



BioModels.net: semantics for model sharing

Nicolas Le Novère, EMBL-EBI, United-Kingdom

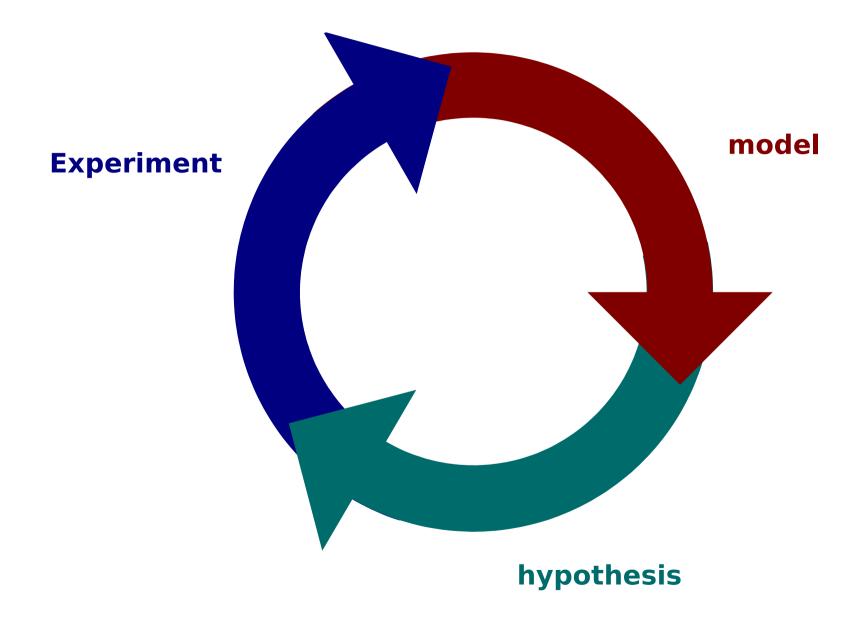
BioPAX

MIRIAM

SBGN



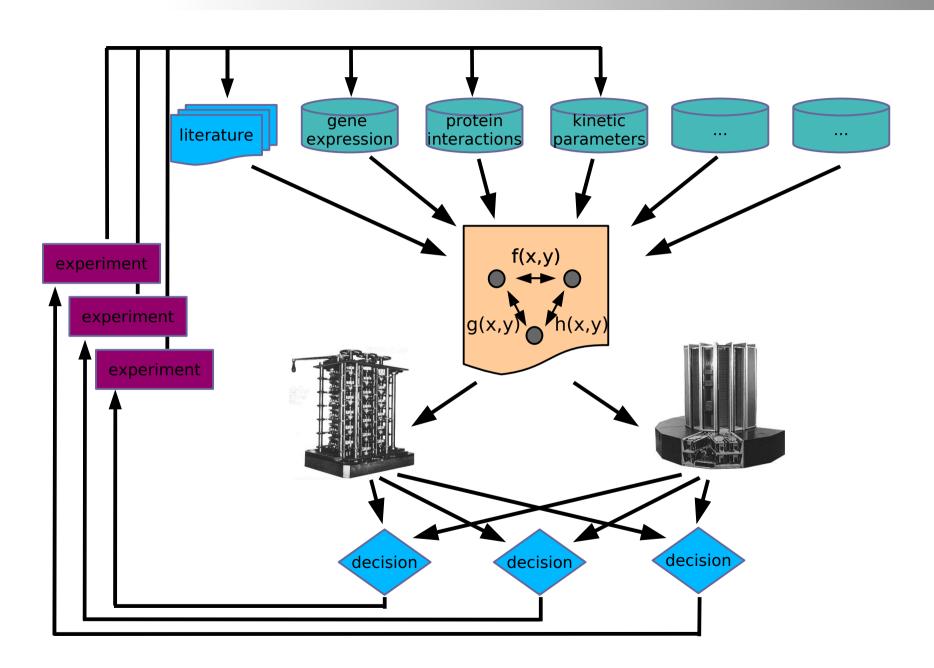








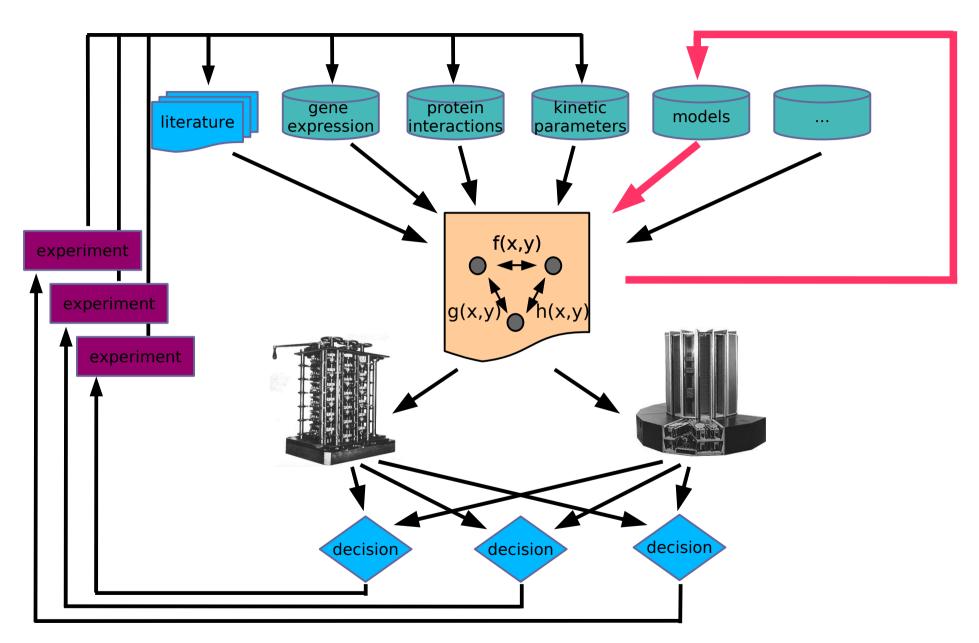
The model as an integrator of knowledge







The model as an integrator of knowledge







A multiscale problem

10-10

 10^{-8}

 10^{-7}

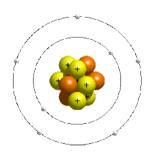
 10^{-5}

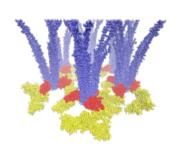
10⁻²

 10^{-0}

m

spatial scale problem





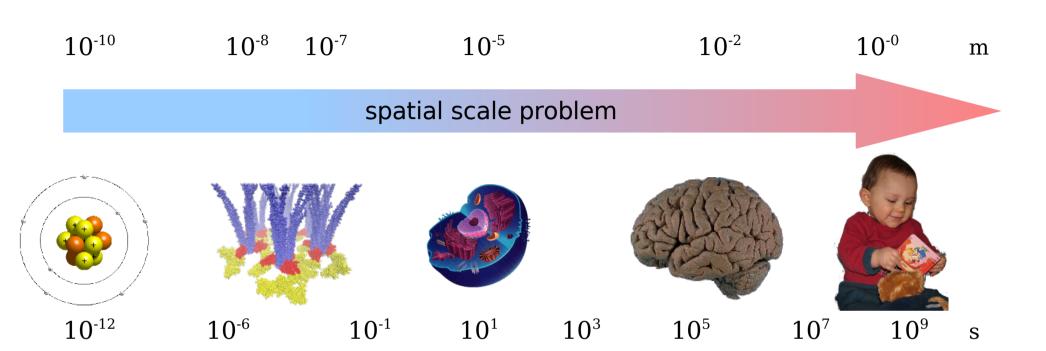








A multiscale problem



time scale problem





A simulation problem

• Molecular dynamics: to simulate $\propto 10^{-12}$ s requires $\propto 1$ s

· Particle diffusion: to simulate $\propto 10^{-6}$ s requires $\propto 1$ s

Stochastic chemical kinetics: to simulate $\propto 1$ s requires $\propto 1$ s

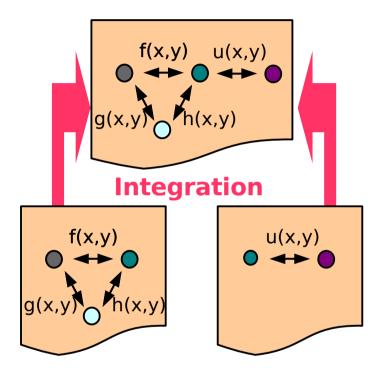
• Continuous ODE: to simulate $\propto 10^3$ s requires $\propto 1$ s

⇒ Humongous stiffness: the speed of the whole simulation is determined by the quickest event





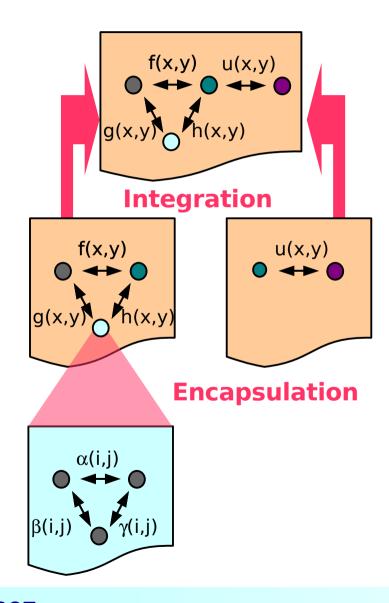






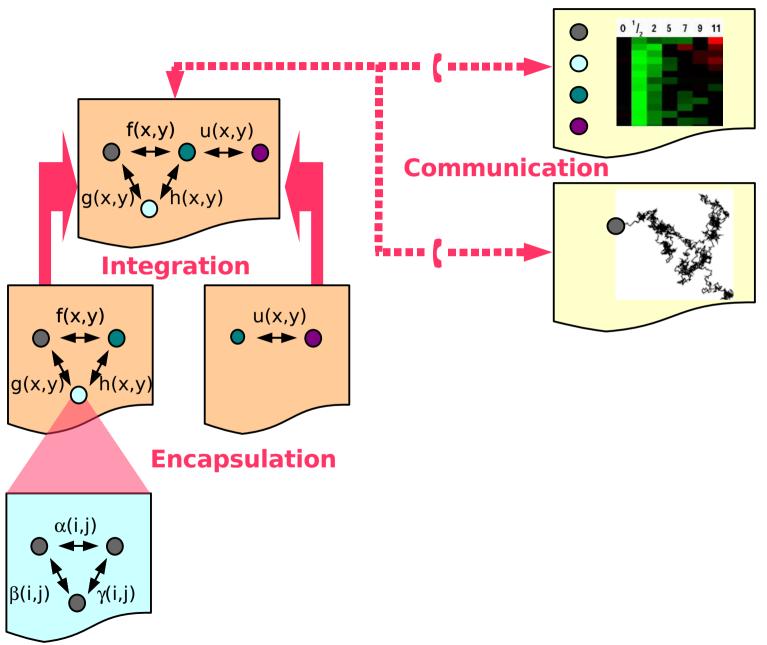








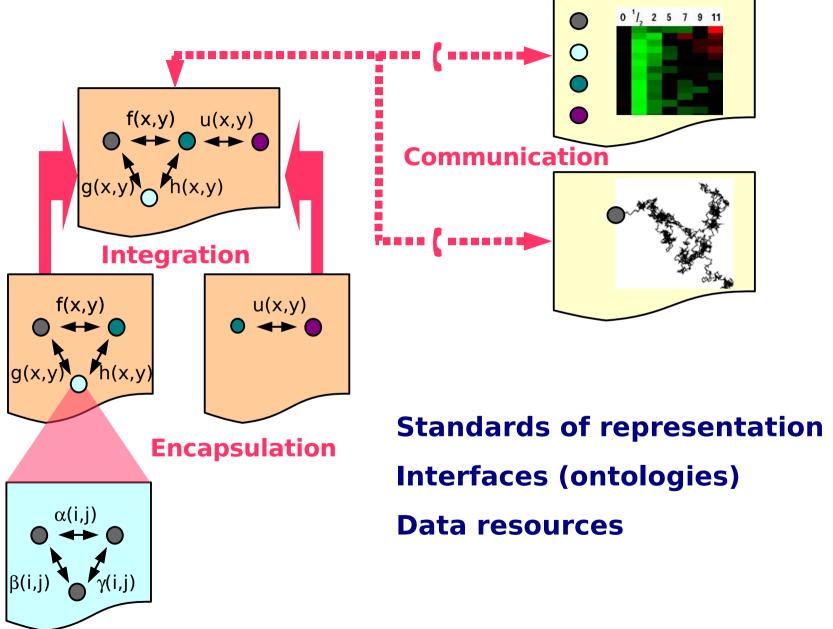
















Standards of representation

"The nice thing about standards is that there are so many to choose from".

Attributed to Andrew S Tanenbaum





Standards of representation



http://www.cellml.org/
Based on modules; scalable;



http://www.neuroml.org/
Flexible (expendable set of classes/schemas);

BrainML.org

http://brainml.org/
Models are XML-schemas



http://www.biopax.org/
No kinetics; deep semantics; OWL

SGN

http://www.sbgn.org/ Graphical representation of interactions

Systems Biology http://sbml.org/
Markup Language ch kinetics; weak semantics; XML





Google	
	Search this site

home • contacts • documents • downloads • FAQs • forums • Level 3 • models • news • online tools • wiki • workshops

The Systems Biology Markup Language (SBML) is a computer-readable format for representing models of biochemical reaction networks. SBML is applicable to metabolic networks, cell-signaling pathways, regulatory networks, and many others.

Internationally Supported and Widely Used

SBML has been evolving since mid-2000 through the efforts of an international group of software developers and users. Today, SBML is supported by over 100 software systems, including the following (where '* indicates SBML support in development):

BALSA	Dizzy	Moleculizer	SBMLR
BASIS	E-CELL	Monod	SBMLSim
BIOCHAM	ecellJ	Narrator	SBMLToolbox
BioCharon	ESS	NetBuilder	SBliD
ByoDyn	FluxAnalyzer	Oscill8	SBToolbox
BioCyc	Fluxor	PANTHER Pathway	SBW
BioGrid	Gepasi	PathArt	SCIpath
BioModels	Gillespie2	PathScout	Sigmoid*
BioNetGen	HSMB	Pathway Analyser	SigPath
BioPathwise	HybridSBML	PathwayLab	SigTran
Bio Sketch Pad	INSILICO discovery	Pathway Tools	SIMBA
BioSens	JACOBIAN	PathwayBuilder	SimBiology
BioSPICE Dashboard	Jarnac	PATIKAweb	Simpathica
BioSpreadsheet	JDesigner	PaVESy	SimPheny*
BioTapestry	JigCell	PET	SimWiz
BioUML	JSim	PNK	SloppyCell
BSTLab	JWS Online	PottersWheel	SmartCell
CADLIVE	Karyote*	Reactome	SRS Pathway Editor
CellDesigner	KEGG2SBML	ProcessDB	StochSim
Cellerator	Kineticon	PROTON	StochKit
CellML2SBML	Kinsolver*	pysbml	STOCKS
Cellware	libSBML	PySCeS	TERANODE Suite
CL-SBML	MathSBML	runSBML	Trelis
CLEML	MesoRD	SABIO-RK	Virtual Cell
COPASI	MetaboLogica	SBML ODE Solver	WebCell
Cyto-Sim	MetaFluxNet	SBML-PET	WinSCAMP
Cytoscape	MMT2	SBMLeditor	XPPAUT

BioNetGen@VCell Release

(October 6, 2006) BioNetGen@VCell is a new release of BioNetGen, a tool for automatically generating a reaction network from user-specified rules for biomolecular interactions on the level of protein domains.

read more

PottersWheel supports SBML

(October 4, 2006) PottersWheel 1.2 beta, a MATLAB systems biology toolbox, supports model creation, fitting data, and designing new experiments.

read more

SBML Level 2 Version 2 Released!

(September 25, 2006) The final version of the SBML Level 2 Version 2 specification is now available!

read more

SBML Wikipedia entry

(September 18, 2006) There is now an updated entry for SBML in Wikipedia. Let us know your suggestions for improvements.

read more

SBML Tutorial at ICSB 2006

(September 8, 2006) Mike Hucka will be leading a tutorial on SBML this year at ICSB 2006 in Japan. The focus will be on the about-to-be-released SBML Level 2 Version 2.

-III:

read more

See older news items.

A Free and Open Language

Modesto

Advances in hiotechnology are leading to larger more complex quantitative models. The systems hiology

SBMLmerge

DBsolve



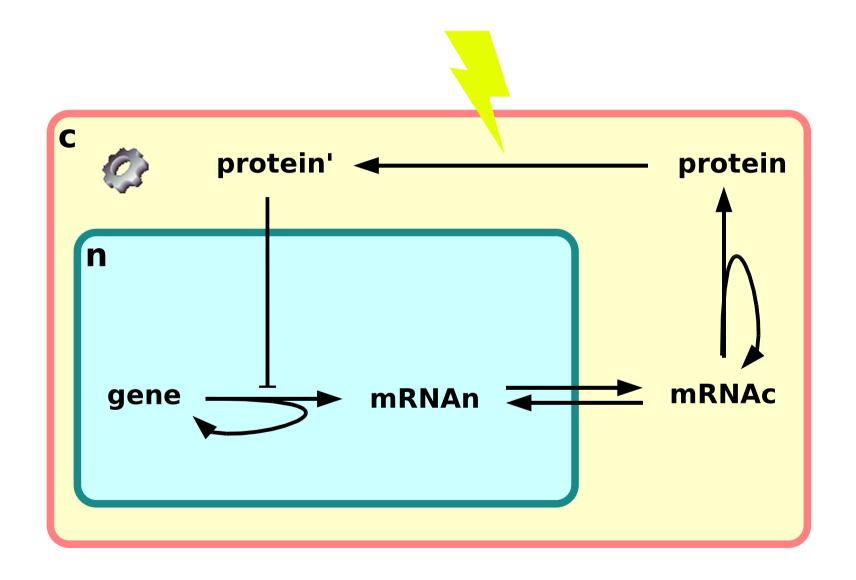


"The goal of SBML is to help people to disagree as precisely as possible". Ed Franck, Argonne National Laboratory













```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="1" xmlns="http://www.sbml.org/sbml/level2">
  <model>
    <listOfCompartments>
      <compartment id="cell" />
    </listOfCompartments>
    <listOfSpecies>
      <species id="A" compartment="cell" initialConcentration="1"/>
      <species id="B" compartment="cell" initialConcentration="0"/>
    </listOfSpecies>
    <listOfParameters>
      <parameter id="kon" value="1"/>
    </listOfParameters>
    <listOfReactions>
      <reaction>
        <listOfReactants>
          <speciesReference species="A" />
        </listOfReactants>
        <listOfProducts>
          <speciesReference species="B" />
        </listOfProducts>
        <kineticLaw>
          <math xmlns="http://www.w3.org/1998/Math/MathML">
            <apply>
              <times />
              <ci>kon</ci>
              <ci>A</ci>
              <ci>ci>cell</ci>
            </apply>
          </kineticLaw>
      </reaction>
    </listOfReactions>
  </model>
</sbml>
```





SBML is not limited to biochemistry!

- Rate Rules can describe the temporal evolution of <u>any quantitative</u> <u>parameter</u>, e.g. transmembrane voltage;
- Events can describe any discontinuous change, e.g. neurotransmitter release;
- A species is an entity participating to a reaction, not always a chemical entity:
 - It can be a molecule
 - It can be a cell
 - It can be an organ
 - It can be an organism
- → Remember, Systems Biology is scale-free!







- · Released on September 25th 2006
- · Simpler and cleaner (units ...)
- Generic entities (compartmentType, speciesType)
 - → path to generalised reactions
- Constraints and initialAssignments
- Controlled (MIRIAM) annotations (+ links to SBO)
- Backward compatible with Level 2 Version 1
- More detailed and bug-free specification ... 145 pages, 10pt, small margin.





- Modular SBML, with core + optional packages
- Graph Layout
- Generalised reactions (probable)
- Model composition (probable)
- Complex species (probable)
- Arrays or sets (maybe)
- · Geometry (maybe)
- Movements (maybe)
- Dynamic compartments (maybe)
- · ???





Is SBML enough? What's missing?

- An SBML model lists participants, but does not identify them.
- An SBML model contains mathematical expressions, but does not tell-us what they "mean", and how they are derived.
- An SBML model constructed for a certain modelling approach cannot be used straight-away within another modelling framework.
 - ⇒ SBML models cannot be easily searched SBML models cannot be easily converted SBML models cannot be easily merged





Minimum Information Requested In the Annotation of biochemical Models

Le Novère N., Finney A., Hucka M., Bhalla U., Campagne F., Collado-Vides J., Crampin E., Halstead M., Klipp E., Mendes P., Nielsen P., Sauro H., Shapiro B., Snoep J.L., Spence H.D., Wanner B.L.

Nature Biotechnology (2005), 23: 1509-1515

http://www.ebi.ac.uk/compneur-srv/miriam/

Reference correspondence

Thou shalt encode your model in a public, standardized, machine-readable format (SBML, CellML, GENESIS ...)

Thy model shalt comply with the standard in which it is encoded!

Thy model shalt be clearly related to a single reference description.

If thy model is composed from different parts, thou shalt still quote a description of the derived/combined model.

The encoded structure of the model must reflect the biological processes listed in the reference description.

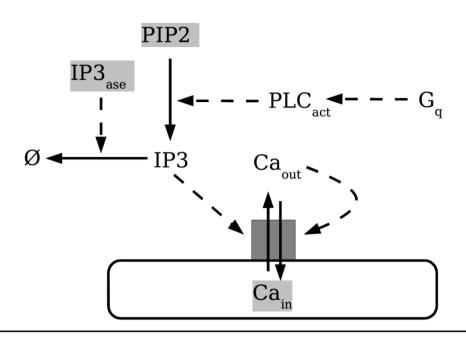
Thy model shalt be instantiated in a simulation: All quantitative attributes shalt be defined, including initial conditions.

When instantiated, thy model must be able to reproduce all results given in the reference description within an epsilon (algorithms, round-up errors)





Model example



$$k_1 = k_2 = k_3 = 1 \, s^{-1}$$

$$Km_1 = 10^{-7} M, Km_2 = 10^{-8}, Km_3 = 2.10^{-6} M$$

$$K_A = 10^{-11}, m = 4, n = 3, \alpha = 0.001$$

$$\frac{d[Ca_{out}]}{dt} = \frac{k_1[IP3R] * ([Ca_{in}] - [Ca_{out}])}{Km_1 + |[Ca_{in}] - [Ca_{out}]|} * \frac{[IP3]^m}{K_A + [IP3]^m}$$

$$\frac{d[IP3]}{dt} = \frac{k_2[PLC_{act}] * [PIP2]}{Km_2 + [PIP2]} - \frac{k_3[IP3_{ase}] * [IP3]}{Km_3 + [IP3]}$$
$$\frac{d[PLC_{act}]}{dt} = \frac{[G_q]^n}{\alpha + [G_q]^n} * [PLC_{tot}]$$

$$[Ca_{in}] = [IP3R] = [PLC_{tot}] = [PIP2] = [IP3_{ase}] = 0.001 M$$

$$[G_q] = 0.01 M, [Ca_{out}] = [IP3] = [PLC_{act}] = 0 M$$



Attribution annotation

Thy model has to be named.

Thou shalt join a citation of the reference description (complete citation, unique identifier, unambigous URL). The citation shalt permit to identify the *authors* of thy model.

Thou shalt join the name and contact of model *creators*.

Thou shalt specify the date and time of creation and last modification. An history is useful but not required.

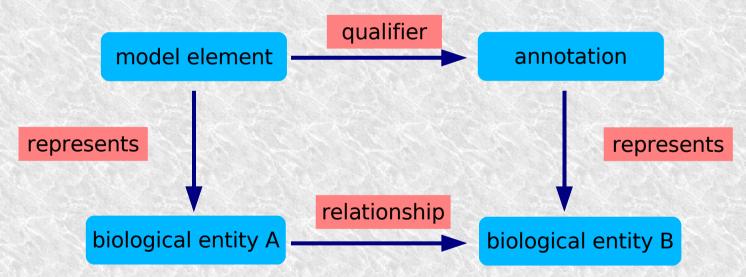
Thou shalt link thy model to a precise statement about the terms of distribution. MIRIAM does <u>not</u> require "freedom of use" or "no cost".



External resource annotation

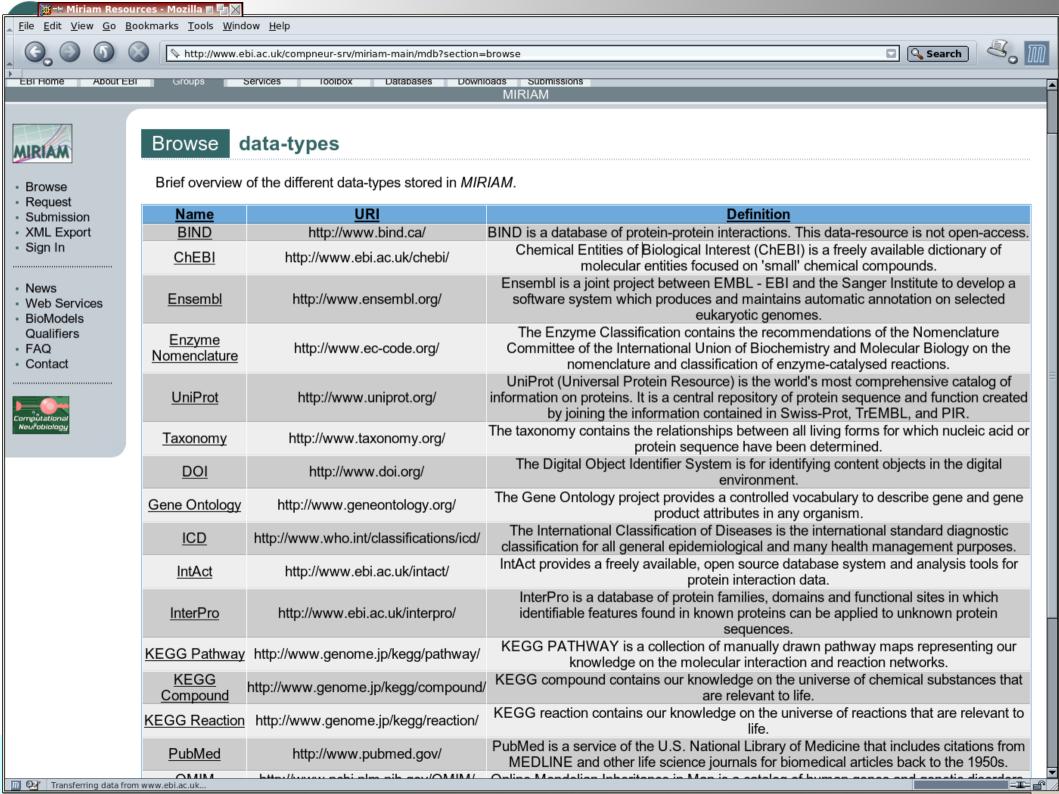
Thy annotation shalt permit to unambiguously relate a piece of knowledge to a model constituent.

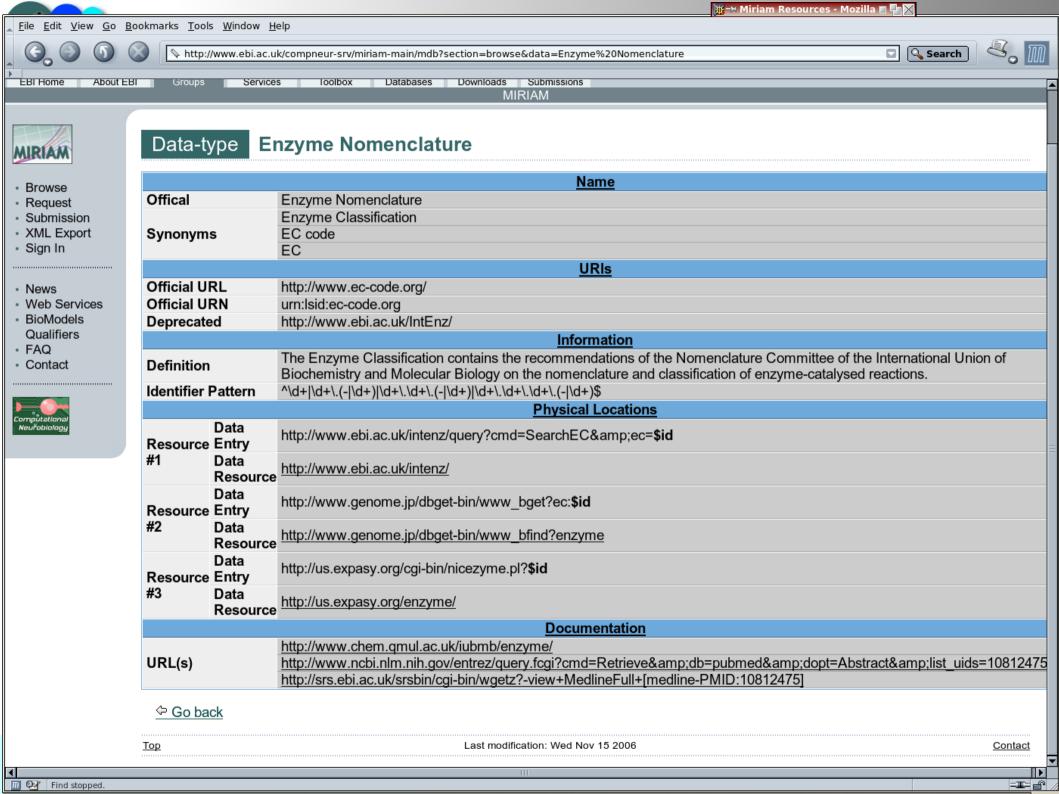
Thy referenced information shalt be described using a triplet {data-type, identifier, qualifier}

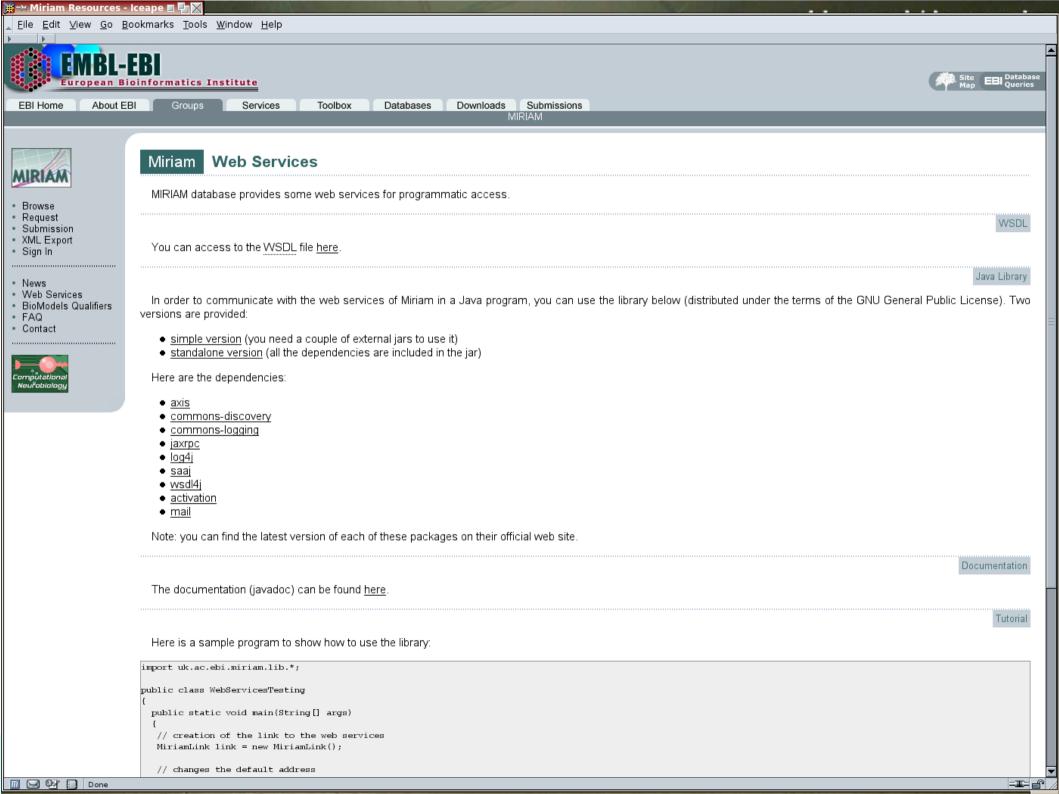


The community has to agree on a set of standard valid data-types. A database and the associated API (WebServices) have been developed at the EBI to provide the generation and interpretation of URIs.









creators creation date last modification	Joe User (juser@eden.com), Anne Other (aother@eden.com) 01 January 0000 31 May 2005				
Constituent	Data Type	Identifier	Qualifier	Meaning	
model	http://www.pubmed.gov/ http://www.taxonomy.org/ http://www.geneontology.org/ http://www.geneontology.org/ http://www.genome.jp/kegg/pathway http://www.genome.jp/kegg/pathway	0000000 9606 GO:0007204 GO:0051279 hsa04020 hsa04070	IsVersionOf IsVersionOf IsPartOf IsPartOf	Homo sapiens positive regulation of cytosolic ca2+ concentration regulation of release of sequestered ca2+ into cytop Calcium signaling pathway—H sapiens Phosphatidylinositol signaling system—H sapiens	
compartment ER	http://www.geneontology.org/	GO:0005790		smooth endoplasmic reticulum	
reactant Ca_{in}	http://www.ebi.ac.uk/chebi/	CHEBI:29108		calcium(2+)	
compartment cytoplasm	http://www.geneontology.org/	GO:0005737		cytoplasm	
reactant Ca _{out}	http://www.ebi.ac.uk/chebi/	CHEBI:29108		calcium(2+)	
reactant IP3	http://www.ebi.ac.uk/chebi/	CHEBI:16595		1D-myo-inositol 1,4,5-tris(dihydrogen phosphate)	
reactant PIP2	http://www.ebi.ac.uk/chebi/	CHEBI:18348		1-phosphatidyl-1D-myo-inositol 4,5-bisphosphate	
reactant IP3R	http://www.uniprot.org/ http://www.uniprot.org/ http://www.uniprot.org/	Q14643 Q14571 Q14573	Has Version Has Version Has Version	Inositol 1,4,5-trisphosphate receptor type 1 Inositol 1,4,5-trisphosphate receptor type 2 Inositol 1,4,5-trisphosphate receptor type 3	
reactant PLC _{act}	http://www.uniprot.org/	Q9NQ66	IsVersionOf	PIP2 phosphodiesterase $\beta 1$	
reactant PLC_{tot}	http://www.uniprot.org/	Q9NQ66		PIP2 phosphodiesterase β 1	
reactant IP3 _{ase}	http://www.uniprot.org/	Q14642		Type I inositol-1,4,5-trisphosphate 5-phosphatase	
reactant G_q	http://www.uniprot.org/	Q6NT27		Guanine nucleotide binding protein Gq	
reaction Ca _{release}	http://www.geneontology.org/ http://www.geneontology.org/	GO:0005220 GO:0008095	IsVersionOf	IP3-sensitive calcium-release channel activity IP3 receptor activity	
reaction IP3 _{production}	http://www.geneontology.org/ http://www.ec-code.org/	GO:0004435 3.1.4.11	IsVersionOf IsVersionOf	phosphoinositide phospholipase C activity phosphoinositide phospholipase C	
reaction IP3 $_{degradation}$	http://www.ec-code.org/	3.1.3.56	IsVersionOf	inositol-polyphosphate 5-phosphatase	
reaction PLC _{activation}	http://www.geneontology.org/	GO:0007200		G-protein signaling coupled to IP3 2nd messenger	



```
<species metaid="metaid 0000055"</pre>
     id="boundEGFReceptor"
                                                  Epidermal Growth Factor receptor
     compartment="cell"
     initialConcentration="0">
    <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="#metaid 0000055">
                <bgbiol:hasPart>
                    <rdf:Baq>
                        <rdf:li rdf:resource="http://www.uniprot.org/#Q9QX70"/>
                        <rdf:li rdf:resource="http://www.uniprot.org/#P07522"/>
                    </rdf:Bag>
                </bgbiol:hasPart>
            </rdf:Description>
        </rdf:RDF>
    </annotation>
</species>
```

Epidermal Growth Factor





Hidden assumptions

```
<reaction>
   <listOfReactants>
     <speciesReference species="S" />
   </listOfReactants>
   stOfProducts>
     <speciesReference species="P" />
   </listOfProducts>
                                                                            S \rightarrow P
   <listOfModifiers>
     <speciesReference species="E" />
                                                           Import in a discrete simulator
   </listOfModifiers>
   <kineticLaw>
     <listOfParameters>
       <parameter id="Km"/>
       <parameter id="kp"/>
     </listOfParameters>
     <math xmlns="http://www.w3.org/1998/Math/MathML">
       <apply>
                                                       E+S \rightleftharpoons ES \rightarrow E+P; k_1 = k_1/K_m
          <divide/><apply>
                       <times/><ci>E</ci>
                                <ci>kp</ci>
                                <ci>S</ci>
                    </apply>
                                                       E+S \rightarrow ES \rightarrow E+P; k_1 = k_n/K_m
                    <apply>
                       <plu><plus/><ci>Km</ci>
                               <ci>S</ci>
                    </apply>
       </apply>
                                                       E+S \stackrel{\cdot}{\rightleftharpoons} ES \stackrel{\cdot}{\Rightarrow} E+P ; k_1 = (k_{-1}+k_p)/K_m
     </kineticLaw>
                                                              k_{-1}
</reaction>
```



The Systems Biology Ontology

http://www.ebi.ac.uk/sbo/



Classifications Vs. Ontologies

- Ontology: A set of elements of knowledge linked with sense-bearing relationships.
- Each term of an ontology is associated to a <u>perennial</u> identifier.
 Once created a term is never destroyed. It can be merged with another, or made obsolete, but it still exists.
- An ontology is an evolving structure: It can cope with an increase or refinement of knowledge. No need to reconstruct everything as with the taxonomies.
- An ontology is a Direct Acyclic Graph, and not a hierarchy. A term can possess more than one parent.
- Ontologies are stored in standard machine-readable formats. They can be subjected to automatic treatments.

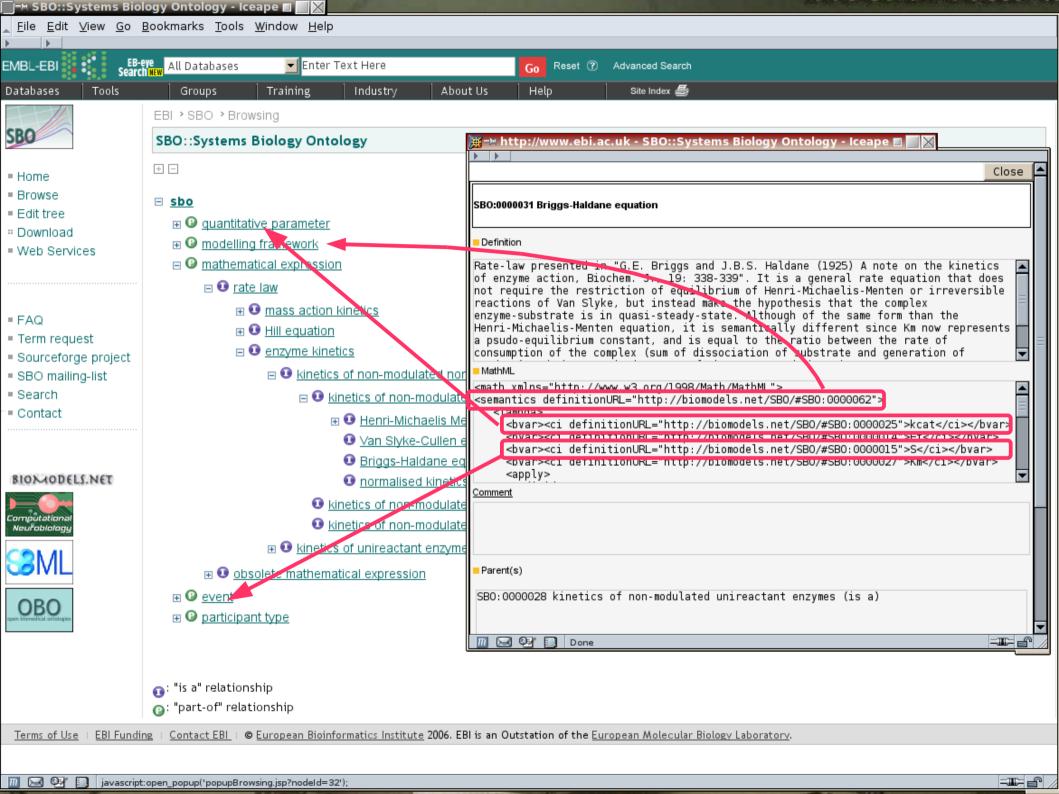


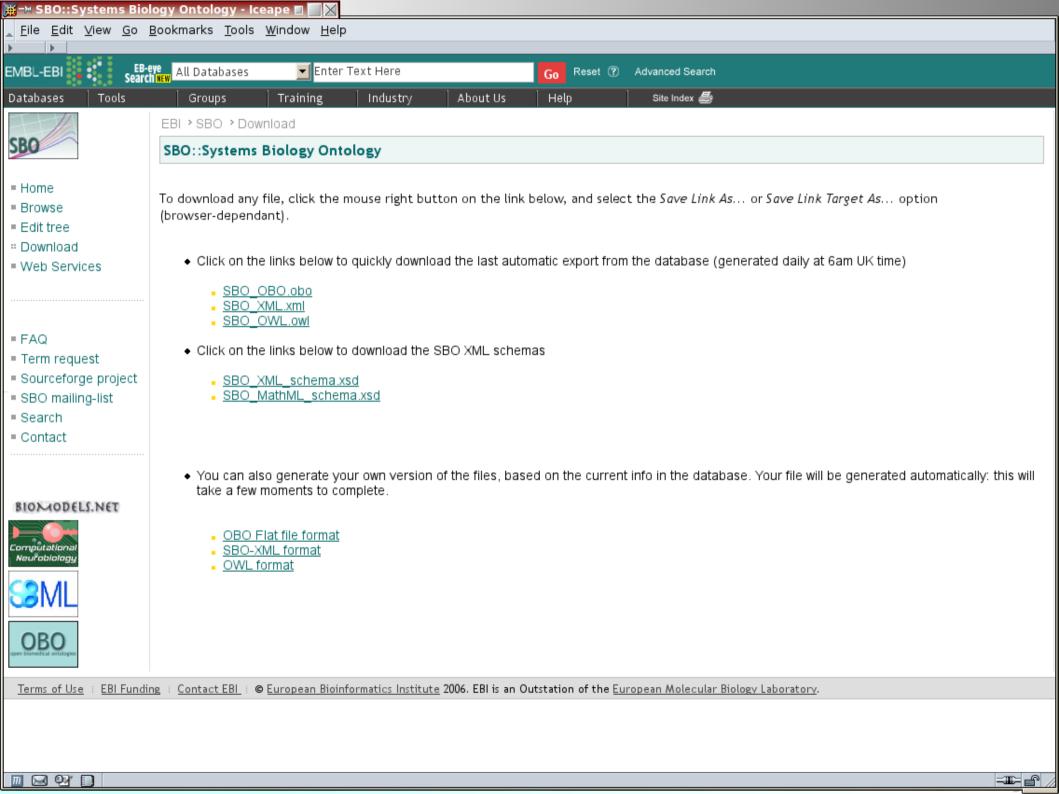


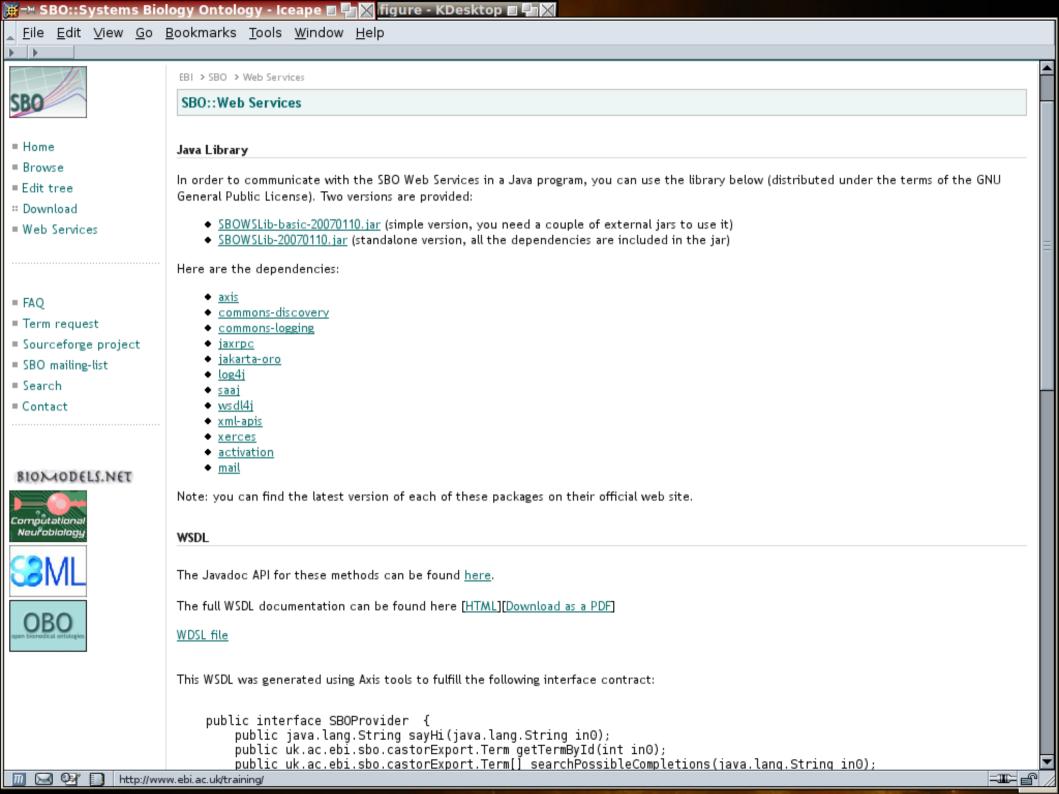
Systems Biology Ontology vocabularies

- Types and roles of reaction participants, including terms like "substrate", "catalyst" etc., but also "macromolecule", or "channel".
- Parameter used in quantitative models. This vocabulary includes terms like "Michaelis constant", "forward unimolecular rate constant"etc. A term may contain a precise mathematical expression stored as a MathML lambda function. The variables refer to other parameters.
- Mathematical expressions. Examples of terms are "mass action kinetics", "Henri-Michaelis-Menten equation" etc. A term may contain a precise mathematical expression stored as a MathML lambda function. The variables refer to the other vocabularies.
- Modelling framework to precise how to interpret the rate-law. E.g. "continuous modelling", "discrete modelling" etc.
- Event type, such as "catalysis" or "addition of a chemical group".



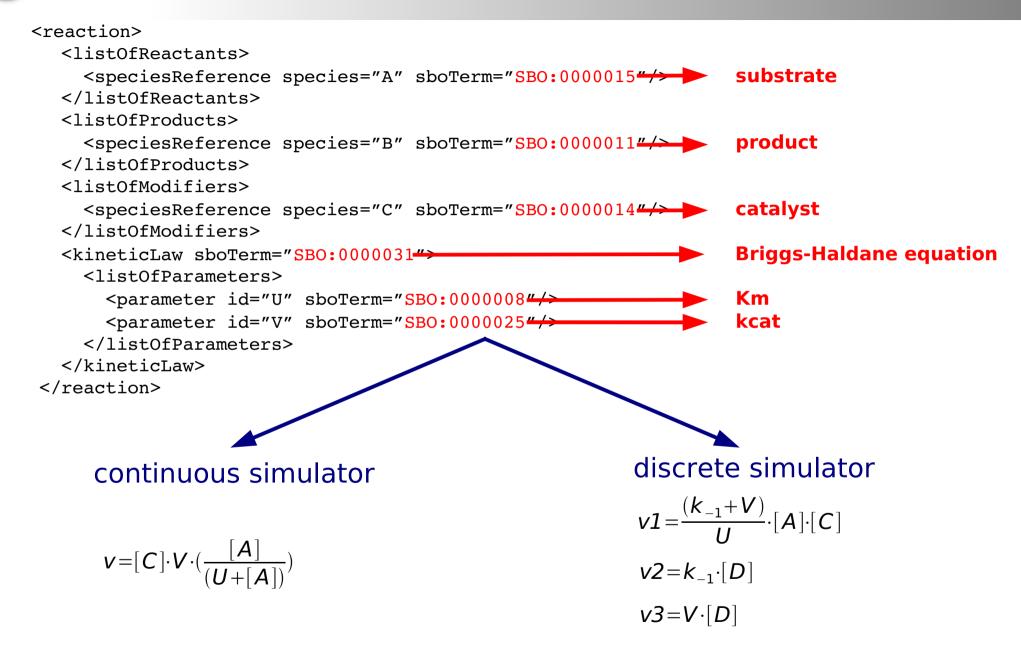








Revealed assumptions







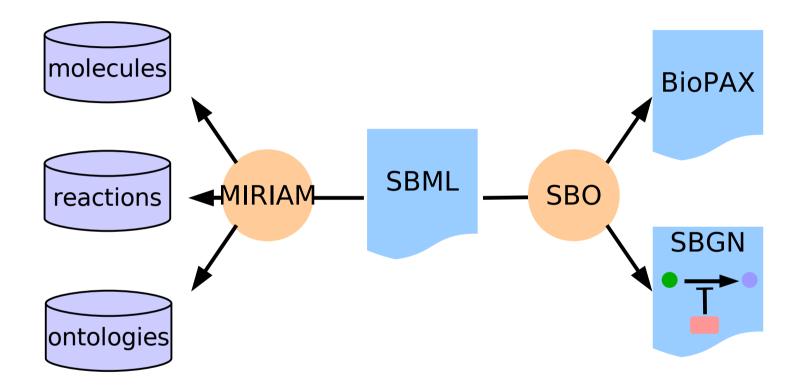
[complete model, with links in species and etc.]

```
<listOfSpecies>
 simple chemical
                                      simple chemical
 macromolecule
 <species id="C" sboTerm="SB0:0000245" />
</listOfSpecies>
<listOfreactions>
    <reaction sboTerm="SBO:0000172">
                                      catalysis
  <listOfReactants>
   </listOfReactants>
  stOfProducts>
   </listOfProducts>
  <listOfModifiers>
   </listOfModifiers>
                                      Briggs-Haldane equation
  <kineticLaw sboTerm="SB0:0000031">>>>
   <listOfParameters>
    <parameter id="V" sboTerm="SBO:0000025"
//>
   </listOfParameters>
  </kineticLaw>
 </reaction>
</listOfreactions>
```













Requirements for a unified model resource

- Neither focussed on a particular biological substrate or process, nor specialised on a given modelling approach
- · Real "searchable" database rather than mere repository
- Models thoroughly verified, structure and results, and annotated
- · International collaboration rather than a one-group effort
- · Freely available and reusable
- Long-term commitment and secure funding





BioModels Database: A Free, Centralized Database of Curated, Published, Quantitative Kinetic Models of Biochemical and Cellular Systems

Le Novère N., Bornstein B., Broicher A., Courtot M., Donizelli M., Dharuri H., Li L., Sauro H., Schilstra M., Shapiro B., Snoep J.L., Hucka M. Nucleic Acids Research, (2006), 34: D689-D691

http://www.ebi.ac.uk/biomodels/



What is BioModels Database?

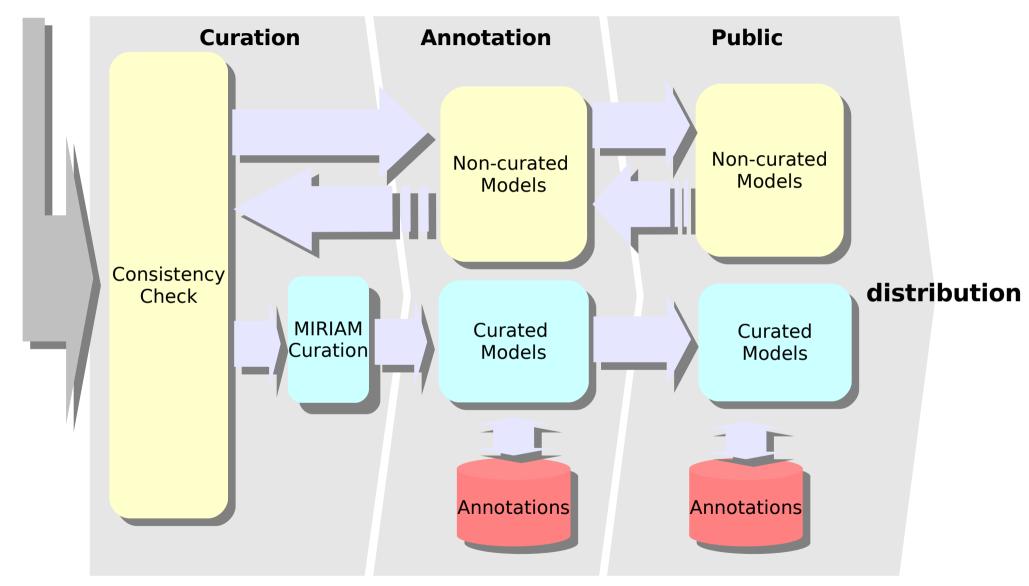
- · Store and serve <u>quantitative</u> models of biomedical interest
- · Only models described in the <u>peer-reviewed</u> scientific literature.
- Models are <u>curated</u>: computer software check the syntax, while human curators check the semantics.
- Models are <u>simulated</u> to check the reference correspondence
- Model components are <u>annotated</u>, to improve identification and retrieval.
- Models are accepted in several formats, and served in several others.



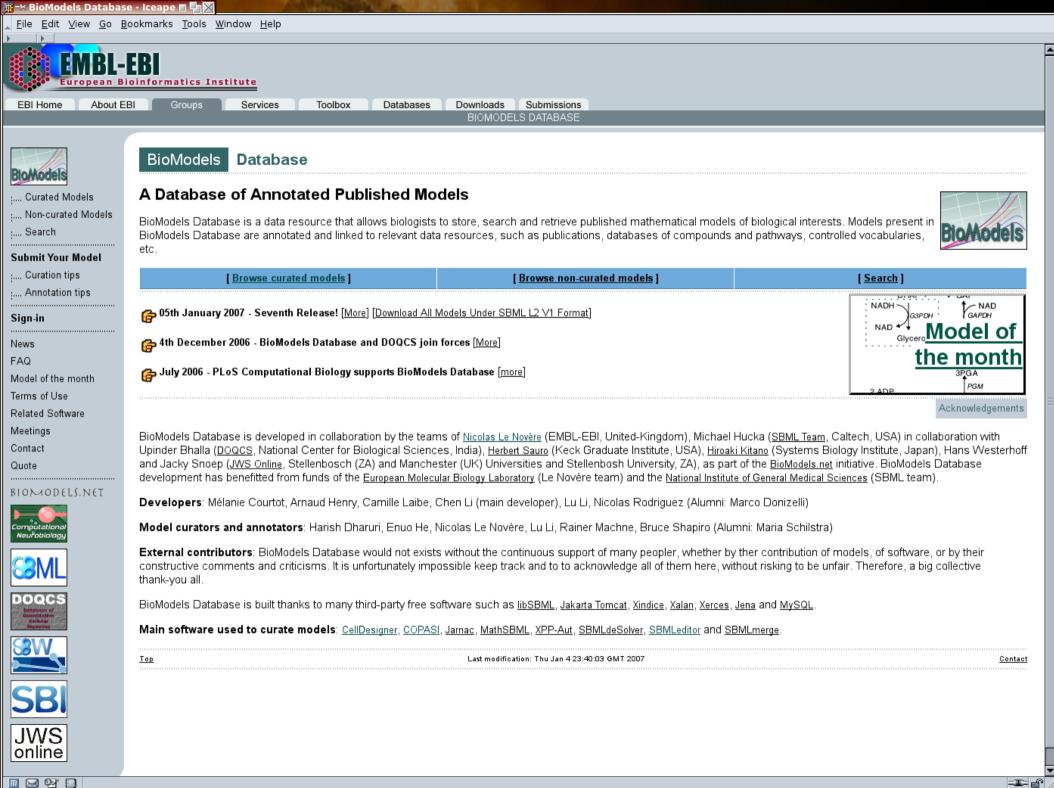


Structure of BioModels Database

Submission







=38 ≥

BIOMODELS DATABASE

BloModels

···· Curated Models

:-- Non-curated Models

:--- Search

Submit Your Model

 $\vdots \cdots \ \mathsf{Curation} \ \mathsf{tips}$

:--- Annotation tips

Sign-in

News

FAQ

Model of the month

Terms of Use

Related Software

Meetings

Contact Quote

BIOMODELS.NET







Search Models

The search totally returned 13 models.

<⇒ <u>New Search</u>

Show 10 Only

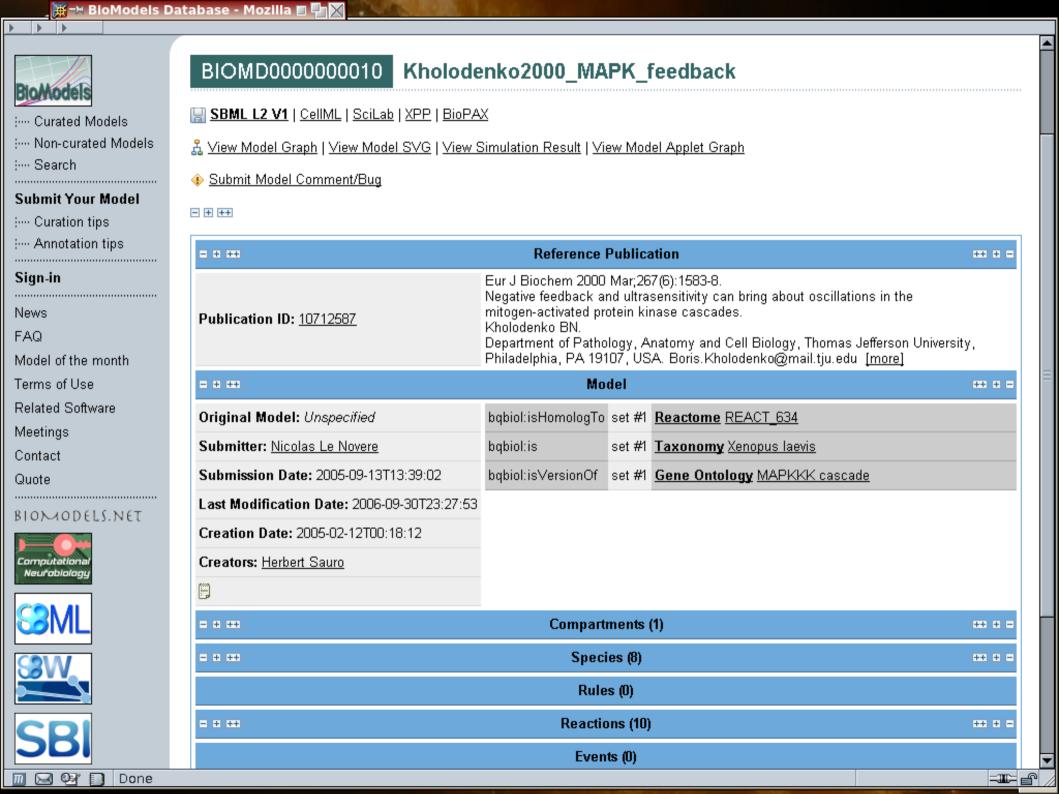
13 Curated Models returned.

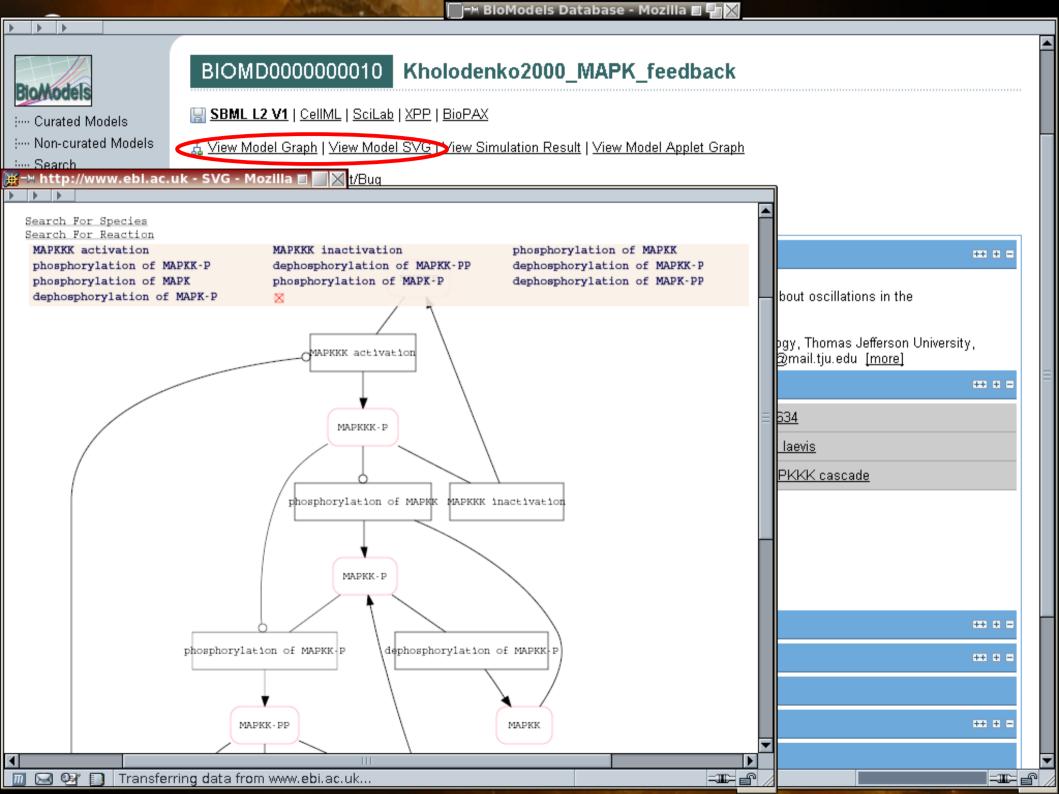
<u>BioModels ID</u> ▼	<u>Name</u>	<u>Publication ID</u>	<u>Last Modified</u>
BIOMD0000000009	Huang1996_MAPK_ultrasens	<u>8816754</u>	2006-09-30T23:18:39
BIOMD0000000010	Kholodenko2000_MAPK_feedback	<u>10712587</u>	2006-09-30T23:27:53
BIOMD0000000011	Levchenko2000_MAPK_noScaffold	<u>10823939</u>	2006-09-15T23:41:42
BIOMD000000014	Levchenko2000_MAPK_Scaffold	<u>10823939</u>	2006-09-18T00:04:02
BIOMD0000000026	Markevich2004_MAPK_orderedElementary	<u>14744999</u>	2006-04-02T18:50:28
BIOMD0000000027	Markevich2004_MAPK_orderedMM	<u>14744999</u>	2006-08-14T13:52:32
BIOMD0000000028	Markevich2004_MAPK_phosphoRandomElementary	<u>14744999</u>	2006-04-02T18:53:13
BIOMD0000000029	Markevich2004_MAPK_phosphoRandomMM	<u>14744999</u>	2006-08-14T13:53:16
BIOMD0000000030	Markevich2005_MAPK_AllRandomElementary	<u>14744999</u>	2006-04-02T18:57:56
BIOMD0000000031	Markevich2004_MAPK_orderedMM2kinases	<u>14744999</u>	2006-04-02T18:58:15
BIOMD000000032	Kofahl2004_pheromone	<u>15300679</u>	2006-08-20T01:25:41
BIOMD000000033	Brown2004_NGF_EGF_signaling	<u>14525003</u>	2006-08-14T13:59:12
BIOMD0000000049	Sasagawa2005_MAPK	<u>15793571</u>	2006-08-24T23:29:11

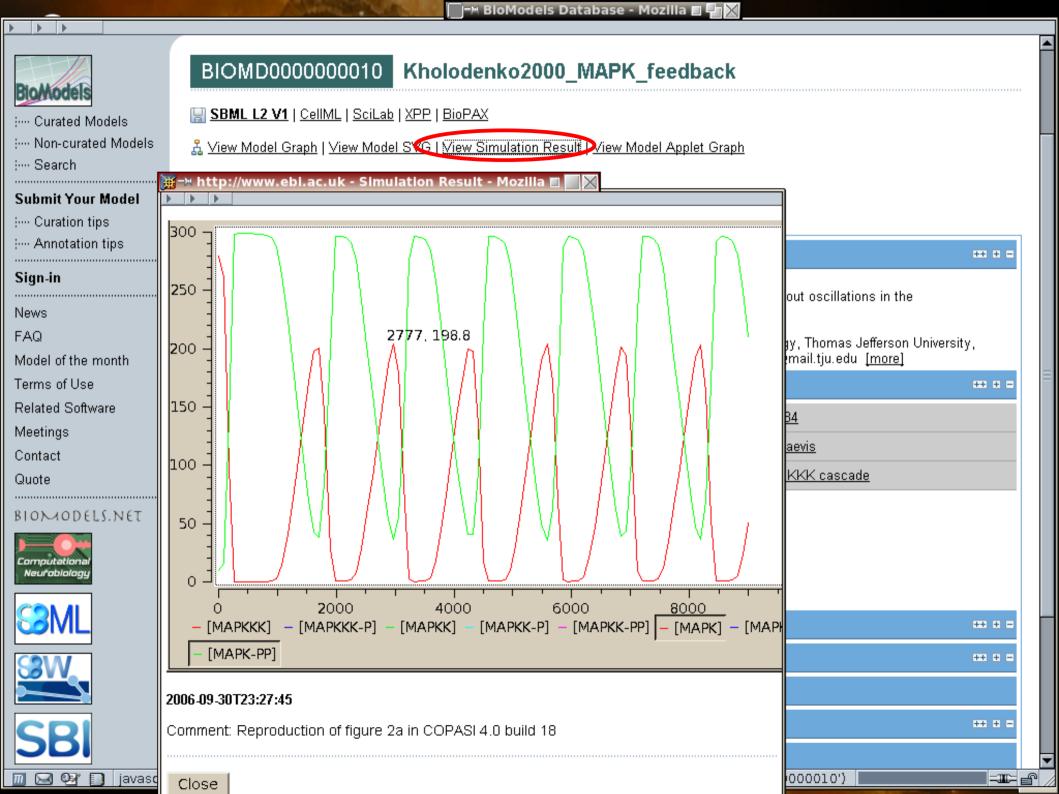
New Search

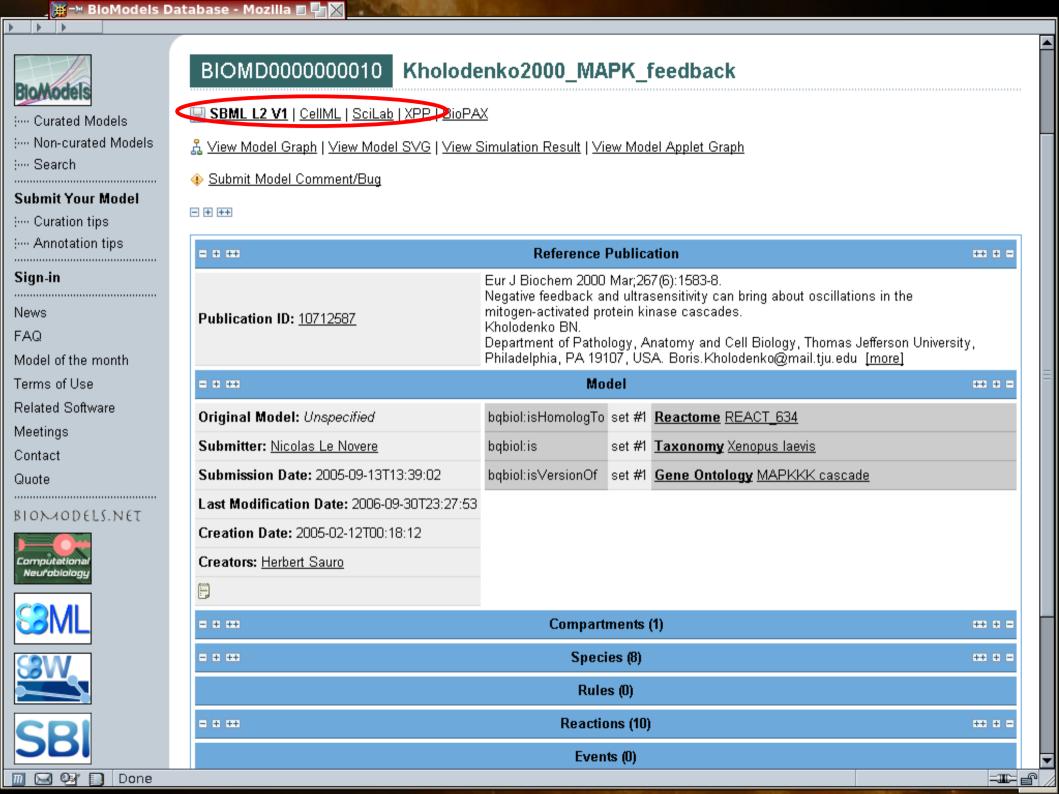


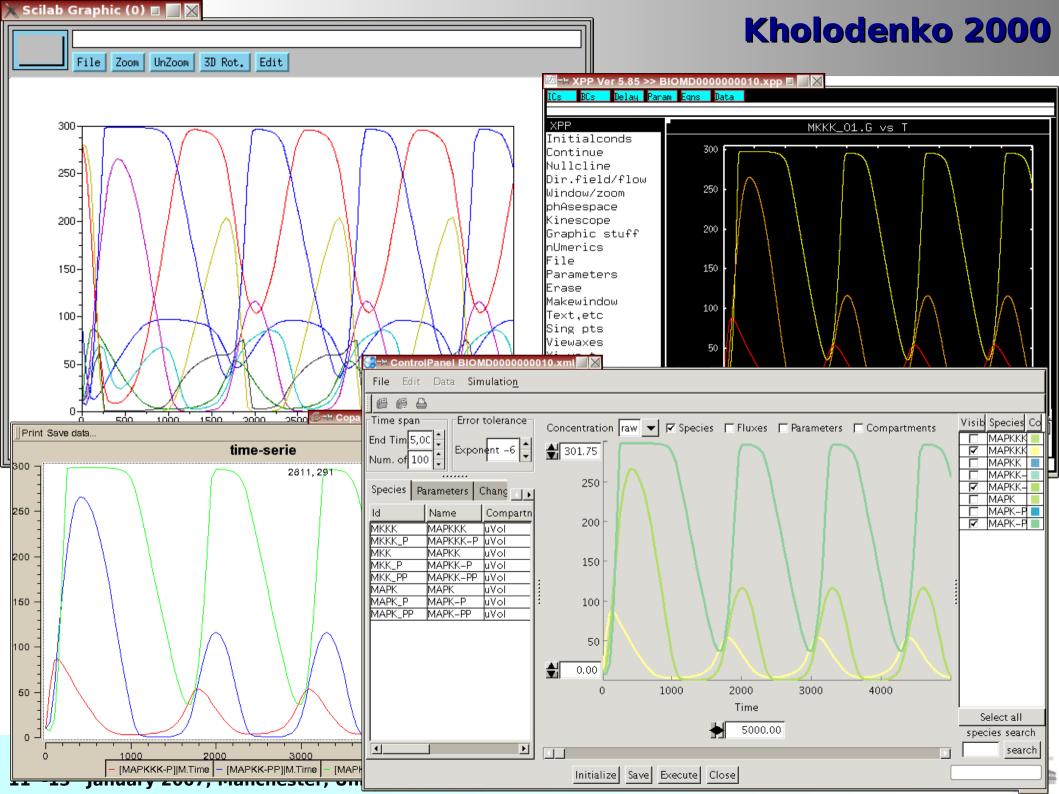














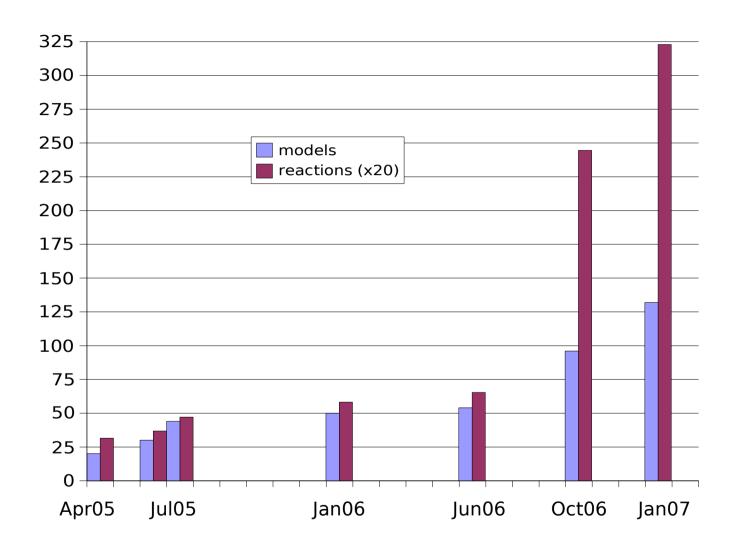
Where are the models coming from

- I) Existing model repositories
- old SBML repository
- JWS Online
- Database Of Quantitative Cell Signalling
- CellML repository
- the Virtual Cell
- II) Individuals
- Members of the SBML community (developers+modellers)
- Authors (prior to grant application, before publication etc.)
- III) Journals (Molecular Systems Biology and PloS Computational Biology advise deposition)
- IV) BioModels DB curators encode new models from literature





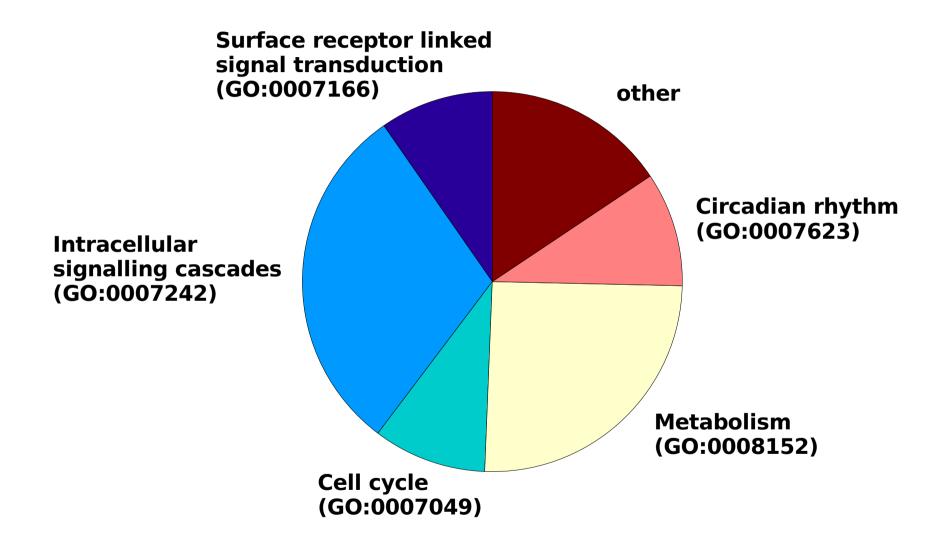








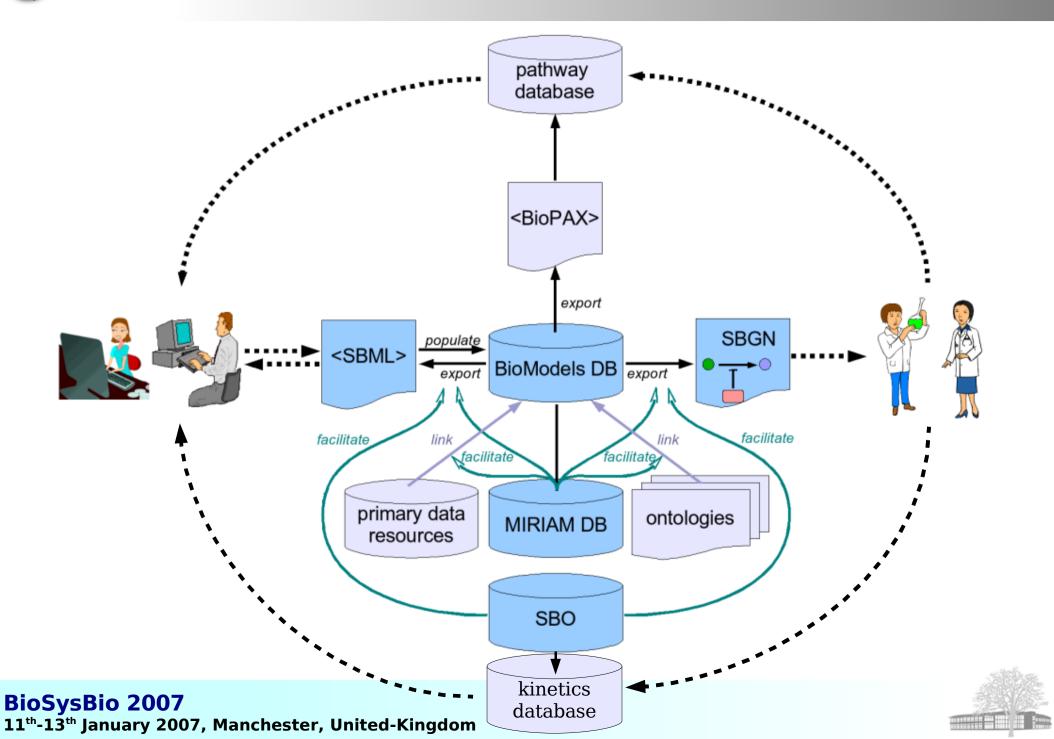








Is the mosaic complete?





The BioModels.net team



Enuo He



Harish Dharuri



Bruce Shapiro



Melanie Stefan

Lu Li

Camille Laibe



Marco Donizelli

Melanie Courtot



Arnaud Henry



Nicolas Le Novère



Michael Hucka







BioSysBio 2007 11th-13th January 2007, Manchester, United-Kingdom





An international collaboration

<u>EBI</u>

- Nicolas Le Novère
- Marco Donizelli
- Chen Li
- Mélanie Courtot
- Lu Li
- Camille Laibe
- Arnaud Henry
- Enuo He
- Nicolas Rodriguez
- Alexander Broicher

· SBML team

- Michael Hucka
- Andrew Finney
- Bruce Shapiro
- Benjamin Borstein
- Maria Schilstra
- Sarah Keating
- Harish Dharuri

NCBS

- Upinder Bhalla
- Harsha Rani

Keck Graduate Institute

Herbert Sauro

Vienna TBI

- Rainer Machne

Systems Biology Institute

- Hiroaki Kitano
- Akira Funahashi

· <u>JWS Online</u>

- Jacky Snoep
- Hans Westerhoff

<u>Journals supporting BioModels Database</u>

- Molecular Systems Biology
- PLoS Computational Biology
- Programs used for curation
 - CellDesigner/SBMLodeSolver
 - COPASI
 - Jarnac/JDesigner
 - MathSBML
 - SBMLeditor
 - XPP-Aut

The community of Systems Biology for their contributions of models and comments.











