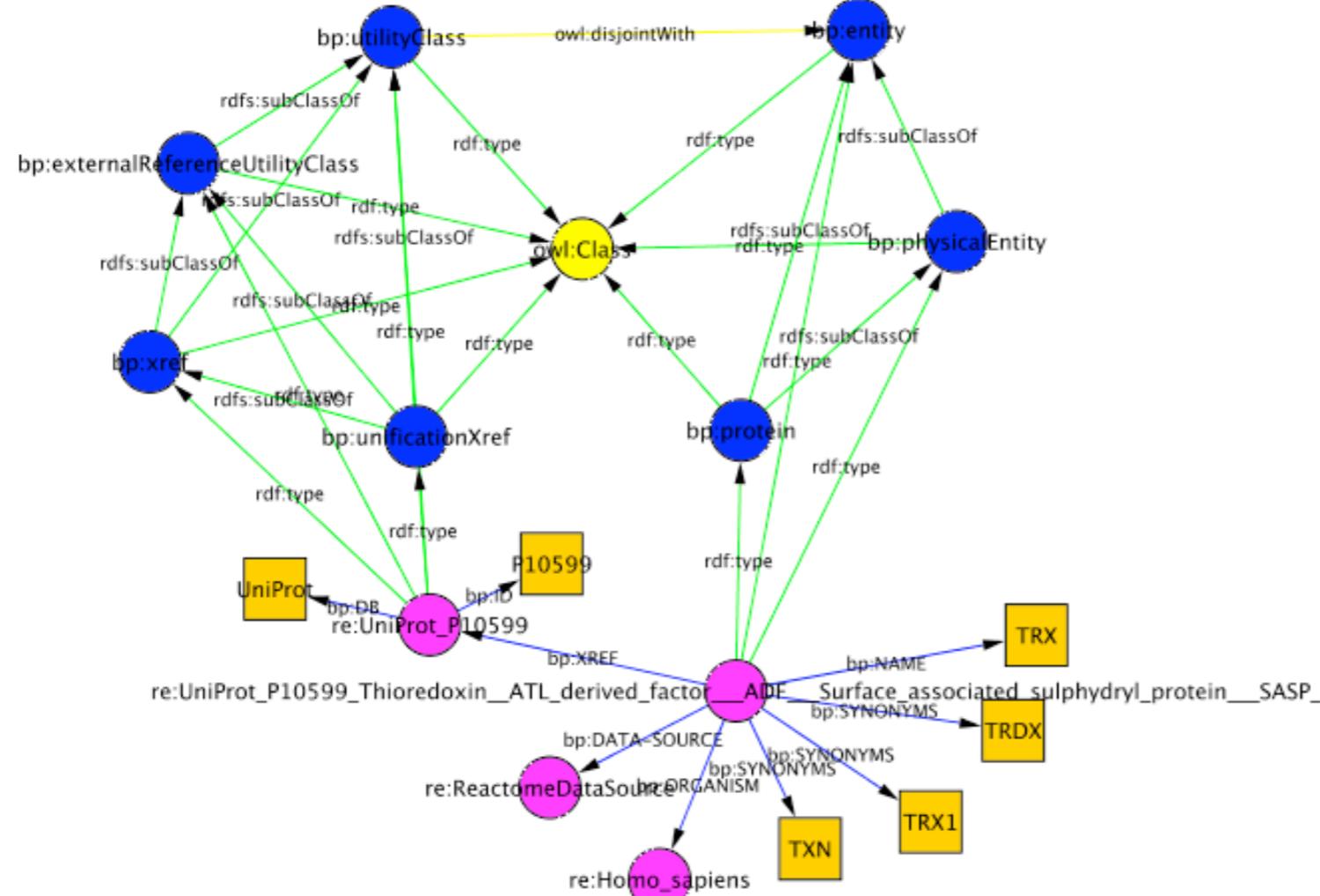


Mapping



- selection/colouring by namespace
- datatype properties as attributes
- URLs represented as multiple nodes

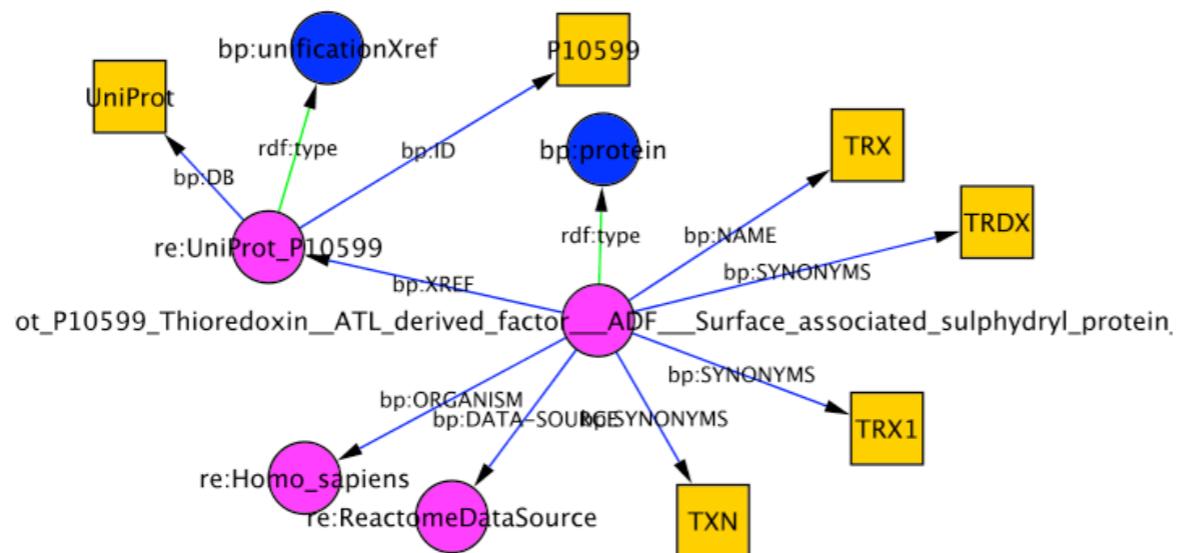
Mapping



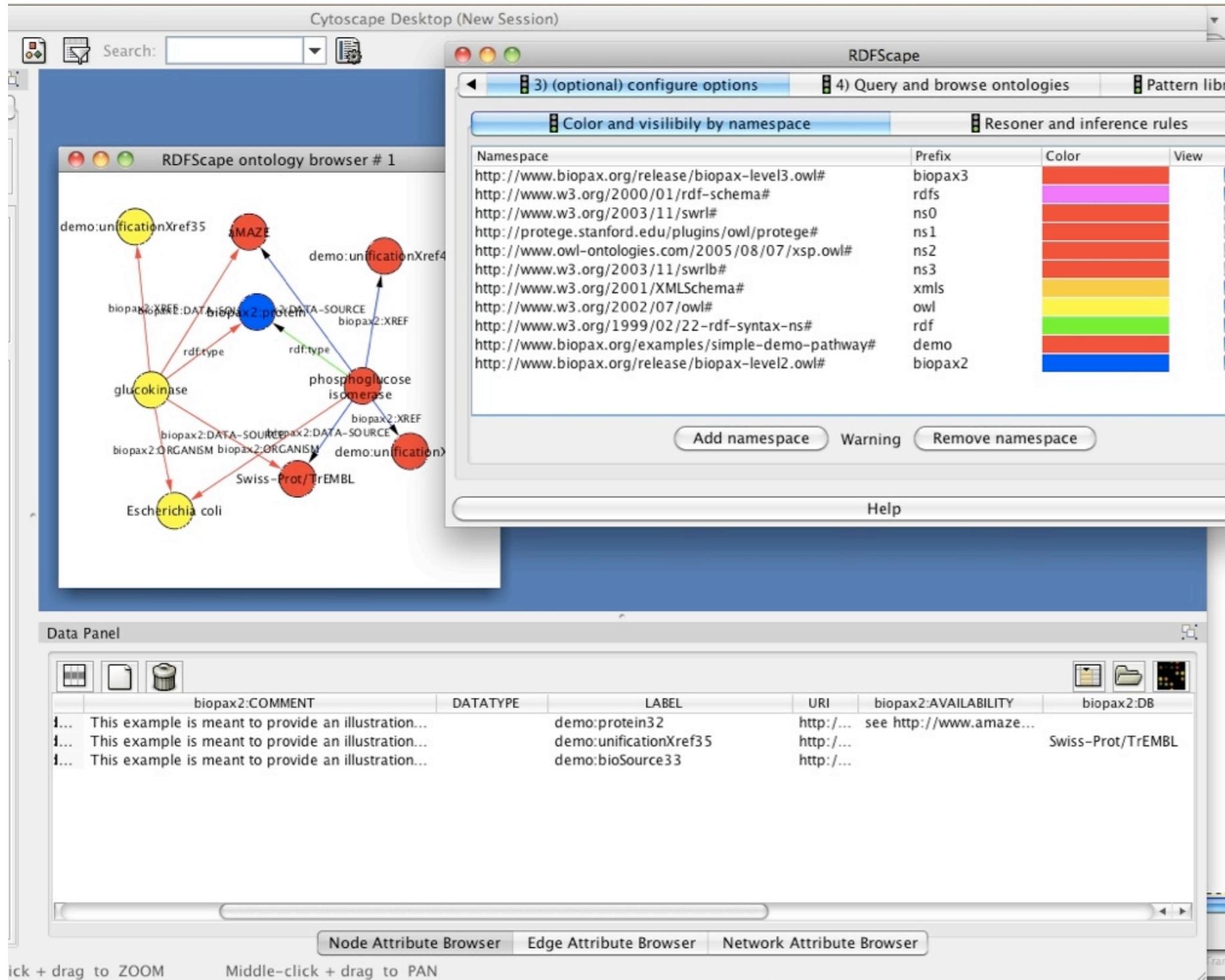
- selection/colouring by namespace
- datatype properties as attributes
- URLs represented as multiple nodes

Mapping

- selection/colouring by namespace
- datatype properties as attributes
- URLs represented as multiple nodes

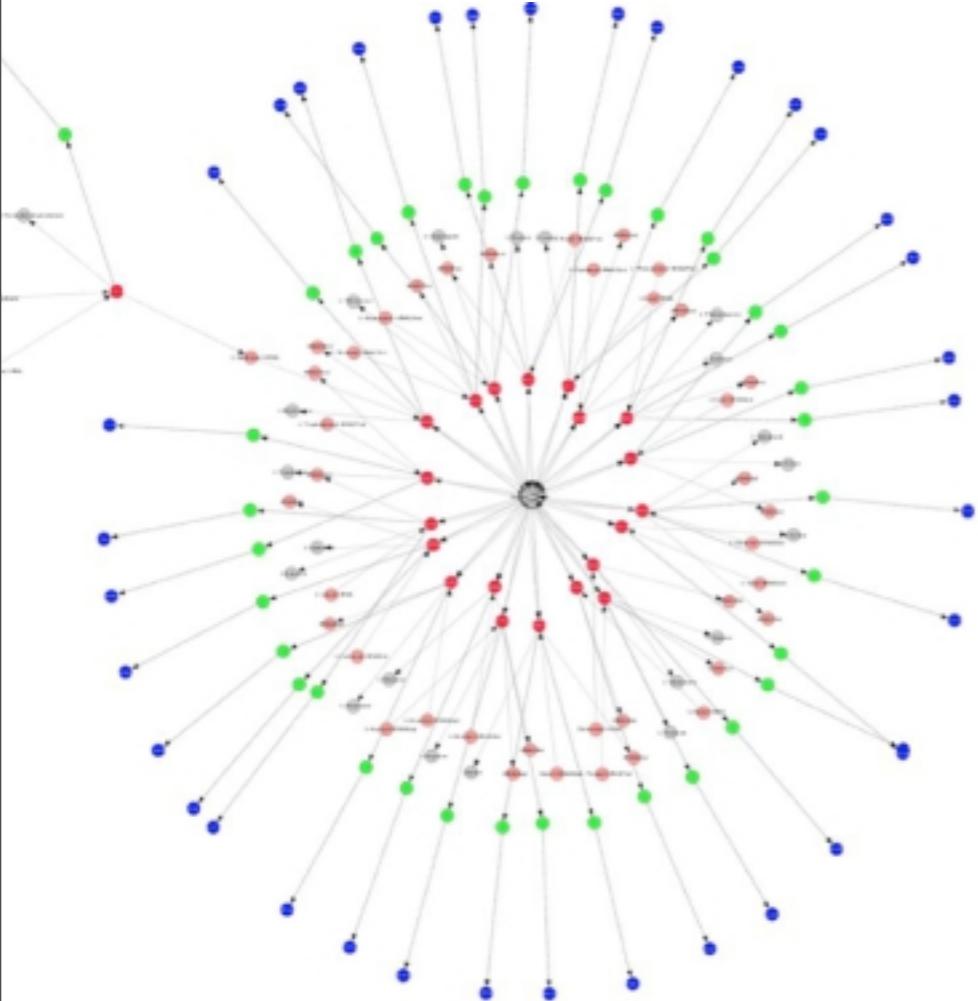


Mapping

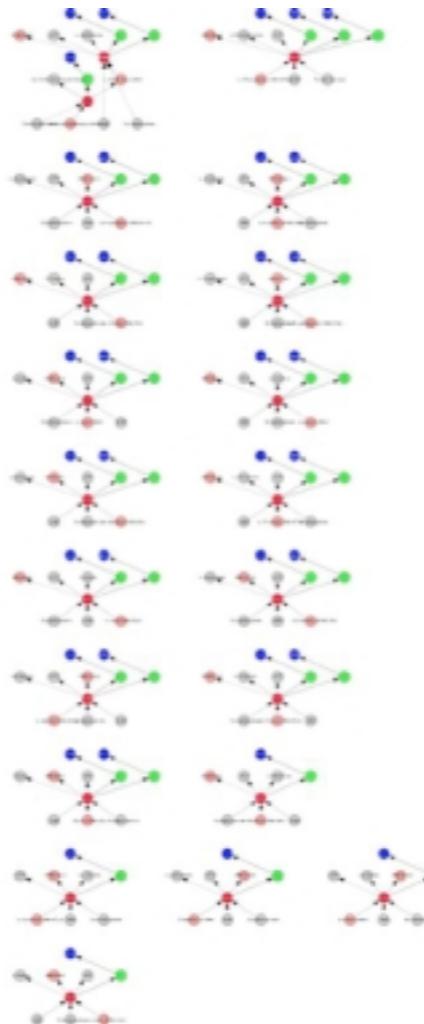


- selection/colouring by namespace
- datatype properties as attributes
- URLs represented as multiple nodes

Mapping



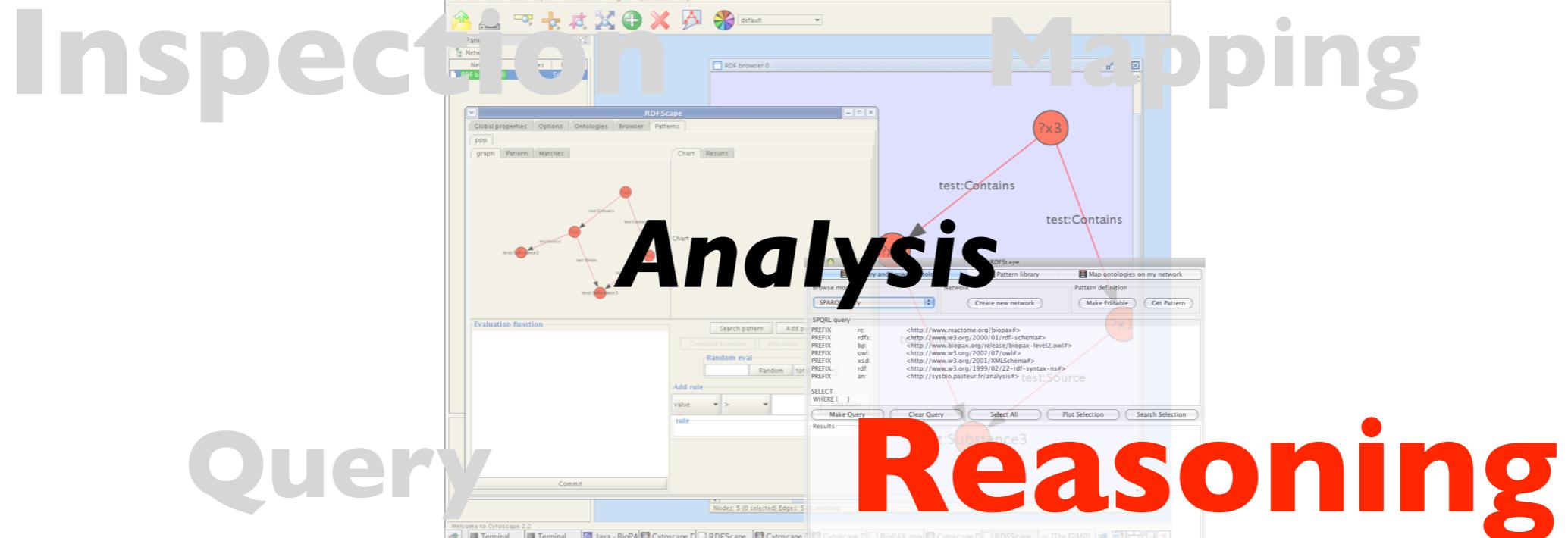
=



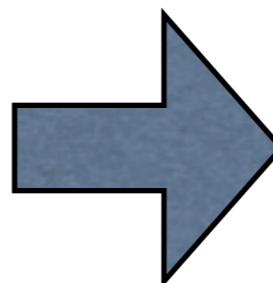
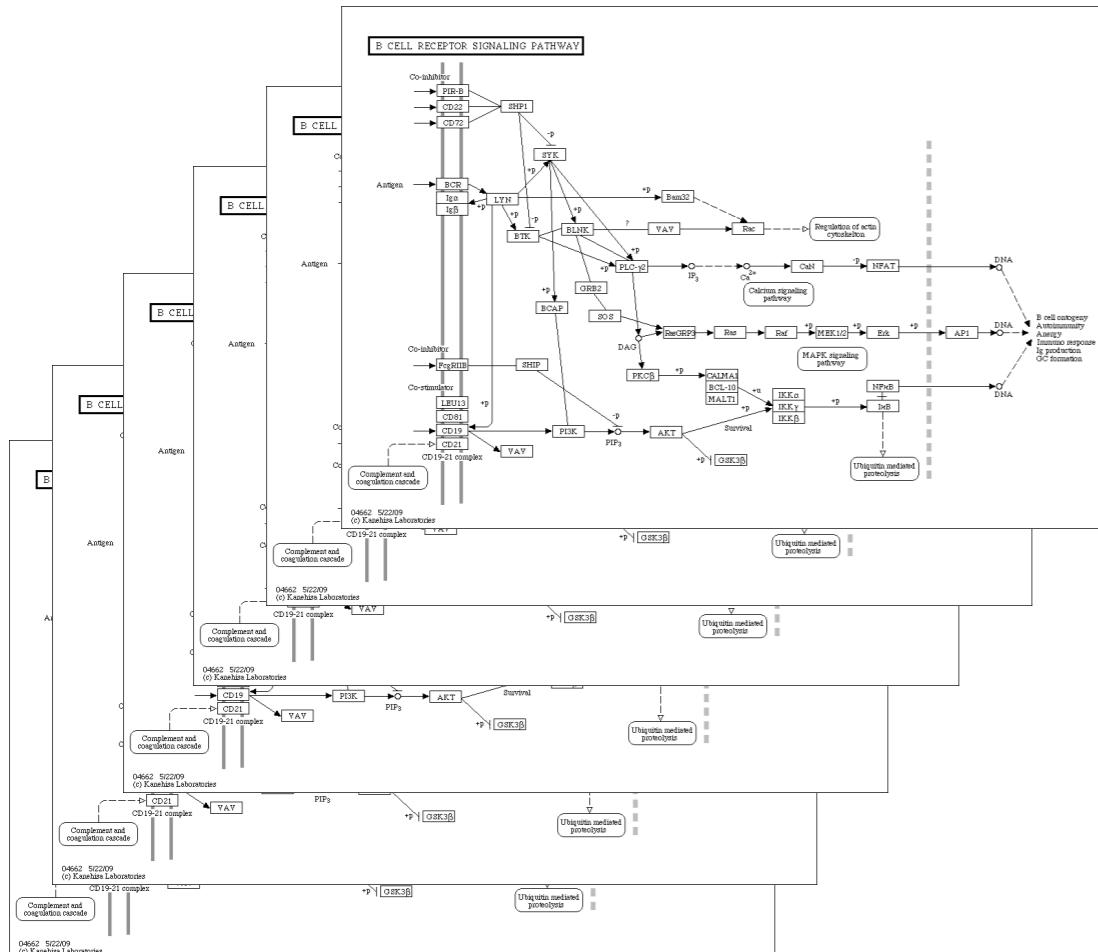
ATP, H₂O...

- selection/colouring by namespace
- datatype properties as attributes
- URIs represented as multiple nodes

RDFscape = Interface



Reasoning

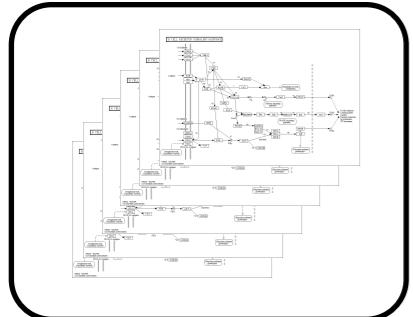


?x6

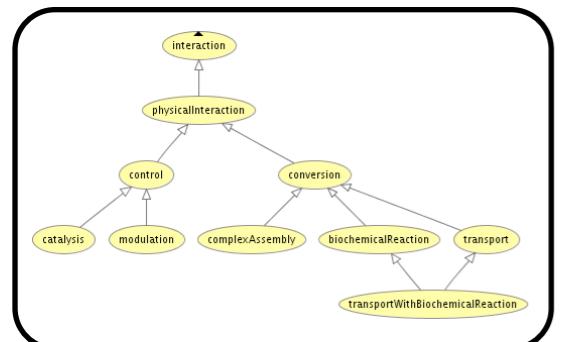
new.interacts

?x7

Reasoning



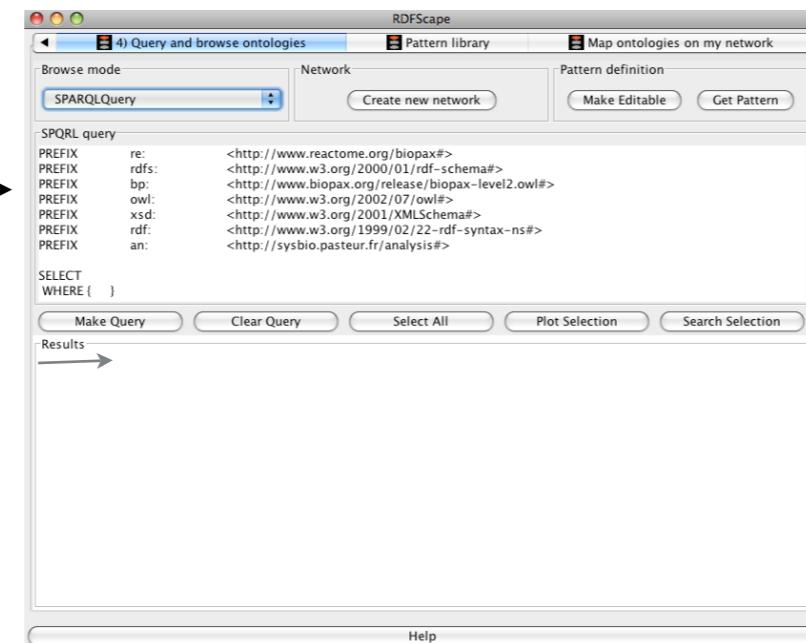
A collection of pathways:
biological facts



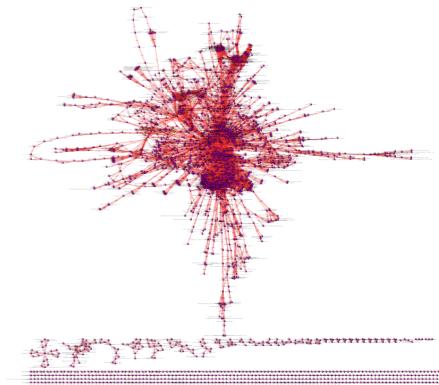
An ontology: a meaning
for facts (interpretation)

```
[Skip-Context1:  
(?i an:directctp ?p) <- (?i bp:PARTICIPANT ?p)  
(?p rdf:type bp:entity)]  
  
[Skip-Context2:  
(?i an:directctp ?p) <- (?i bp:PARTICIPANTS ?pep)  
(?pep bp:PHYSICAL ENTITY ?p)  
(?p rdf:type bp:entity)]  
  
[Interacts:  
(?x an:interacts ?y) <- (?z rdf:type bp:interaction)  
(?z and:directctp ?x) (?z an:directctp ?y)]
```

inference rules:
more meaning...



“Semantic-enabled”
Interface (RDFscape)



Visualization



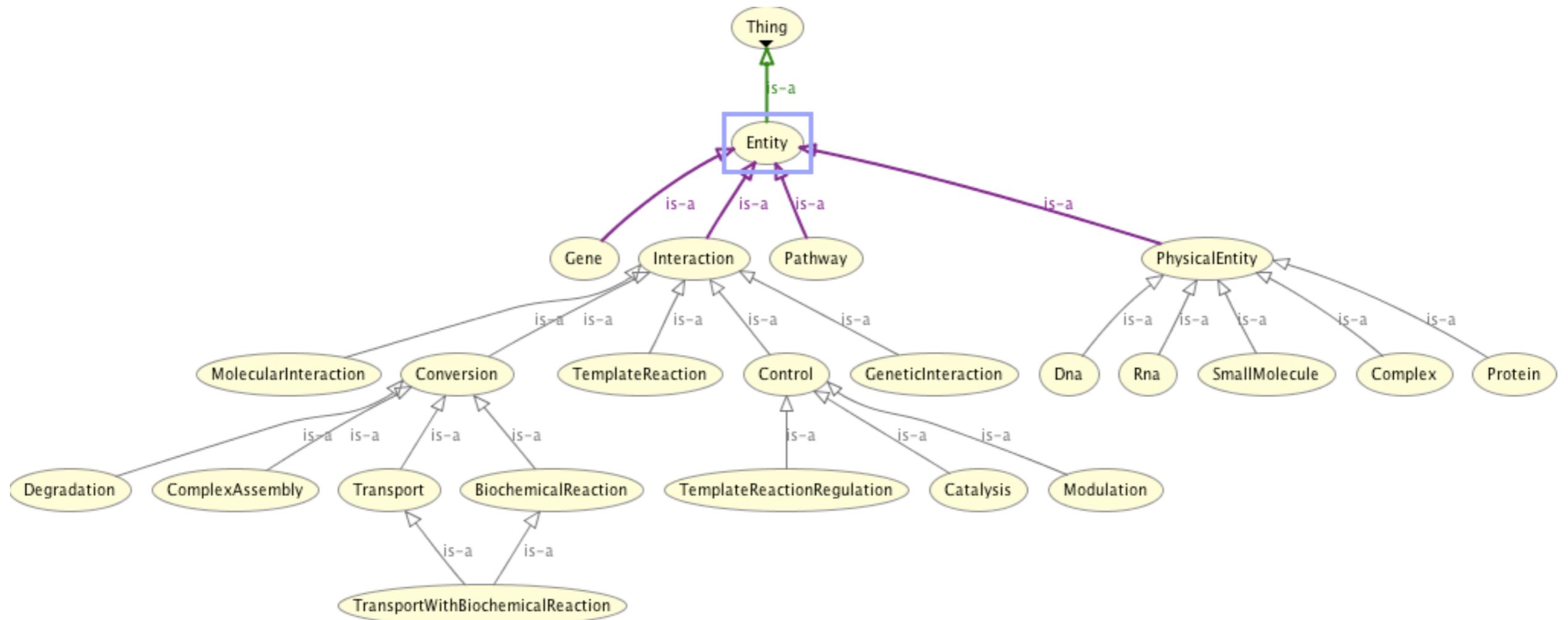
Analysis

Abstraction of pathways
as interactions

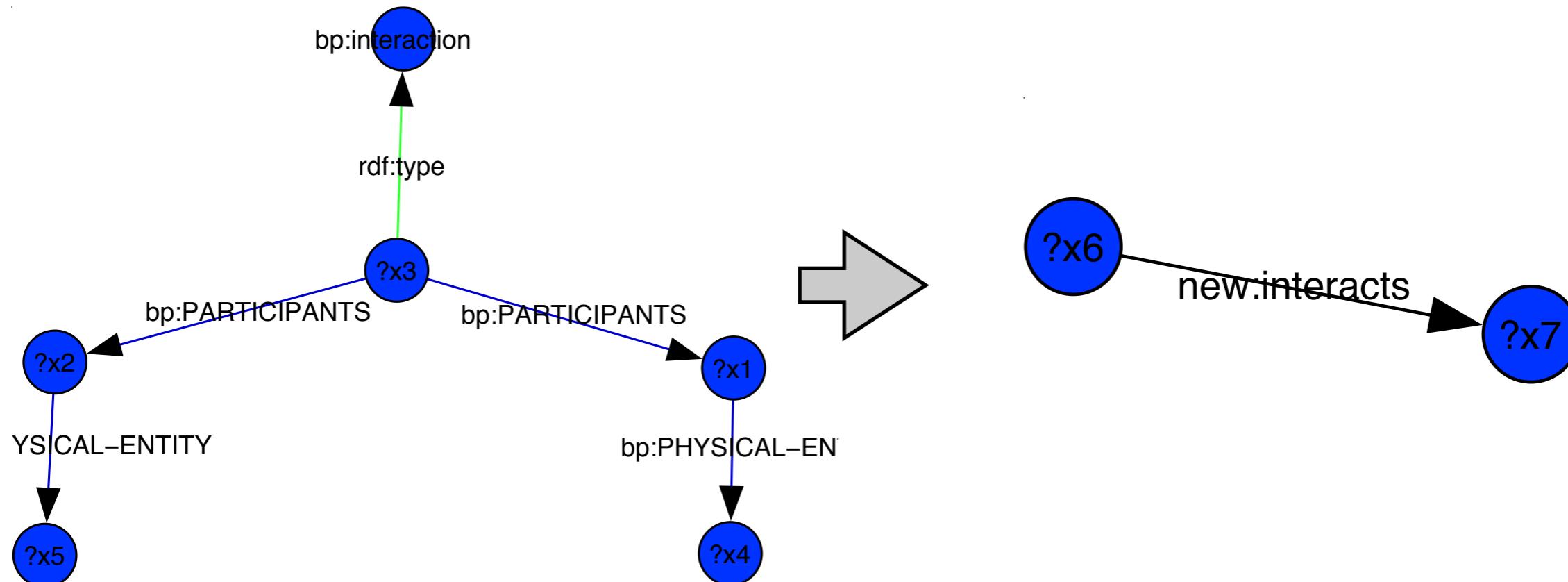


Queries

Reasoning



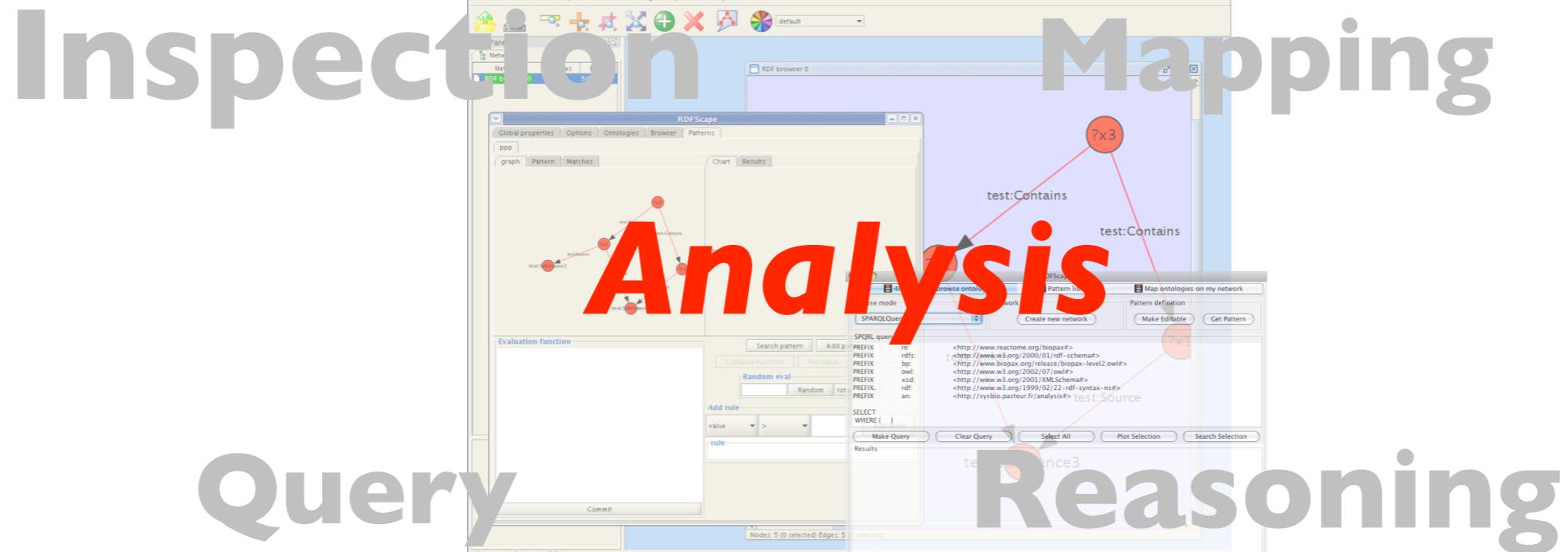
Reasoning



[Direct-interaction:
(?interactor new:dirInt ?p)
<-
(?x rdf:type bp:interaction)
(?x bp:PARTICIPANTS ?px)
(?px bp:PHYSICAL-ENTITY ?p)]

[Define-interaction:
(?pe1 new:interacts ?pe2)
<-
(?x rdf:type bp:interaction)
(?x new:dirInt ?pe1)
(?x new:dirInt ?pe2)]

RDFScape = Interface

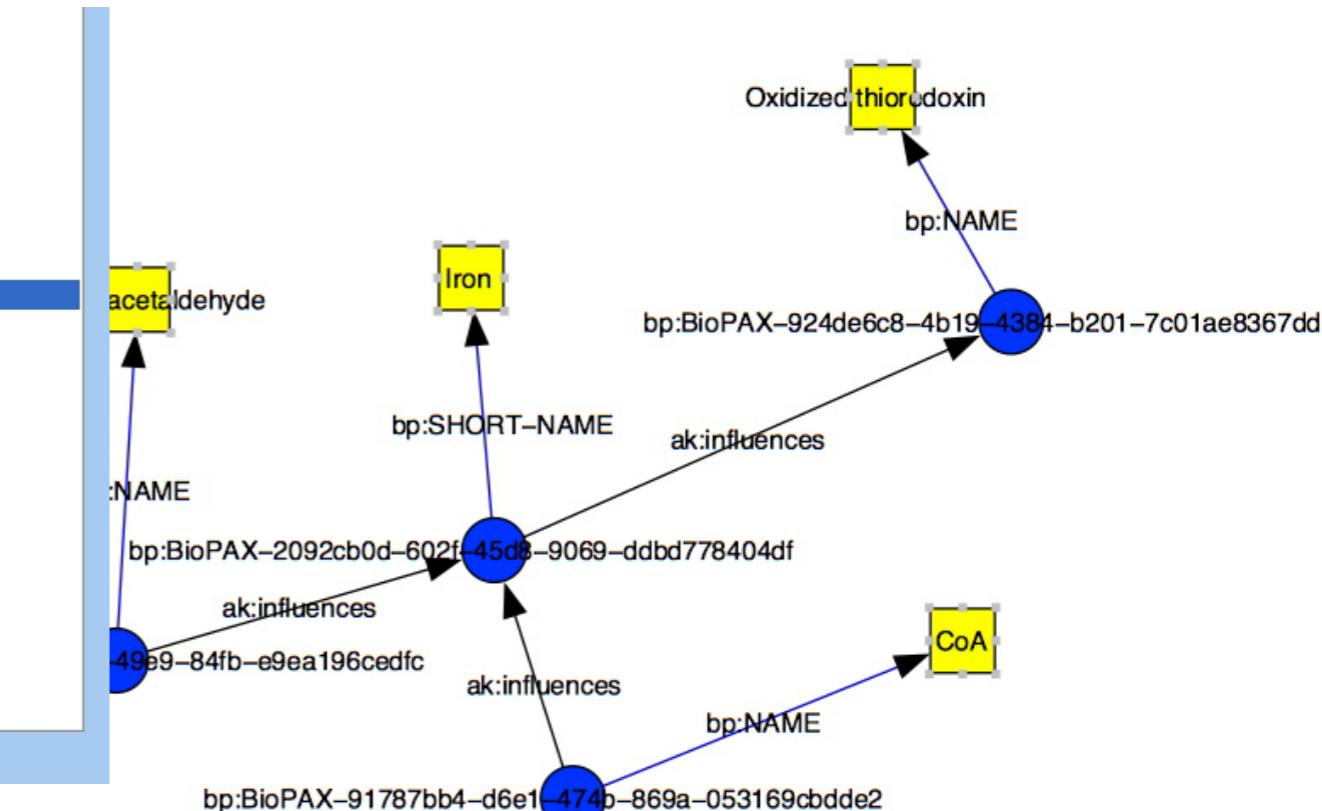


Analysis

(Reasoning on pathways)



```
[influence: (?x ak:influences ?y) <- (?p1 rdf:type
bp:biochemicalReaction)
(?p2 rdf:type bp:biochemicalReaction)
(?p1 bp:RIGHT ?k1)(?k1 bp:PHYSICAL-ENTITY ?k)
(?p2 bp:LEFT ?k2)(?k2 bp:PHYSICAL-ENTITY ?k)
(?p1 bp:LEFT ?k3)(?k3 bp:PHYSICAL-ENTITY ?x)
(?p2 bp:RIGHT ?k4)(?k4 bp:PHYSICAL-ENTITY ?y) ]
```



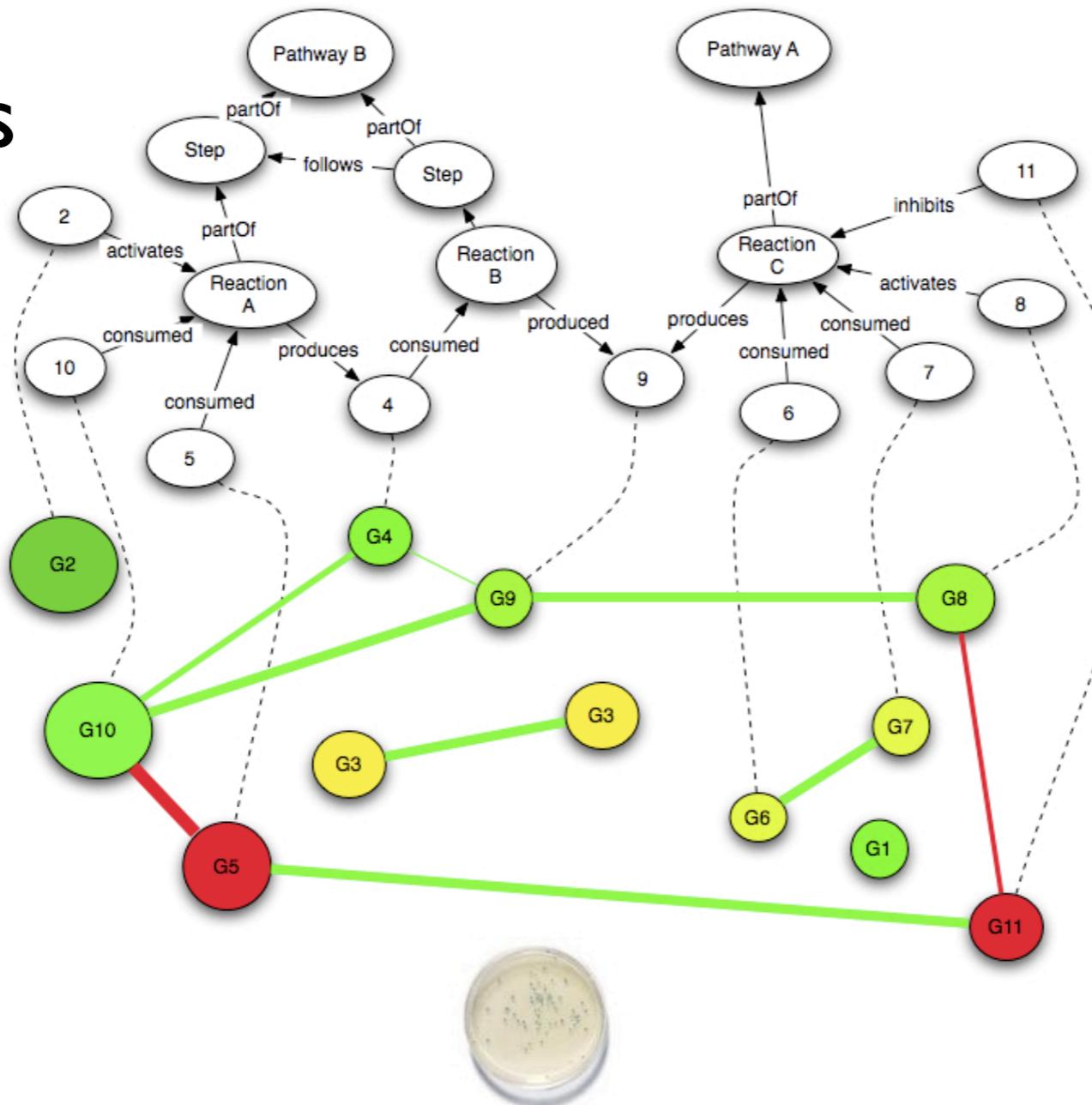
View of pathways as an influence graph

inference rules:
definition of “influence”

Analysis

(Ontologies and data)

- Ontologies
- General
- A-priori

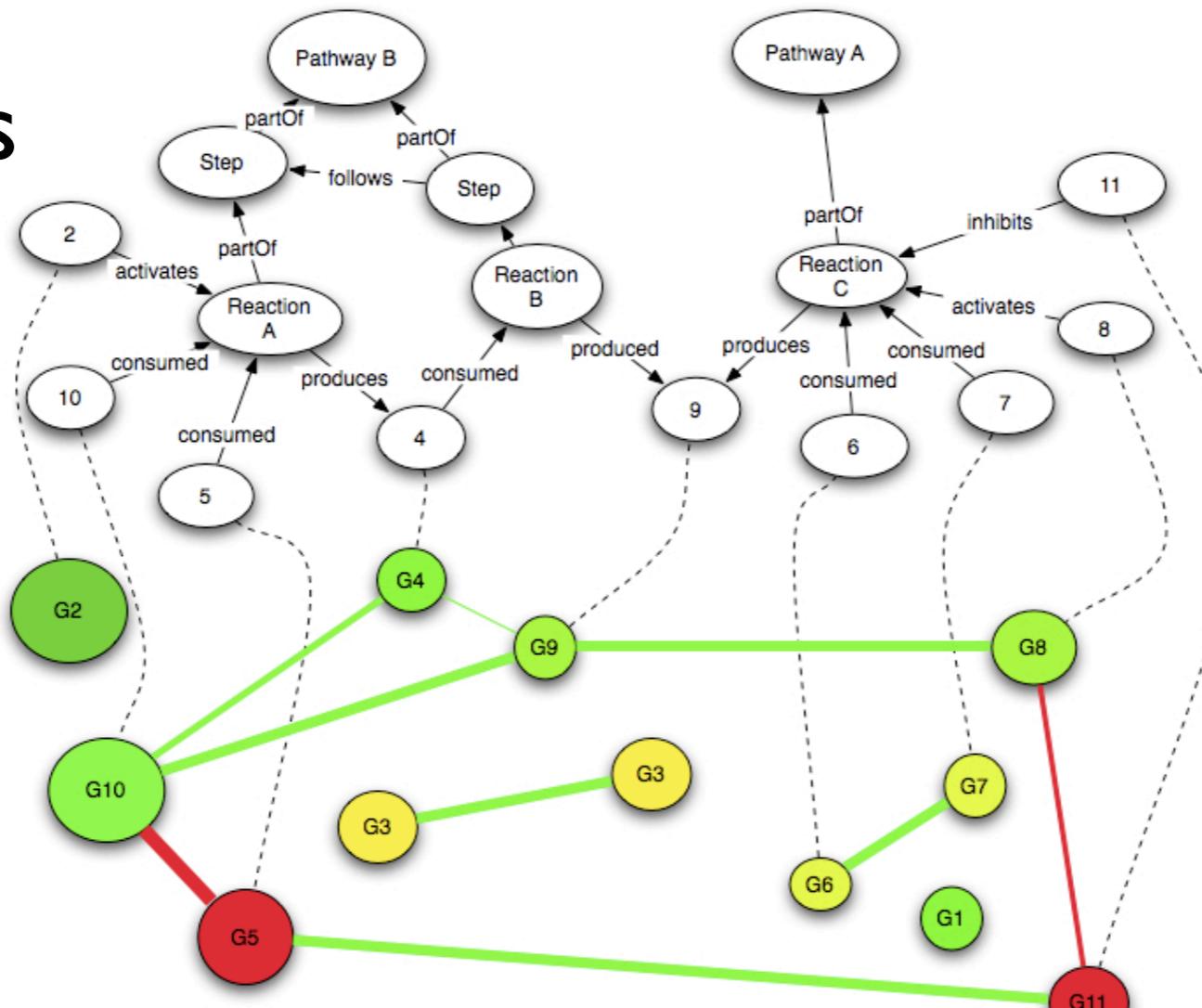


- Data
- Specific
- A-posteriori

Analysis

(Ontologies and data)

- Ontologies
- General
- A-priori

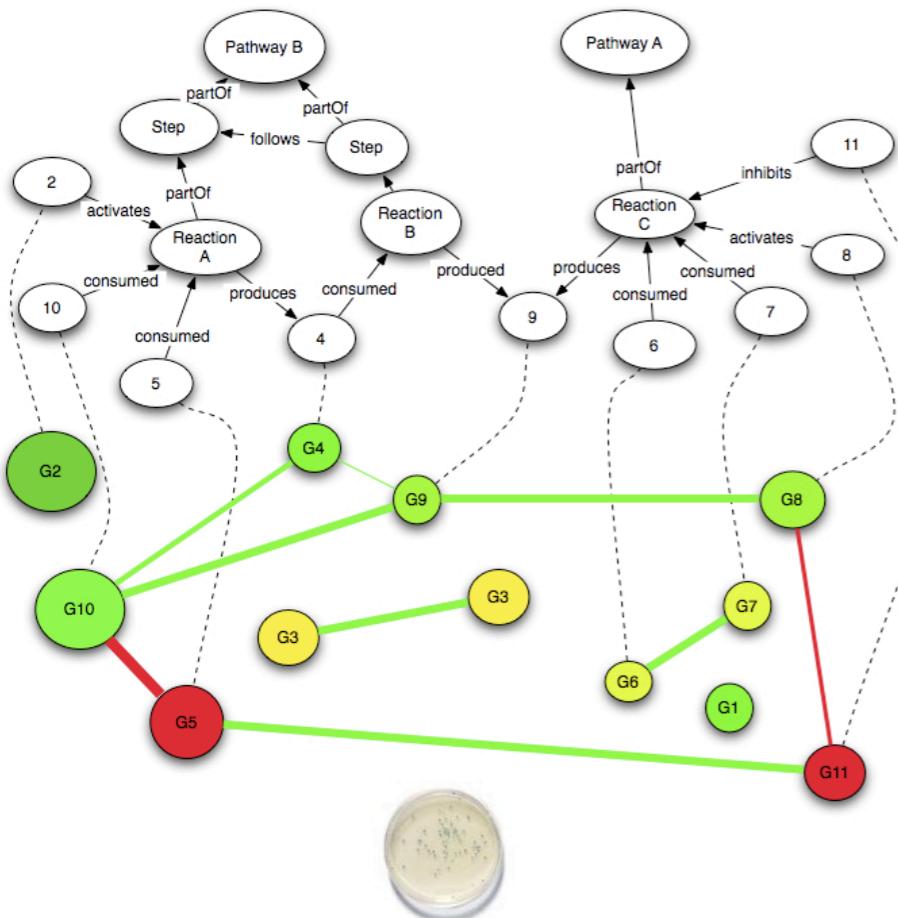


- Data
- Specific
- A-posteriori

The user open a correlation network in Cytoscape, and this is connected (through RDFScape) to ontologies

Analysis

(Ontologies and data)



Defining
the
mapping

The screenshot shows the Cytoscape interface with the RDFscape plugin. A rule named "sgd" is defined with the following pattern:

```
(?x <bp:XREF> ?y) (?y <bp:DB> ?z) (?y <bp:ID> ?w) AND ?z eq <SGD>
```

Details of the rule:

- Name: sgd
- Pattern: ($?x <\text{bp:XREF}> ?y$) ($?y <\text{bp:DB}> ?z$) ($?y <\text{bp:ID}> ?w$)
AND $?z \text{ eq } <\text{SGD}>$
- ID: $?w$
- URI: $?x$

Below the rule definition, there are buttons for New, Del, Resolve (which is highlighted in blue), and Merge. The status bar at the bottom indicates "Nodes: 331 (0 selected) Edges: 362 (0 selected)".

Analysis

(Queries on ontologies and data)

RDFscape

4) Query and browse ontologies Pattern library Map ontologies on my network
1) Select an analysis type! 2) Load ontologies 3) (optional) configure options

Browse mode: SPARQLQuery Network: Create new network Pattern definition: Make Editable, Get Pattern

SPARQL query:

```
PREFIX def: <http://www.leafbioscience.com/ontologies/test1#>
PREFIX ns0: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ns1: <http://www.w3.org/2001/XMLSchema#>
PREFIX ns2: <http://www.w3.org/2002/07/owl#>
PREFIX ns3: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ns4: <http://www.daml.org/2001/03/daml+oil#>
PREFIX n4: <http://rdfscape/network4#>

SELECT ?x ?relation ?y ?diffX ?stdvY
WHERE { ?x def:follows ?y .
         ?x n4:diff ?diffX .
         ?y n4:stdv ?stdvY .
         FILTER( ?diffX ) .
         FILTER( ?stdvY > 0.2 ) }
```

Make Query Clear Query Select All Plot Selection Search Selection

Results:

x	relation	y	diffX	stdvY
def:Peak3		def:Peak2	true	0.25
def:Peak2		def:Peak1	true	0.29

Analysis

(Queries on ontologies and data)

All possible relations (in the ontology) between ?x and ?y (in the a network)

PREFIX def: <http://www.eearbioscience.com/ontologies/test1#>
PREFIX s0: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX s1: <http://www.w3.org/2001/XMLSchema#>
PREFIX s2: <http://www.w3.org/2002/07/owl#>
PREFIX s3: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX n4: <http://www.daml.org/2001/03/daml+oil#>
PREFIX r4: <http://rdfscape/network4#>

SELECT ?x ?relation ?y ?diffX ?stdvY
WHERE { ?x def:follows ?y .
?x n4:diff ?diffX .
?y n4:stdv ?stdvY .
FILTER(?diffX) .
FILTER (?stdvY > 0.2) }

The screenshot shows a software interface for querying ontologies and networks. At the top, there's a toolbar with buttons for 'configure options', 'initiation', 'editable', and 'Get Pattern'. Below the toolbar is a text area containing SPARQL code. A large blue arrow points from the text area down towards the results table. The results table has columns for 'x', 'relation', 'y', 'diffX', and 'stdvY'. There are two rows of data: one where 'x' is 'def:Peak3' and 'y' is 'def:Peak2', and another where 'x' is 'def:Peak2' and 'y' is 'def:Peak1'. Both rows have 'true' in the 'relation' column and values '0.25' and '0.29' respectively in the 'diffX' and 'stdvY' columns.

x	relation	y	diffX	stdvY
def:Peak3	true	def:Peak2	0.25	
def:Peak2	true	def:Peak1	0.29	

Analysis

(Queries on ontologies and data)

All possible relations (in the ontology) between ?x and ?y (in the a network)

PREFIX def: <http://www.eearbioscience.com/ontologies/test1#>
PREFIX s0: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX s1: <http://www.w3.org/2001/XMLSchema#>
PREFIX s2: <http://www.w3.org/2002/07/owl#>
PREFIX s3: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX s4: <http://www.daml.org/2001/03/xhtml#>
PREFIX s5: <http://rdfscape/network#>

```
SELECT ?x ?relation ?y ?diffX ?stdvY  
WHERE { ?x def:follows ?y .  
        ?x n4:diff ?diffX .  
        ?y n4:stdv ?stdvY .  
        FILTER( ?diffX ) .  
        FILTER( ?stdvY > 0.2 ) }
```

Such that ?y follows ?x
(inferred)

The screenshot shows a SPARQL query editor interface. At the top, there's a configuration window with tabs for 'configurable options' and 'Get Pattern'. Below it is a text area containing the SPARQL query. At the bottom, there are buttons for 'Make Query' (highlighted in yellow), 'Clear Query', 'Select All', 'Plot Selection', and 'Search Selection'. The 'Results' section displays a table with the following data:

x	relation	y	diffX	stdvY
def:Peak3		def:Peak2	true	0.25
def:Peak2		def:Peak1	true	0.29

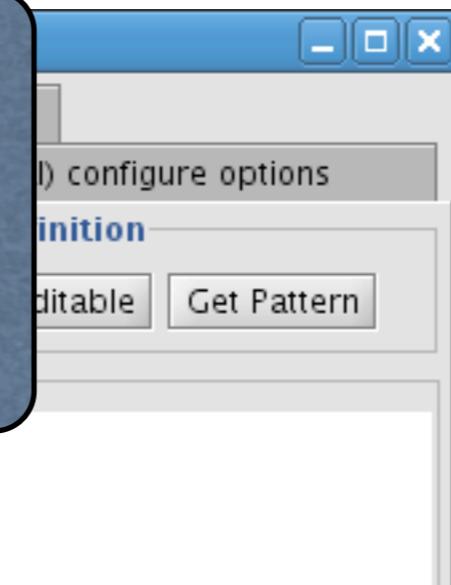
Analysis

(Queries on ontologies and data)

All possible relations (in the ontology) between ?x and ?y (in the a network)

PREFIX def: <<http://www.eearbioscience.com/ontologies/test1#>>
PREFIX s0: <<http://www.w3.org/2000/01/rdf-schema#>>
PREFIX s1: <<http://www.w3.org/2001/XMLSchema#>>
PREFIX s2: <<http://www.w3.org/2002/07/owl#>>
PREFIX s3: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>
PREFIX n4: <<http://www.daml.org/2001/03/rdf-schema#>>
PREFIX n5: <<http://rdfescape.netwerk42.com/ns#>>

SELECT ?x ?relation ?y ?diffX ?stdvY
WHERE { ?x def:follows ?y .
 ?x n4:diff ?diffX .
 ?y n4:stdv ?stdvY .
 FILTER(?diffX > 0) .
 FILTER(?stdvY > 0.2) }

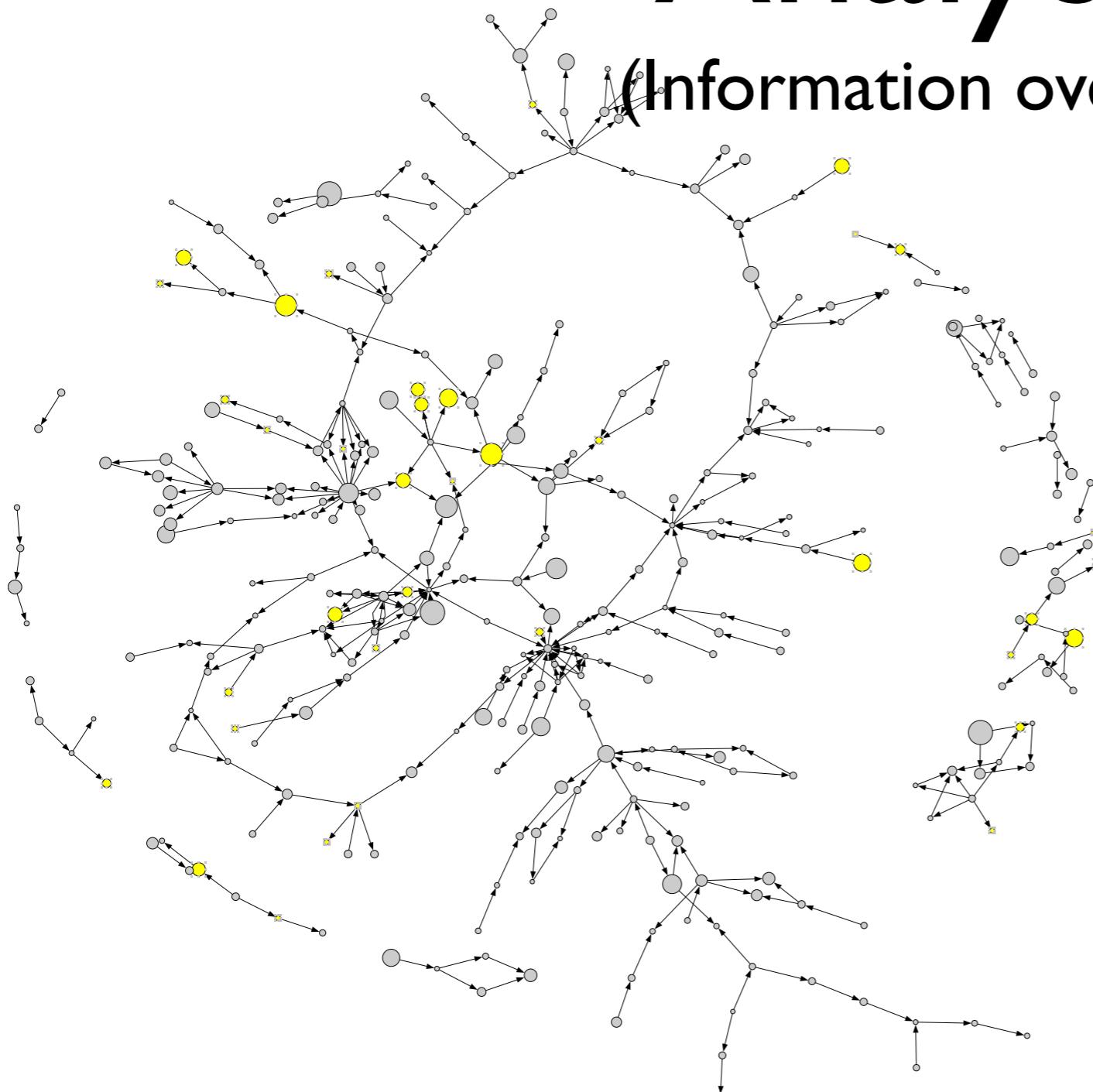


Such that ?y follows ?x
(inferred)

Such ?x is differentially expressed and ?y has variance>0.2 (from data)

Analysis

(Information overlay)



Selected elements in this p-p interaction graphs correspond to “controllers” in Kegg. The size of the nodes is proportional to the variance in gene expression

Analysis

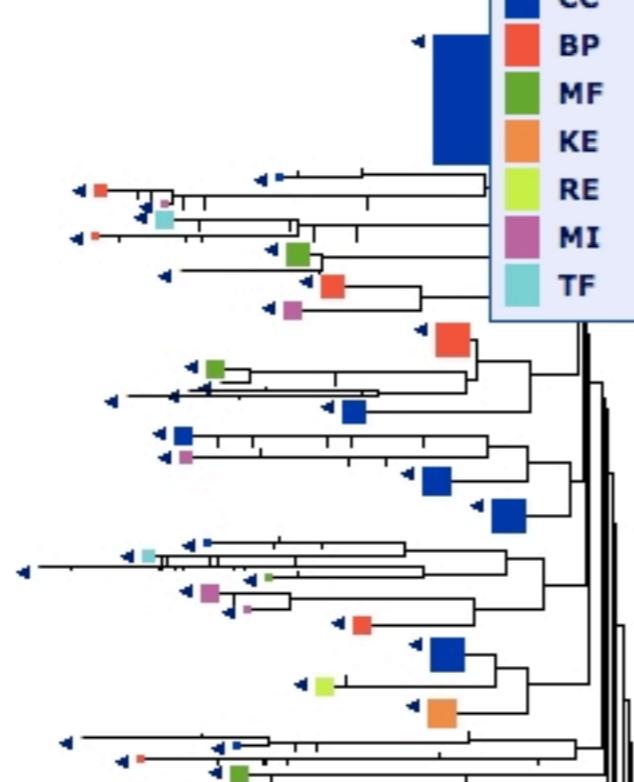
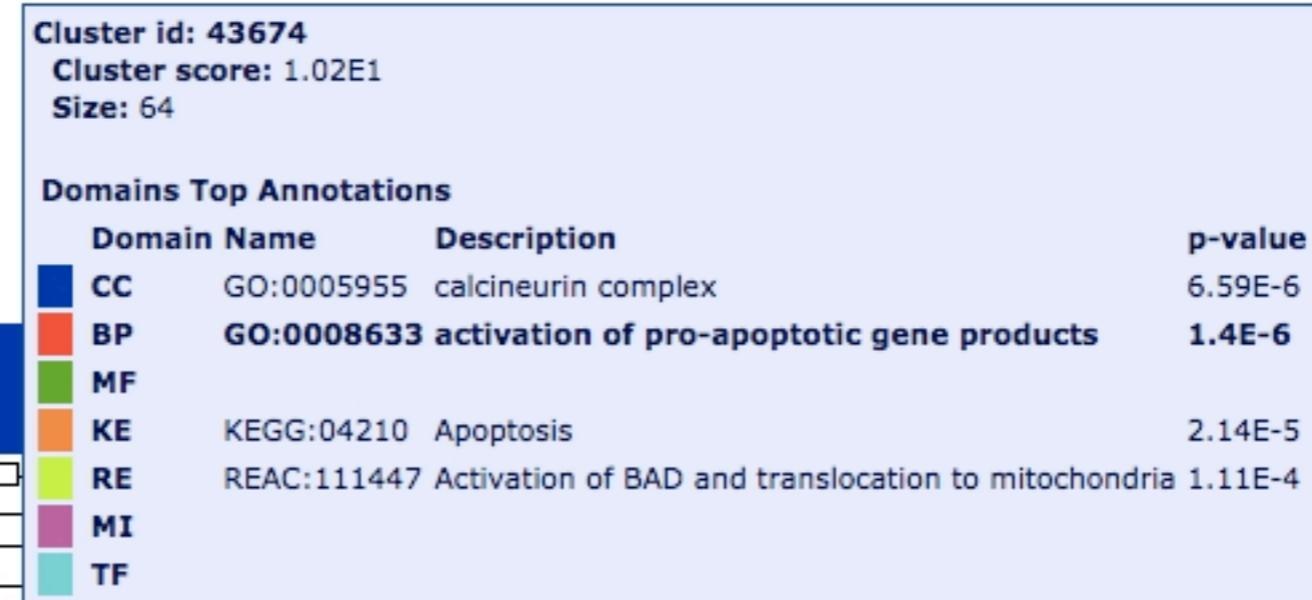
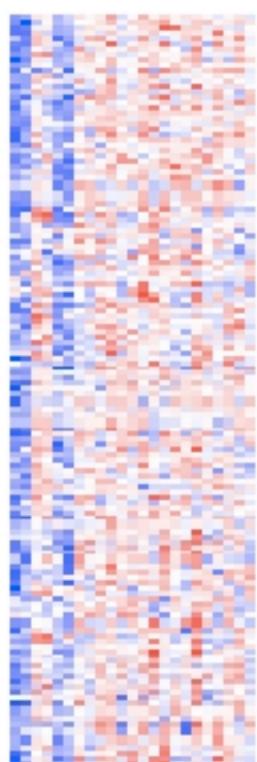
(Using ontologies for data analysis)

[HOME](#) >> [DATASET ANALYSIS](#)

[HeLP](#)

DATASET: GDS1023 HEMATOPOIETIC STEM CELL ENGRAFTMENT IN GOAT

Search for particular genes:



Source:
<http://biit.cs.ut.ee/vishic/>

Analysis

(Using ontologies for data analysis)

RDFscape

4) Query and browse ontologies Pattern library Map ontologies on my network

1) Select an analysis type! 2) Load ontologies 3) (optional) configure options

graph Pattern Matches

ayGenes

Matches

...	x1	x2	...
...	ns...	def:FakeInhibited	...
...	ns...	def:FakeActivated2	...
...	ns...	def:Peak3	...
...	ns...	def:Peak1	...
...	ns...	def:Pathway	...
...	ns...	def:FakeInhibited2	...
...	ns...	def:FakeActivated	...
...	ns...	def:Peak2	...

Chart Results

Pattern	Value	Pvalue
http://www.leaf...		
http://www.leaf...	0.158	
http://www.leaf...	0.073	
http://www.leaf...		
http://www.leaf...		
http://www.leaf...		
http://www.leaf...	0.109	
http://www.leaf...	0.18	

Evaluation function

$f(?x3)=\text{SUM}(\text{corr}(?x4 ?x5))/\text{NUM}(?x4)*\text{SUM}(\text{stdev}(?x4))/\text{NUM}(?x4))$

Search variables Add variables

Search patterns Add patterns Delete pattern

Compute Funct... Plot value Plot pvalue

Add rule

value > pattern Add R...

rule

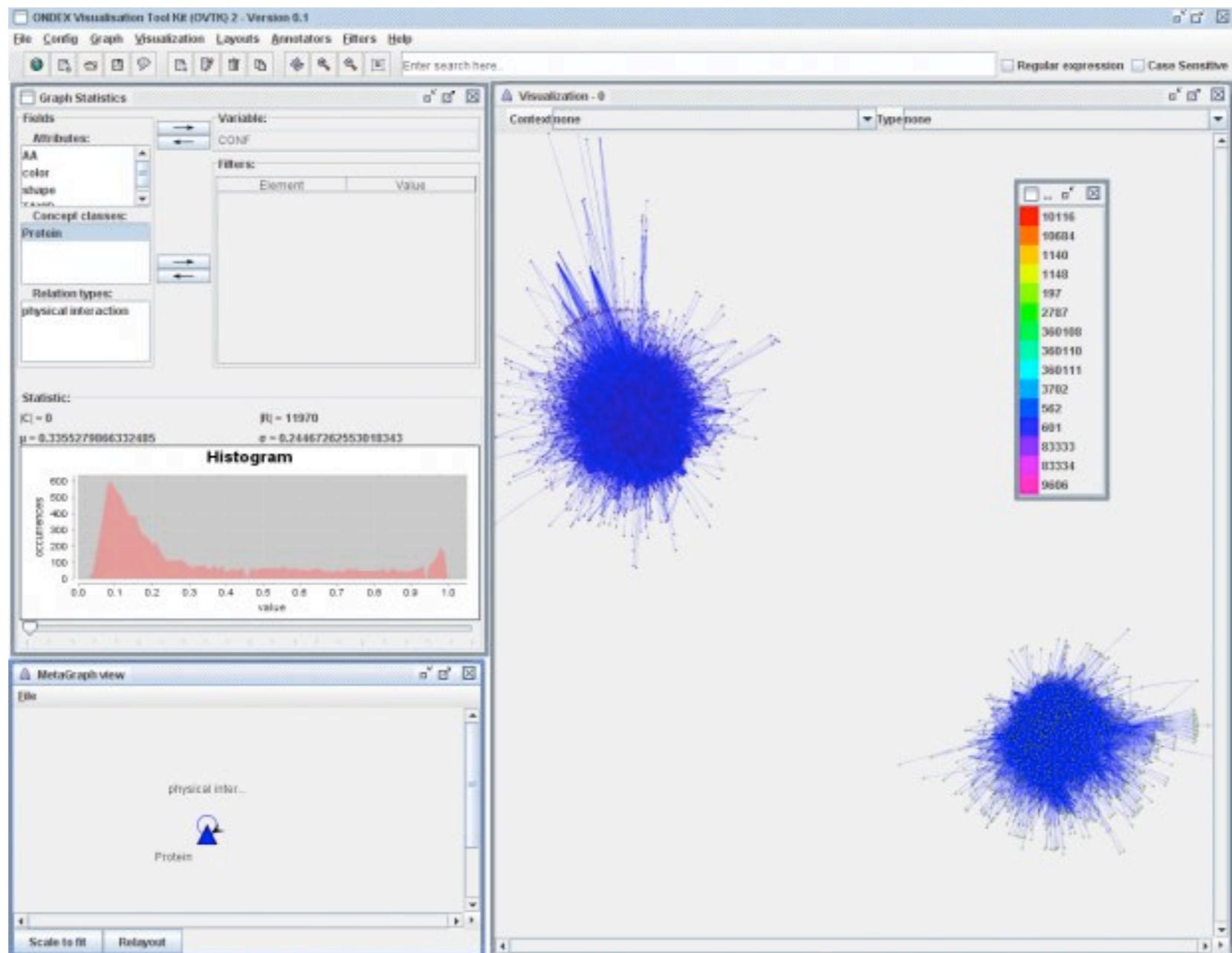
Random eval

Commit

Help

The screenshot shows the RDFscape interface for analyzing biological networks. On the left, a query graph is displayed with nodes for 'def:Pathway' (red), 'def:Gene' (red), and three question-mark placeholder nodes (grey). Edges include 'ns3:type' from pathway to gene, and various 'def:' predicates connecting the placeholder nodes. In the center, a 'Matches' table lists nine rows of triples, mostly in red. To the right, a 'Results' table shows eight rows with numerical values and P-values. At the bottom, an 'Evaluation function' input field contains a complex formula involving correlation and standard deviation. The interface includes tabs for 'graph', 'Pattern', and 'Matches', and various configuration and analysis tools.

Analysis



Source:
<http://ondex.org/>

Can generic analysis tools be brought on the Semantic Web ?

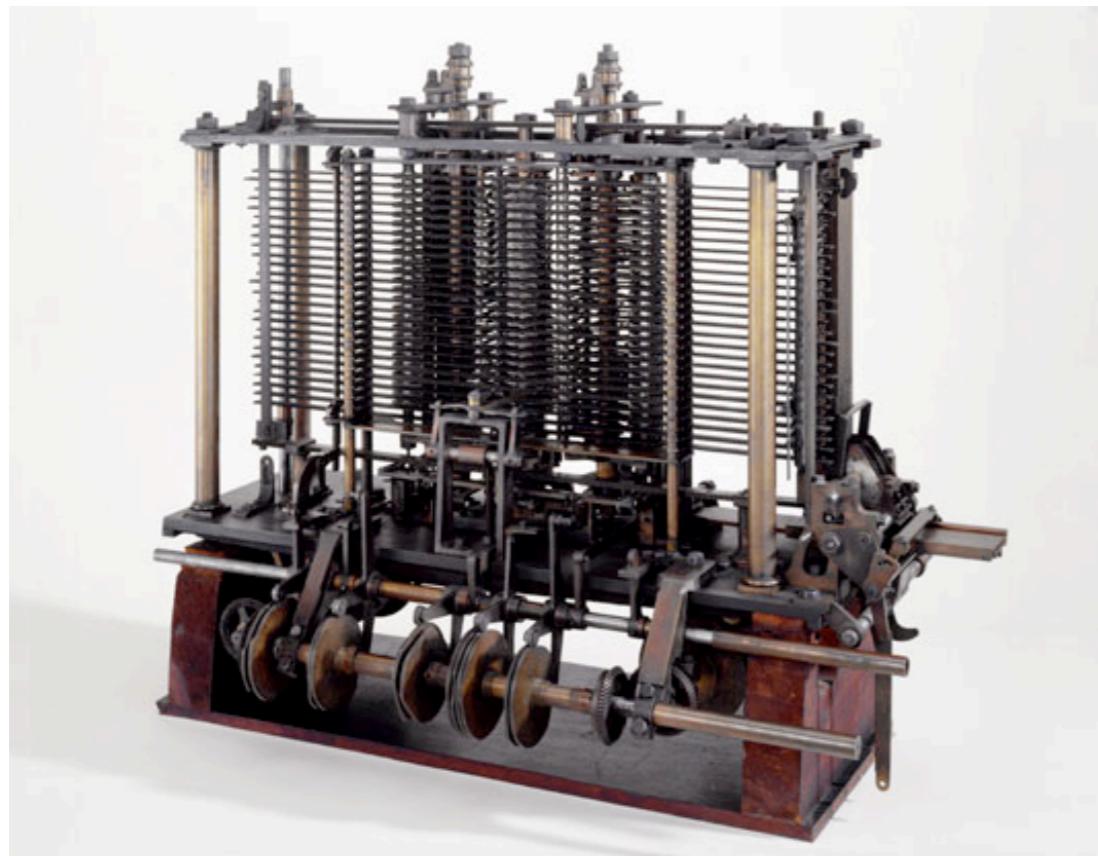


Conclusions @Hackathon

RDFScape:
Interaction+Query+Mapping+Reasoning+Analysis=
Interface

Conclusions @Hackathon

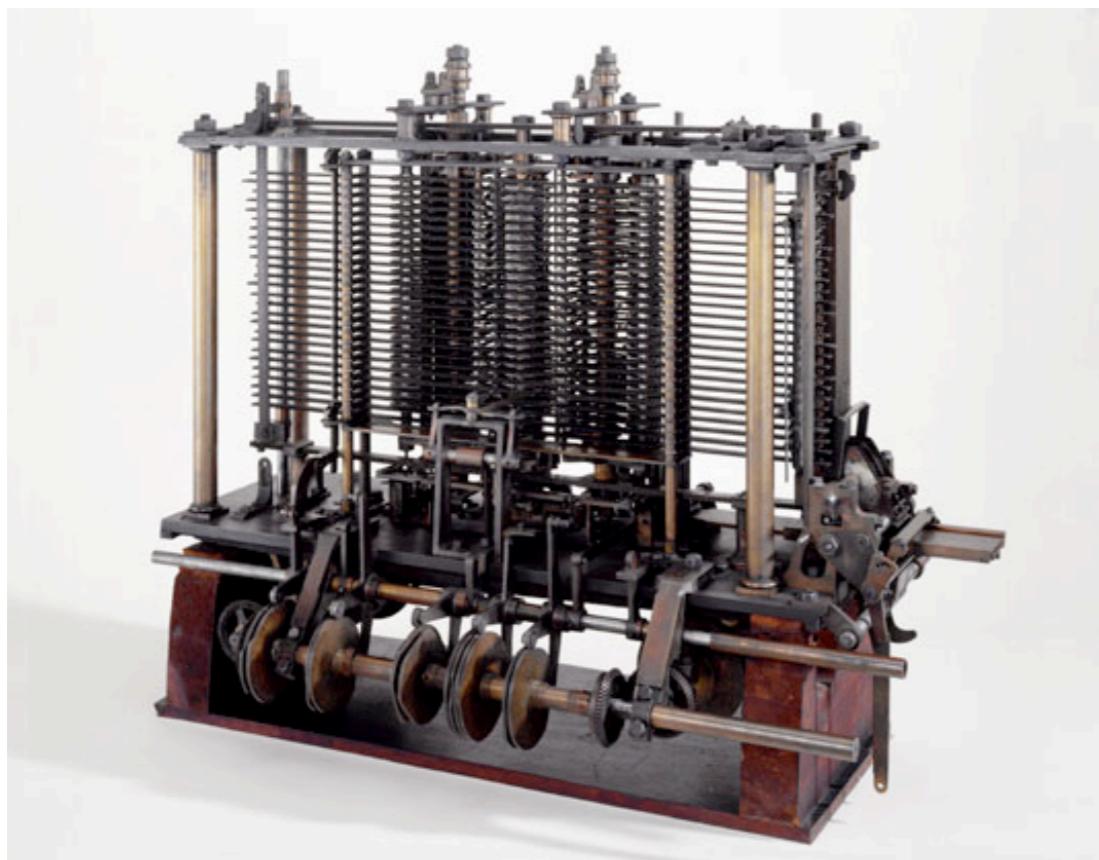
RDFScape:
Interaction+Query+Mapping+Reasoning+Analysis=Interface



RDFScape today

Conclusions @Hackathon

RDFscape:
Interaction+Query+Mapping+Reasoning+Analysis=Interface



RDFscape today



RDFscape after the Hackaton

References

RDFScape and network analysis (inference):
<http://www.biomedcentral.com/1471-2105/9/S4/S6/>

Semantic Web user interfaces (discussion):
<http://CEUR-WS.org/Vol-544/>

andrea.splendiani@bbsrc.ac.uk

The end

Thanks