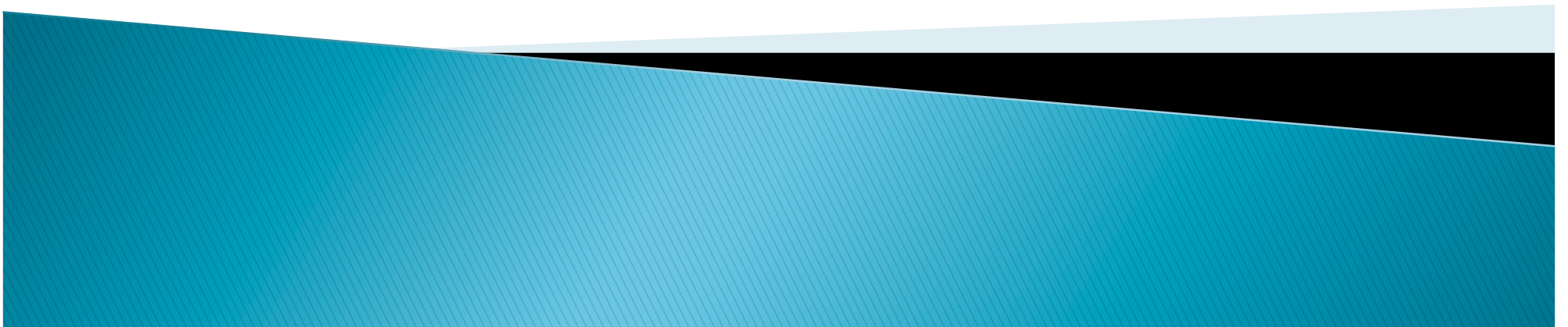


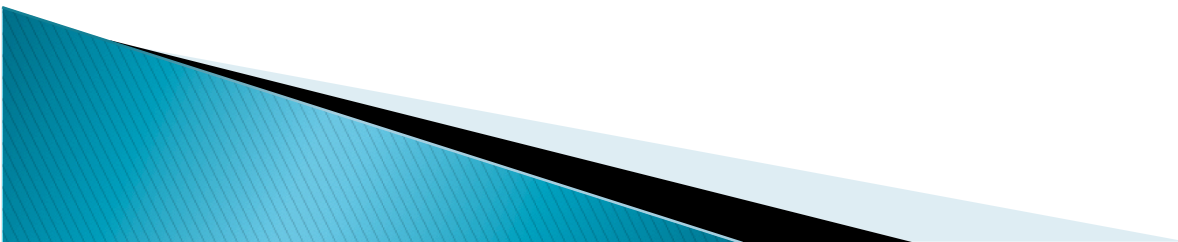
Biological annotation of systems biology models

Abhishek Tiwari
Auckland Bioengineering Institute



Metadata/Annotation

- ▶ Both CellML and SBML use RDF for metadata/annotation representation
- ▶ In SBML biological information is embedded as RDF in the language elements using annotation element while in CellML use of RDF is freestyle
- ▶ To report Biological informations SBML uses BioModels.net/ MIRIAM qualifier elements while CellML employs a **cmeta:bio_entity** element
- ▶ SBML elements can be annotated with **sboTerm** which adds additional semantic information
- ▶ Ongoing efforts to adopt the BioModels.net/ MIRIAM qualifier in CellML (proposed **cmeta:biomodels** ?)



```

<species metaid="MyModelElement" id="ALL" name="ActiveACh2" compartment="comp1" initialAmount="0">
  <notes>
    <body xmlns="http://www.w3.org/1999/xhtml">
      <p>biliganted active state</p>
    </body>
  </notes>
  <annotation>
    <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
             xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
             xmlns:bqmodel="http://biomodels.net/model-qualifiers/">
      <rdf:Description rdf:about="#MyModelElement">
        <bqbiol:hasPart>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:uniprot:P04551" />
            <rdf:li rdf:resource="urn:miriam:uniprot:P10815" />
          </rdf:Bag>
        </bqbiol:hasPart>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>

```

SBML

```

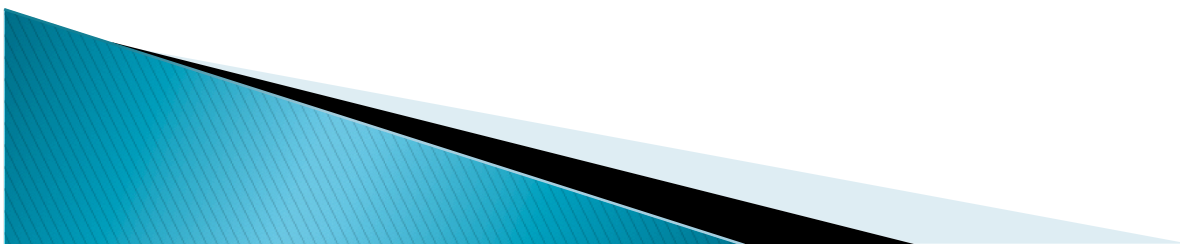
<rdf:RDF
  xmlns:cmeta="http://www.cellml.org/metadata/1.0#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <rdf:Description rdf:about="#cellml_element_id">
    <cmeta:bio_entity>
      <rdf:Bag>
        <rdf:li rdf:parseType="Resource">
          <dc:title>calmodulin</dc:title>
          <dcterms:alternative>CaM</dcterms:alternative>
          <cmeta:identifier rdf:parseType="Resource">
            <cmeta:identifier_scheme>SWISS-PROT</cmeta:identifier_scheme>
            <rdf:value>CALM_HUMAN</rdf:value>
          </cmeta:identifier>
        </rdf:li>
        <rdf:li rdf:parseType="Resource">
          <dc:title>troponin C</dc:title>
        </rdf:li>
        <rdf:li rdf:parseType="Resource">
          <cmeta:identifier rdf:parseType="Resource">
            <cmeta:identifier_scheme>SWISS-PROT</cmeta:identifier_scheme>
            <rdf:value>PRVA_HUMAN</rdf:value>
            <rdfs:label>parvalbumin</rdfs:label>
          </cmeta:identifier>
        </rdf:li>
      </rdf:Bag>
    </cmeta:bio_entity>
  </rdf:Description>
</rdf:RDF>

```

CellML

Annotating while developing

- ▶ Software tools like CellDesigner, SBMLeditor and COPASI support MIRIAM annotations
- ▶ These tools can be used for annotation during the model development process itself
- ▶ Wizard or Form based metadata/annotation editor interface



CellDesigner

BIOMD0000000001.xml

Grid Snap OFF

Species Proteins Genes RNAs asRNAs Reactions Compartments Parameters Functions UnitDefinitions Rules

Edit Export

class	id	name	speciesT...	comp...	positi...	included	quanti...	initialQuantity	su...	has...	b
PROTEIN	BLL	BasalACh2		comp1	inside		Amount	0.0		false	fa
PROTEIN	IL	Intermediate...		comp1	inside		Amount	0.0		false	fa
PROTEIN	AL	ActiveACh		comp1	inside		Amount	0.0		false	fa
PROTEIN	A	Active		comp1	inside		Amount	0.0		false	fa
PROTEIN	BL	BasalACh		comp1	inside		Amount	0.0		false	fa
PROTEIN	B	Basal		comp1	inside		Amount	1.660577881...		false	fa
PROTEIN	DLL	Desensitised...		comp1	inside		Amount	0.0		false	fa
PROTEIN	D	Desensitised		comp1	inside		Amount	0.0		false	fa
PROTEIN	ILL	Intermediate...		comp1	inside		Amount	0.0		false	fa
PROTEIN	DL	Desensitised...		comp1	inside		Amount	0.0		false	fa

NOTE MIRIAM

Relation

bqbiol.isVersionOf

New Edit Remove

DataType ID

InterPro IPR002394

Gene Ontology GO:0005892

New Edit Remove

SBMLeditor

sbml

- level: 2
- metaid: _000000
- version: 1
- xmlns: http://www.sbml.org/sbml/level2
- model: EPSP_Edelstein (Edelstein1996_EPSP_AChEvent)
 - id: EPSP_Edelstein
 - metaid: _000001
 - name: Edelstein1996_EPSP_AChEvent
 - notes: hidden
 - Creators
 - Family: Le Novère
 - Given: Nicolas
 - EMAIL: lenov@ebi.ac.uk
 - Orgname: EMBL-EBI
 - Links
 - is
 - urn:miriam:taxonomy:7787
 - isVersionOf
 - urn:miriam:obo.go:GO%3A0007274
 - urn:miriam:obo.go:GO%3A0007166
 - annotation:
 - bqmodel:is
 - urn:miriam:biomodels.db:BIOMD0000000001
 - bqmodel:isDescribedBy
 - urn:miriam:pubmed:8983160

Add Relation

Qualifier:

ID:

URI:

URIs: OMIM
Reactome
PubMed
DOI
Taxonomy
Gene Ontology
KEGG Pathway
BioModels Database
ICD

Relation set:

OK Reset Cancel

COPASI

BIOMD000000001 - COPASI 4.5 (Build 30) /home/.../curated/BIOMD000000001.cps

File Tools Help

Concentrations

Copasi

- Model
 - Biochemical
 - Mathematical
 - Diagrams
- Tasks
- Output
- Functions

Model | Annotation | RDF Browser

Created at 02/02/2005 14:56:11

Authors

	Status	Family Name	Given Name	Email	Organization
1		Le Novère	Nicolas	lenov@ebi.ac.uk	EMBL-EBI
2					

References

	Status	Resource	ID	Description
1		PubMed	8983160	
2		-- select --		

Description

	Status	Relationship	Resource	ID
1		is	BioModels Database	BIOMD000000001
2		is	Taxonomy	
3		is version of	Gene Ontology	GO:0007274
4		is version of	Gene Ontology	GO:0007166
5		-- select --	-- select --	

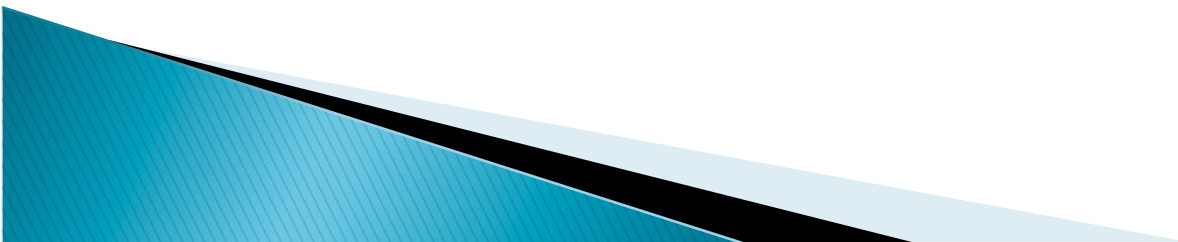
Modified at

	Status	Date and Time Modified
1		05/06/2009 11:40:04
2		00/00/0000 00:00:00

Commit Revert Clear Delete/Undelete New

Minimal vs Well annotated

- ▶ Tools described in previous slides require massive web hunting for the identifier values, for instance *In urn:miriam:uniprot:P62158 modellers have to find the uniprot id P62158 for the corresponding protein using keyword search*
- ▶ Finding identifier values of all miriam:dataType for all components/elements in a large model can be mammoth task
- ▶ Additional complexity introduced due to relationship between MIRIAM resources in form of “has a”, “is version of”, “is homolog to”, etc.
- ▶ Not good enough for full fledged annotation process resulting into a well annotated model

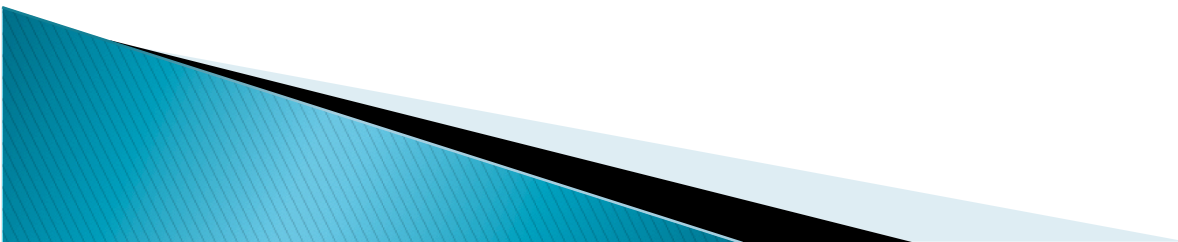


So how do we annotate model with these many resources?



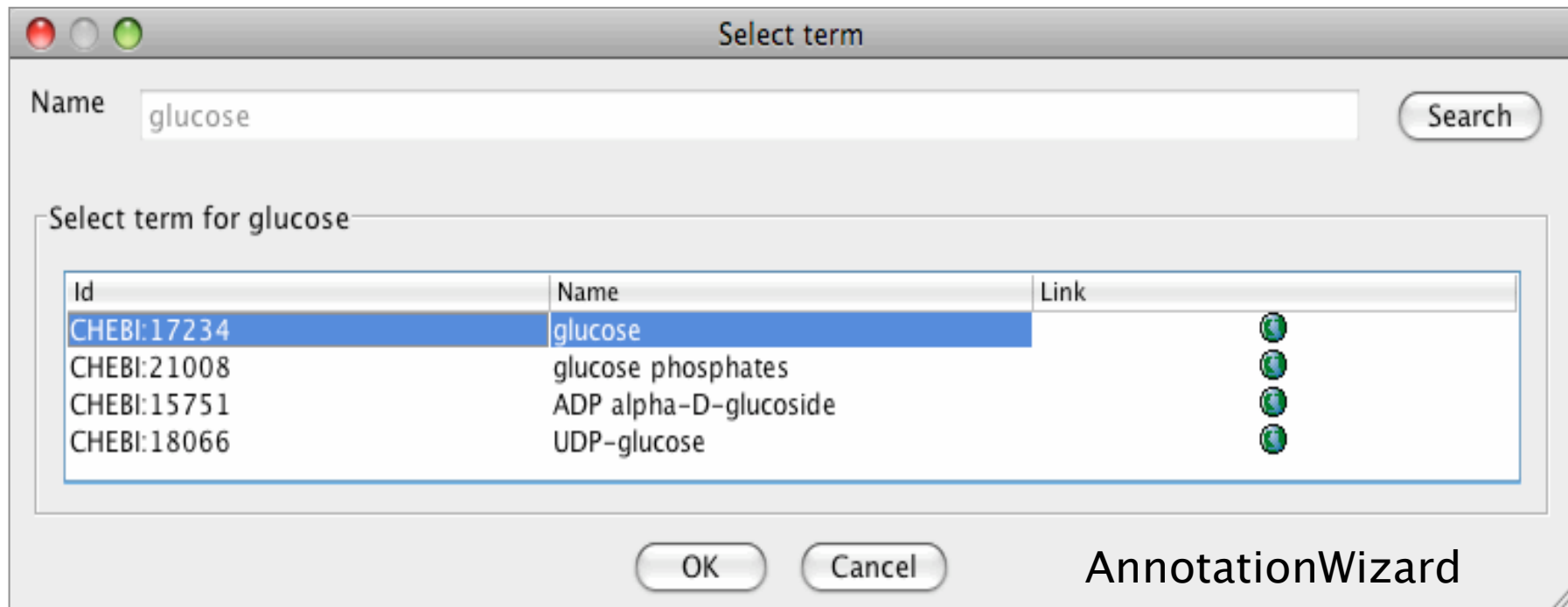
Creating annotation pipeline using webservices

- ▶ Individually querying web service for each of these databases using keywords and update the model
- ▶ Unified interface which allows to interact with a list of web services in a asynchronous way.
- ▶ Tools such as libAnnotationSBML, Saint and semanticSBML can automate this process
- ▶ Dynamic linking of model to databases is in early stage



libAnnotationSBML

- ▶ Java library and tools



semanticSBML

The screenshot displays the semanticSBML application window. The main interface is divided into a left-hand tree view and a right-hand details panel. The tree view shows a hierarchical model structure for 'HOG_Phosphorelay_module_11_06_09', including Models, Compartments, Species, Parameters, Reactions, and Rate rules. The 'Species' section is expanded, showing 'Sln1p' selected. The details panel on the right, titled 'Sln1p in cytosol', shows 'Current Annotations' with two entries: 'bio:is Gene Ontology GO:0004673 protein histidine kinase activity' and 'bio:is SGD S000001409'. Each entry has a 'Bio:is' dropdown, a 'change Qualifier' button, and a 'remove Annotation' button. Below the annotations are 'Copy and paste annotations' buttons (Copy, Paste), a 'Search Annotation by Name' section with a search box containing 'water' and a 'search' button, and an 'Add Annotation' section with a table for Database, Qualifier, and Identifier, and an 'add' button.

File Help

Main Build model Configure Annotations Checks

HOG_Phosphorelay_module_11_06_09

Model HOG_Phosphorelay_module_11_06_09

- Models
 - Phosphorelay module [ID: Model_1]
- Compartments
 - cytosol [ID: compartment_1]
- Species
 - Sln1p [ID: species_1]
 - Sln1pP [ID: species_2]
 - Ssk1p [ID: species_3]
 - Ssk1pP [ID: species_4]
 - Ypd1p [ID: species_5]
 - Ypd1pP [ID: species_6]
- Parameters
- Reactions
 - v1_mod1 [ID: reaction_1]
 - v3_mod1 [ID: reaction_2]
 - v4_mod1 [ID: reaction_3]
 - v2_mod1 [ID: reaction_4]
- Assignment rules
 - (Unnamed assignmentRule) [ID: parameter_4 [Assignment Rule]]
- Rate rules
 - (Unnamed rateRule) [ID: compartment_1 [Rate Rule]]

MIRIAM SBO Information

Sln1p in cytosol

Current Annotations

bio:is [Gene Ontology GO:0004673](#)
protein histidine kinase activity

Bio:is change Qualifier remove Annotation

bio:is [SGD S000001409](#)

Bio:is change Qualifier remove Annotation

Copy and paste annotations Copy Paste

Search Annotation by Name

water search

Add Annotation

Database	Qualifier	Identifier	
ChEBI	Bio:is		add

Wish list?

- ▶ MIRIAM annotation in CellML
- ▶ Better web services

