

# Sharing enriched computational models, a cornerstone for Integrative Biology

Nicolas Le Novère, EMBL-EBI



## **European Molecular Biology Laboratory**

- EMBL is a basic research institute funded by public research monies from 20 member states and numerous external funding sources
- 1530 staffs, over 60 nationalities



Basic research in molecular biology
Administration
EMBO



Structural biology



Structural biology

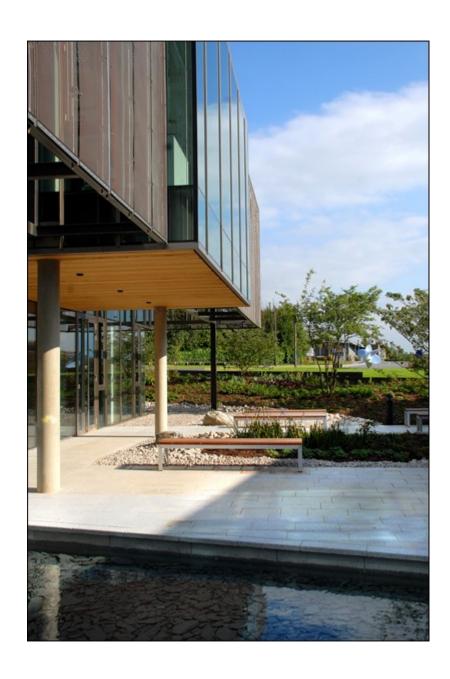


**Bioinformatics** 



Mouse biology

## **EMBL-European Bioinformatics Institute**



Provision of services, research and training in Bioinformatics

From the molecule to the cell, and the genome to the individual

Based on the Wellcome-Trust Genome Campus near Cambridge, UK













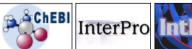












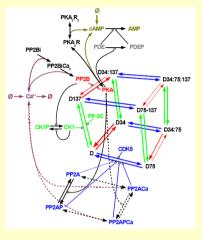


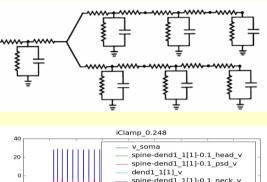


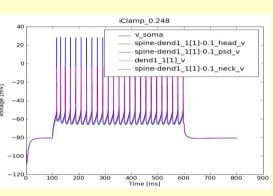


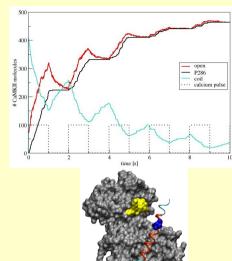
## Themes and projects of the "compneur" group

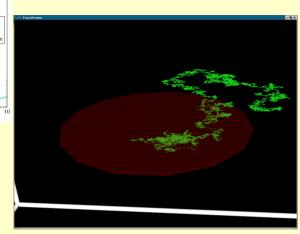
## Computational Neurobiology









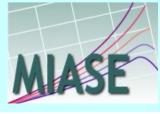


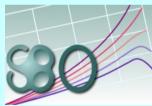


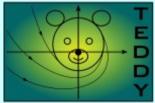


Computational Systems Biology











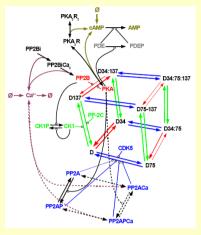


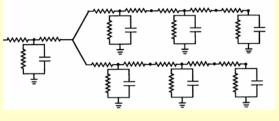


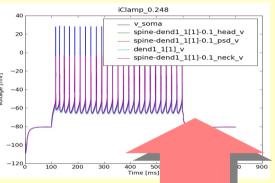


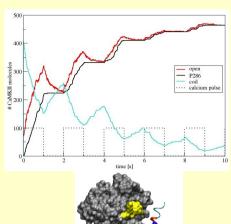
## Themes and projects of the "compneur" group

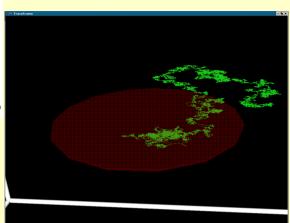
## Computational Neurobiology

















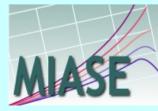






Systems Biology









## What happened to biology at the end of XX<sup>th</sup> century?

Annu. Rev. Genomics Hum. Genet. 2001. 2:343-72 Copyright © 2001 by Annual Reviews. All rights reserved



## A New Approach to Decoding Life:

Systems Biology

Trey Ideker<sup>1,2</sup>, Timothy Galitski<sup>1</sup>, and Leroy Hood<sup>1,2,3,4,5</sup> Institute for Systems Biology<sup>1</sup>, Seattle, Washington 98105; Departments of

Systems of Life

Systems Biology

Basic science

REVIEW

## Systems Biology: A Brief Overview

Hiroaki Kitano

1 MARCH 2002 VOL 295 SCIENCE www.sciencemag.org



## What happened to biology at the end of XX<sup>th</sup> century?

### **RESEARCH** ARTICLE

## Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,<sup>1</sup> John I. Glass,<sup>1</sup> Carole Lartigue,<sup>1</sup> Vladimir N. Noskov,<sup>1</sup> Ray-Yuan Chuang,<sup>1</sup> Mikkel A. Algire,<sup>1</sup> Gwynedd A. Benders,<sup>2</sup> Michael G. Montague,<sup>1</sup> Li Ma,<sup>1</sup> Monzia M. Moodie,<sup>1</sup> Chuck Merryman,<sup>1</sup> Sanjay Vashee,<sup>1</sup> Radha Krishnakumar,<sup>1</sup> Nacyra Assad-Garcia,<sup>1</sup> Cynthia Andrews-Pfannkoch,<sup>1</sup> Evgeniya A. Denisova,<sup>1</sup> Lei Young,<sup>1</sup> Zhi-Qing Qi,<sup>1</sup> Thomas H. Segall-Shapiro,<sup>1</sup> Christopher H. Calvey,<sup>1</sup> Prashanth P. Parmar,<sup>1</sup> Clyde A. Hutchison III,<sup>2</sup> Hamilton O. Smith,<sup>2</sup> J. Craig Venter<sup>1,2</sup>\*

2 JULY 2010 VOL 329 SCIENCE www.sciencemag.org

# Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors

Kazutoshi Takahashi1 and Shinya Yamanaka1,2,\*

Department of Stem Cell Biology, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan

<sup>2</sup> CREST, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan

\*Contact: yamanaka@frontier.kyoto-u.ac.jp DOI 10.1016/j.cell.2006.07.024

Cell 126, 663-676, August 25, 2006 @2006 Elsevier Inc. 663



## EXTREME GENETIC ENGINEERING

An Introduction to Synthetic Biology



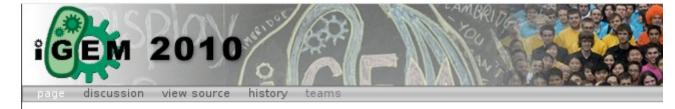


## A synthetic oscillatory network of transcriptional regulators

Michael B. Elowitz & Stanislas Leibler

Departments of Molecular Biology and Physics, Princeton University, Princeton New Jersey 08544, USA

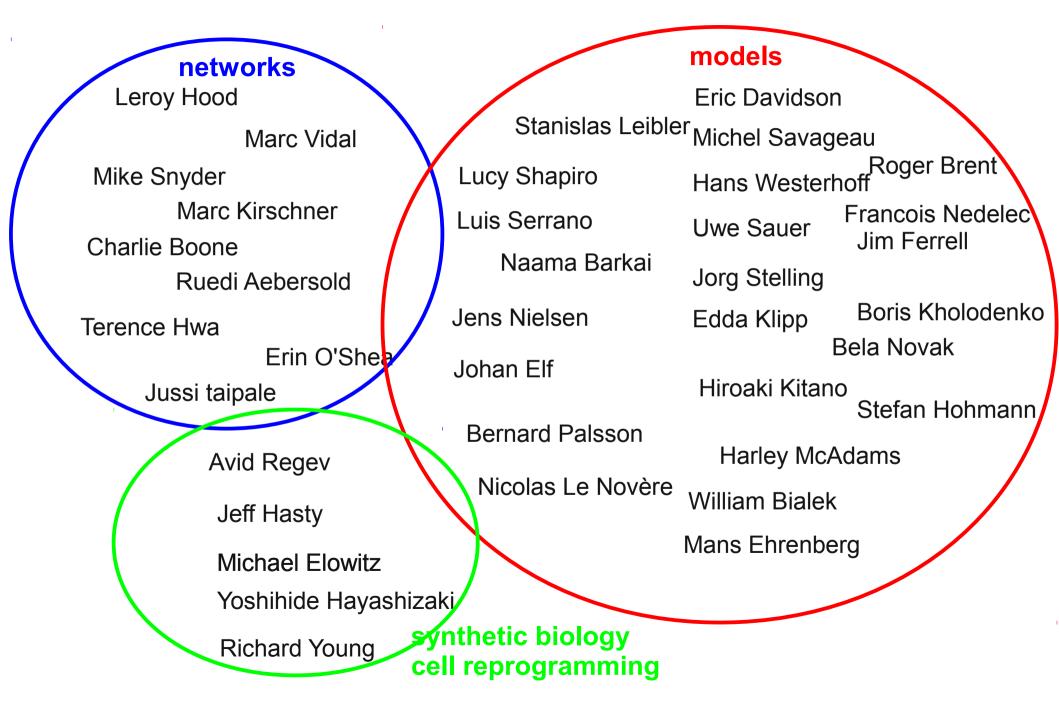
NATURE | VOL 403 | 20 JANUARY 2000 | www.nature.com



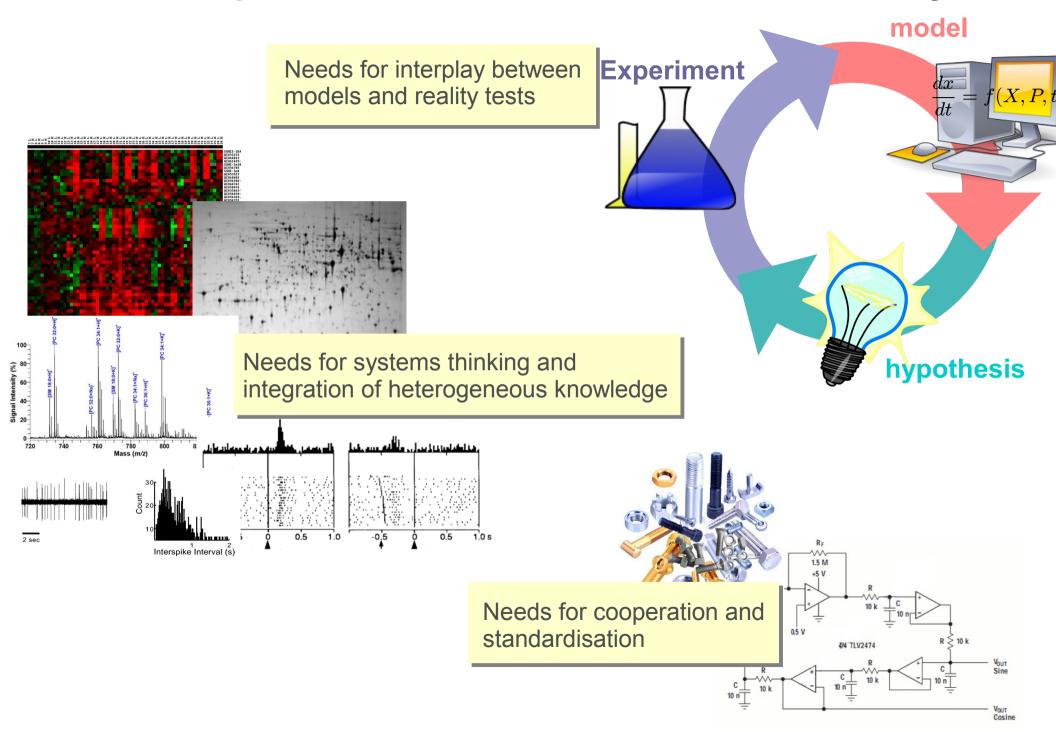
#### About

The International Genetically Engineered Machine competition (iGEM) is Biology competition. Student teams are given a kit of biological parts at the beginnin Standard Biological Parts. Working at their own schools over the summer, they use t

## Nobel Symposium on Systems Biology (June 2009)



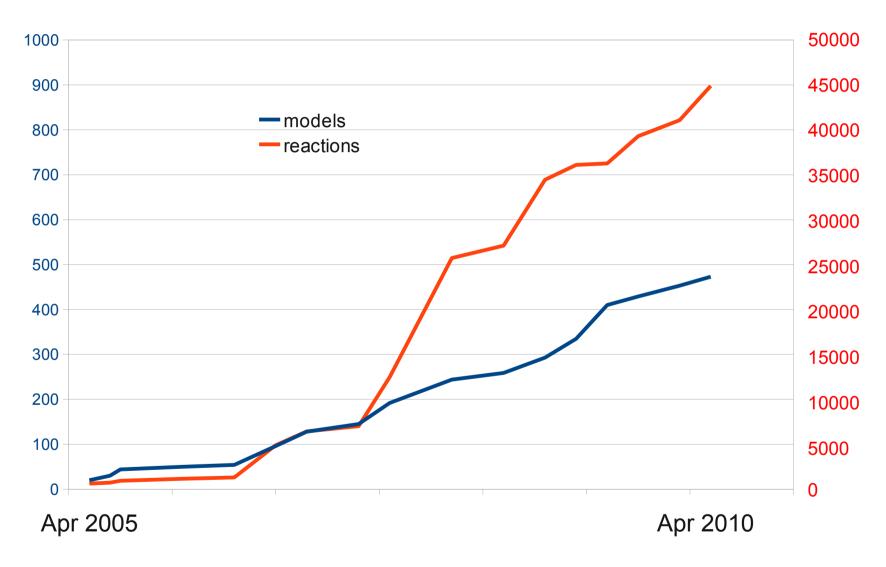
## Consequences of this revolution on our activity



## Computational modelling for biology and medicine left the niches in the last decades

- Metabolic networks (Herrgård et al. A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat Biotechnol* 2008)
- Signalling pathways (Bray et al. Receptor clustering as a cellular mechanism to control sensitivity. Nature 1998; Bhalla ad Iyengar. Emergent properties of signaling pathways. Science 1998, Schoeberl et al. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. Nat Biotechnol 2002; Nelson et al. Oscillations in NF-kB Signaling Control the Dynamics of Gene Expression. Science 2004; Ashall et al. Pulsatile Stimulation Determines Timing and Specificity of NF-kappa B-Dependent Transcription. Science 2009)
- Gene regulatory networks (McAdams and Shapiro. Circuit simulation of genetic networks. Science 1995; Yue et al. Genomic cis-regulatory logic: Experimental and computational analysis of a sea urchin gene. Science 1998; Von Dassow et al. The segment polarity network is a robust developmental module. Nature 2000)
- Pharmacokinetic/dynamic models (Labrijn et al. Therapeutic IgG4 antibodies engage in Fab-arm exchange with endogenous human IgG4 in vivo. *Nat Biotechnol* 2009)
- Physiological models (Noble. Modeling the heart from genes to cells to the whole organ. Science 2002; Izhikevich and Edelman. Large-scale model of mammalian thalamocortical systems. PNAS 2008)
- Infectious diseases (Perelson et al. HIV-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time. Science 1996; Neumann et al. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. Science 1998)

## Computational models on the rise



Growth of BioModels Database between its creation and release 17

BioModels Home

EMBL-EBI

Databases

Models

Sign in

Search

Support

About BioModels

#### BioModels Database - A Database of Annotated Published Models

Submit

BioModels Database is a data resource that allows biologists to store, search and retrieve published mathematical models of biological interests. Models present in BioModels Database are annotated and linked to relevant data resources, such as publications, databases of compounds and pathways, controlled vocabularies, etc.

Go to the model Advanced search



#### Browse models

- Curated models (249)
- Browse models using GO
- Non-curated models (224)

Simulate in JWS Online

Submit a model

Mirror at California Institute of Technology http://biomodels.caltech.edu

BioModels AT SourceForge http://sourceforge.net/projects/biomodels/

Web Services http://www.ebi.ac.uk/biomodels-main/webservices

Download archived models http://www.ebi.ac.uk/biomodels/models-main/tars/

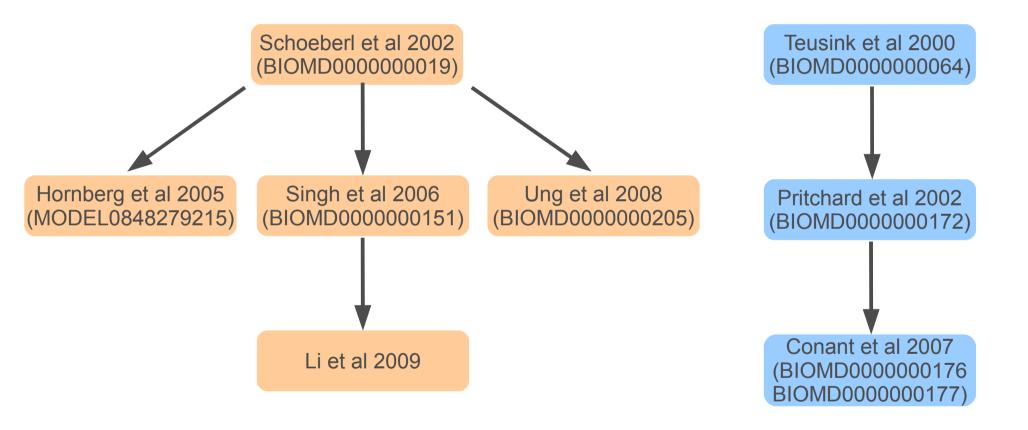
#### Model of the month July, 2010 The activity of the transcription factor NF-kB, is largely controlled by three IkB isoforms (IkBa, IkBB, IkBE) which upon binding to NF-kB. prevents its association with DNA causing its localization to the cytoplasm. A computational model that describes the temporal control of NF-kB activation by IkB proteins is described here.

Read more...



N. Le Novère, B. Bornstein, A. Broicher, M. Courtot, M. Donizelli, H. Dharuri, L. Li, H. Sauro, M. Schilstra, B. Shapiro, J.L. Snoep, M. Hucka. BioModels Database: A Free, Centralized Database of Curated, Published, Quantitative Kinetic Models of Biochemical and Cellular Systems. (2006) Nucleic Acids Research, 34: D689-D691.

## Direct model re-use: EGFR signalling and glycolysis



## Models at the core of integrative systems bio/physio/neurobiology

- As any kind of output from scientific research, models must be available to the scientific community: Results based on a model that is not distributed cannot be verified and falsified. This is not valid and useful science
- Computational models must be exchanged: computer storable and readable
- Computational models must be related to relevant experimental datasets: expressive relationships and robust links
- Computational models must be amenable to many different analyses: standard API
- Computational models must allow modification, split, merge: Encoded according a suitable design

## Models at the core of integrative systems bio/physio/neurobiology

- As any kind of output from scientific research, models must be available to the scientific community: Results based on a model that is not distributed cannot be verified and falsified. This is not valid and useful science
- Computational models must be exchanged: computer storable and readable
- Computational models must be related to relevant experimental datasets: expressive relationships and robust links
- Computational models must be amenable to many different analyses: standard API
- Computational models must allow modification, split, merge: Encoded according a suitable design

→ Standardisation is unavoidable!

## A "complete" (?) mosaic of standards for Computational Systems Biology models

	Models	Simulation	Results
Minimal requirements	MIRIAM	MIASE	
Data-models	SIML SIGN	SED ML	SBRML
Ontologies	<b>S30</b>	KISAO	TEDDY

## **Model description**

	Models	Simulation	Results
Minimal requirements	MIRIAM	MIASE	
Data-mode		ML	SBRML
Ontologies	<b>S30</b>	KISAO	TEDDY



### The Systems Biology Markup Language

🕊 News Documents Downloads Forums Facilities Community Events About 氝 📴

Welcome to the portal for the Systems Biology Markup Language (SBML), a computer-readable format for representing models of biological processes. SBML is suitable for models of metabolism, cell signaling, and other processes, and has been evolving since 2000 thanks to an international community of researchers.



#### For the curious

What is SBML? Read our introduction, then perhaps browse the mailing lists to get a sense for what's going on with SBML today.



#### For modelers

Looking for software that supports SBML? Our software guide lists over 180 systems today. Are you instead looking for models? Visit the BioModels Database &, where you can find hundreds!



#### For software developers

Interested in supporting SBML in your software? Read our basic introduction and then the SBML specifications to understand how to use SBML. After that, you may want to look at libSBML.

No matter how you use SBML, we invite you to sign up for news updates either through our RSS feed, our Twitter feed &, or one of the mailing lists, and get involved with community efforts to help keep improving SBML. You can also

#### SBML News

#### SBMLToolbox 3.1.2!

(27 Apr. '10) Another minor bug-fix release of our free MATLAB toolbox for SBML is now available.

#### SBMLToolbox 3.1.1!

(12 Apr.'10) A minor bug-fix release of our free MATLAB toolbox for SBML is now available.

#### Older news ...

#### Community News

#### Cain 1.4 released! 🚱

(31 Mav.'10) Cain № is a stochastic simulator with highly efficient implementations of many methods.

#### SBW 2.8.0 released!

(10 May.'10) SBW is a component-based application framework. The update improves support for simulation of SBML models and Windows 7 support...

#### CBO workshop @ ICSB @

(2 May.'10) There will be a



Hucka M., Bolouri H., Finney A., Sauro H.M., Doyle J.C., Kitano H.Arkin A.P., Bornstein B.J., Bray D., Cornish-Bowden A., Cuellar A.A., Dronov S., Ginkel M., Gor V., Goryanin I.I., Hedley W.J., Hodgman T.C., Hunter P.J., Juty N.S., Kasberger J.L., Kremling A., Kummer U., Le Novère N., Loew L.M., Lucio D., Mendes P., Mjolsness E.D., Nakayama Y., Nelson M.R., Nielsen P.F., Sakurada T., Schaff J.C., Shapiro B.E., Shimizu T.S., Spence H.D., Stelling J., Takahashi K., Tomita M., Wagner J., Wang J.et al (2003). The Systems Biology Markup Language (SBML): A Medium for Representation and Exchange of Biochemical Network Models. *Bioinformatics*, 19: 524-531.



### What is SBML?

The Systems Biology Markup Language is a way to **exchange and reuse** (and hopefully **interface**) descriptions of quantitative models in "Systems Biology", in fact mostly well-stirred chemical kinetics so far.



It is not a procedural language.

It is not a programming language.

It is not a format for specific software configuration files (only 3 of the 7 SBML founding software are still maintained today).

Development philosophy: Start small, get good support, extend.

SBML itself re-uses other standards: XML, MathML, XHTML, RDF, existing ontologies.

It is supported by a community large, diverse, active and evolving.



### Global structure of a SBML file

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="3" version="1".</pre>
      xmlns="http://www.sbml.org/sbml/level3/version1/core">
  <model>
    <listOfFunctionDefinitions> </-- --> </listOfFunctionDefinitions>
    <listOfUnitDefinitions> </-- --> </listOfUnitDefinitions>
    <list0fCompartments> <!-- --> </list0fCompartments>
    <list0fSpecies> </-- --> </list0fSpecies>
    <list0fParameters> </-- --> </list0fParameters>
    <list0fInitialAssignments> </-- --> </list0fInitialAssignments>
    <list0fRules> </-- --> </list0fRules>
    <list0fConstraints> </-- --> </list0fConstraints>
    <listOfReactions> </-- --> </listOfReactions>
    <list0fEvents> </-- --> </list0fEvents>
  </model>
</sbml>
```



### Global structure of a SBML file

```
<?xml version="1.0" encoding="UTF-8"?>
                                                             variables
<sbml level="3" version="1".</pre>
     xmlns="http://www.sbml.org/sbml/level3/version1/core">
 <model>
   <listOfFunctionDefinitions> </-- --> </listOfFunctionDefinitions>
   <listOfUnitDefinitions> </-- --> </listOfUnitDefinitions>
   <list0fCompartments> <!-- --> </list0fCompartments>
   <list0fSpecies> </-- --> </list0fSpecies>
   <list0fParameters> </-- --> </list0fParameters>
   <list0fInitialAssignments> </-- --> </list0fInitialAssignments>
   <list0fConstraints> </-- --> </list0fConstraints>
   <listOfReactions> </-- --> </listOfReactions>
   <list0fEvents> </-- --> </list0fEvents>
 </model>
</sbml>
```

relationships



```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="4" xmlns="http://www.sbml.org/sbml/level2/version4">
 <model>
   <compartment id="cell" />
   </list0fCompartments>
   <species id="A" compartment="cell" initialConcentration="1"/>
    <species id="B" compartment="cell" initialConcentration="0"/>
   </listOfSpecies>
   <parameter id="kon" value="1"/>
   <reaction>
      <speciesReference species="A" />
      Ist0fProducts>
        <speciesReference species="B" />
      <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>
   </list0fReactions>
 </model>
</sbml>
```



```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="4" xmlns="http://www.sbml.org/sbml/level2/version4">
 <model>
   <compartment id="cell" />
   </list0fCompartments>
   <species id="A" compartment="cell" initialConcentration="1"/>
    <species id="B" compartment="cell" initialConcentration="0"/>
   </listOfSpecies>
   <parameter id="kon" value="1"/>
   <reaction>
      <speciesReference species="A" />
      Ist0fProducts>
        <speciesReference species="B" />
      <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>
   </list0fReactions>
 </model>
</sbml>
```



```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="4" xmlns="http://www.sbml.org/sbml/level2/version4">
 <model>
   <compartment id="cell" />
   <species id="A" compartment="cell" initialConcentration="1"/>
    <species id="B" compartment="cell" initialConcentration="0"/>
   </list0fSpecies>
   <parameter id="kon" value="1"/>
   <reaction>
      <speciesReference species="A" />
      Ist0fProducts>
       <speciesReference species="B" />
      <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>
   </list0fReactions>
 </model>
</sbml>
```



```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="4" xmlns="http://www.sbml.org/sbml/level2/version4">
 <model>
   <compartment id="cell" />
   </list0fCompartments>
   <species id="A" compartment="cell" initialConcentration="1"/>
     <species id="B" compartment="cell" initialConcentration="0"/>
   </listOfSpecies>
   <parameter id="kon" value="1"/>
   </listOfParameters>
   <reaction>
      <speciesReference species="A" />
      Ist0fProducts>
        <speciesReference species="B" />
      <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>
   </list0fReactions>
 </model>
</sbml>
```



```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="4" xmlns="http://www.sbml.org/sbml/level2/version4">
 <model>
   <compartment id="cell" />
   </list0fCompartments>
   <species id="A" compartment="cell" initialConcentration="1"/>
     <species id="B" compartment="cell" initialConcentration="0"/>
   </listOfSpecies>
   <parameter id="kon" value="1"/>
   <reaction>
      <speciesReference species="A" />
      </listOfReactants>
      <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>
   </list0fReactions>
 </model>
</sbml>
```



```
<?xml version="1.0" encoding="UTF-8"?>
    <sbml level="2" version="4" xmlns="http://www.sbml.org/sbml/level2/version4">
      <model>
       <compartment id="cell" />
       </list0fCompartments>
       <species id="A" compartment="cell" initialConcentration="1"/>
         <species id="B" compartment="cell" initialConcentration="0"/>
       </listOfSpecies>
       <parameter id="kon" value="1"/>
       IistOfReactions>
         <reaction>
          <speciesReference species="A" />
          <speciesReference species="B" />
          <kineticLaw>
            <math xmlns="http://www.w3.org/1998/Math/MathML">
             <apply>
               <times />
               <ci>kon</ci>
MathML
               <ci>A</ci>
               <ci>ci>cell</ci>
             </apply>
            </kineticLaw>
         </reaction>
       </model>
    </sbml>
```



## A more realistic example ...

```
<species ·</pre>
           id="A".
           name="a-tubulin"
           compartment="cell"
           initial Amount="1000"
           substanceUnits="item"
           hasOnlySubstanceUnits="true"
           boundaryCondition="true"
           constant="false"
           charge="0"
           metaid="PX"
                                               biological semantics
           sboTerm="SB0:0000245" >
         <notes>
           <body xmlns="http://www.w3.org/1999/xhtml">
XHTML
              One of the components of a microtubule
           </body>
         </notes>
         <annotation>
           <rdf:RDF.
                xmlns:bqbiol="http://biomodels.net/biology-qualifiers/".
                xmlns:bqmodel="http://biomodels.net/model-qualifiers/"
                xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
              <rdf:Description rdf:about="#PX">
                <br/>dpiol:is>
RDF
                  <rdf:Bag>
                    <rdf:li rdf:resource="urn:miriam:uniprot:P68370"/>
                    <rdf:li rdf:resource="urn:miriam:obo.go:G0%3A0045298"/>
                  </rdf:Bag>
                </bgbiol:is>
              </rdf:Description>
           </rdf:RDF>
         </annotation>
       </species>
```



## 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 or repolarisation;

A species is an entity participating to a reaction, **not always** a **chemical** entity:

It can be a molecule

It can be a cell

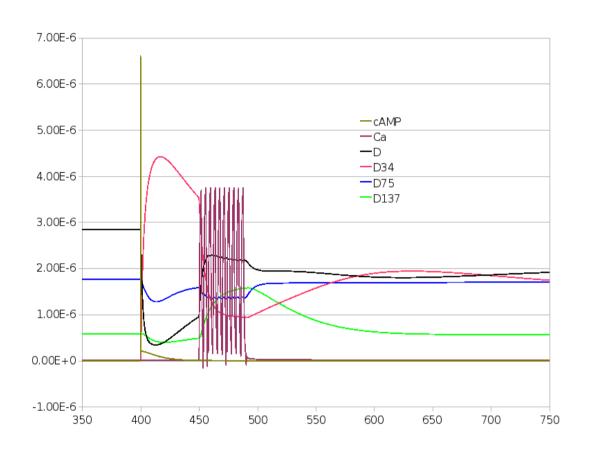
It can be an organ

It can be an organism

→ SBML is about process descriptions



## Model of signalling pathways

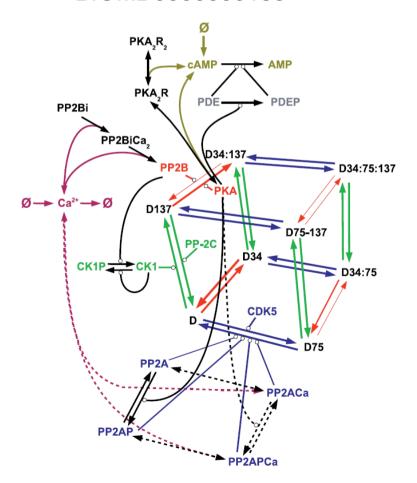


reaction:

$$v_{on1} = k_{on1} \times [D] \times [CDK5] \times Vol$$

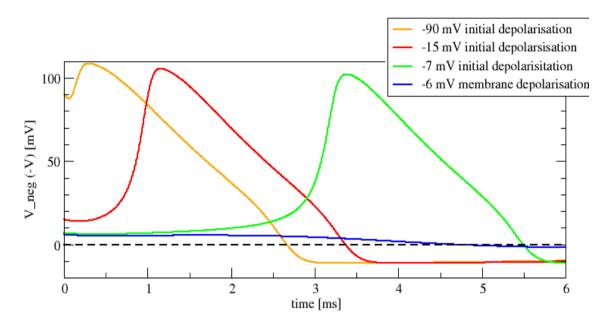
Fernandez et al. DARPP-32 is a robust integrator of dopamine and glutamate signals *PLoS Comput Biol* (2006) 2: e176.

#### BIOMD000000153





### Conductance-based model



Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* (1952) 117:500-544.

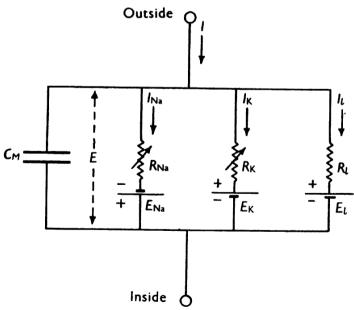
BIOMD000000020

rate rule:

$$\frac{dv}{dt} = \frac{I - (i_{Na} + i_K + i_L)}{C_m}$$

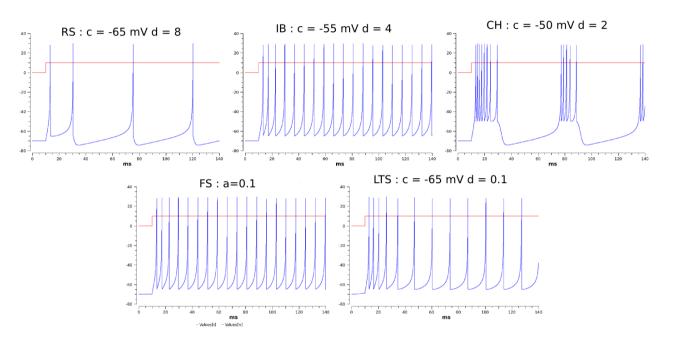
assignment rule:

$$i_{Na} = g_{Na} \times m^3 \times h \times (V - E_{Na})$$





## Single-compartment neurons



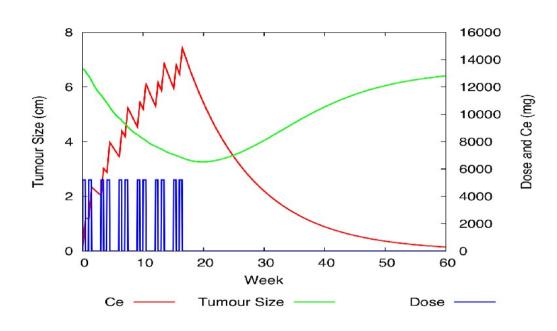
Izhikevich EM. Simple model of spiking neurons. *IEEE Trans Neural Netw* (2003) 14(6):1569-1572.

BIOMD000000127

rate rule: 
$$\frac{dv}{dt} = 0.04^2 + 5 \times V + 140 - U + i$$

event: when 
$$v > V_{thresh} \left\{ egin{aligned} v = c \\ U = U + d \end{aligned} \right.$$

## Pharmacokinetic/dynamic model



Tham et al (2008) A pharmacodynamic model for the time course of tumor shrinkage by gemcitabine + carboplatin in non-small cell lung cancer patients.

Clin Cancer Res. 2008 14(13): 4213-8.

BIOMD000000234

#### rate rule:

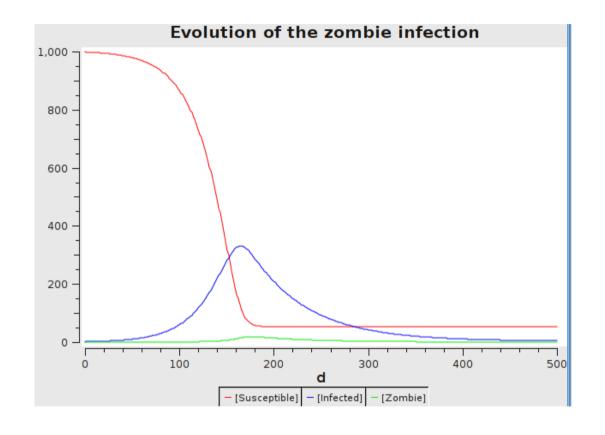
$$\frac{Size}{dt} = (Rate_{in} \times \textit{Effect} - K_{over} \times Size) \times Size$$

### assignment rule:

$$Effect = 1 - \frac{E_{max} - Ce}{Amt_{50} + Ce}$$

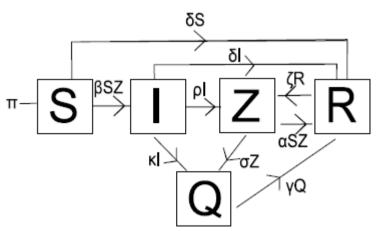


## **Spread of infection diseases ...**



Munz P et al. When zombies attack!: Mathematical modelling of an outbreak of zombie infection. in "Infectious Disease Modelling Research Progress", (2009)133-150

MODEL1008060001





### Difference between SBML L1, L2 and L3

function definitions function definitions predefined functions all math in MathML proprietary infix all math in MathMI math notation ■ no reserved namespaces for ■ no reserved namespaces for reserved namespaces for annotations annotation annotations controlled no controlled controlled RDF annotation annotation RDF annotation discrete events no discrete events discrete events monolithic monolithic modular default values default values no default values

Progressive simplification, generalisation and externalisation



## SBML Level 3 packages

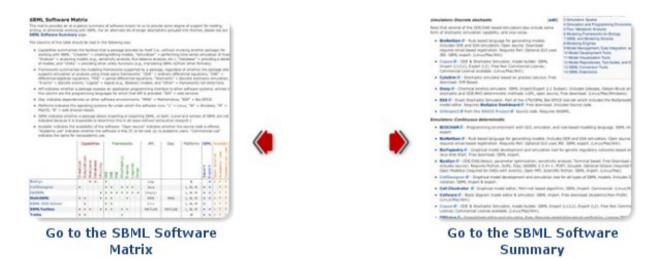
- Core package Release candidate
- Graph Layout specification finalised
- Complex species specification finalised
- Groups specification finalised
- Model composition specification under discussion
- Qualitative models specification under discussion
- Distributions and ranges specification under discussion
- Graph rendering specification proposed
- Arrays and sets specifications proposed
- Geometry specification proposed
- Spatial diffusion specification proposed
- Dynamic structures needed

#### SBML Software Guide

The following summarize all SBML-compatible systems known to us. The *matrix* provides an at-a-glance summary, whereas the *summary* provides longer descriptions of each software or project grouped by themes.

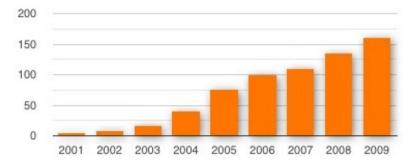
Number of software packages listed in the matrix today: 188

Please use the survey form of to notify us about additions and suggestions.



#### Historical trend

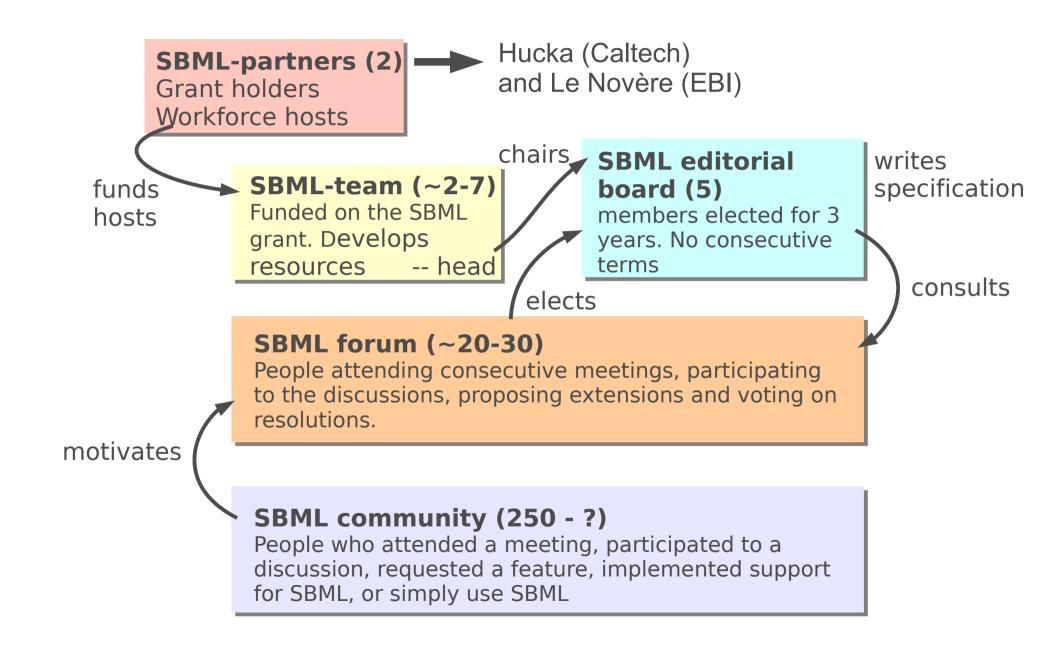
The following graph shows the total number of known SBML-compatible software packages each year, as counted by the SBML Team. The counts shown are for approximately the middle of each year.



(Note: the flat period in 2007 is an artifact of inadequate record keeping rather than a lull in SBML software development.)

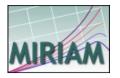


### **Current structure of the SBML community**



#### **Model semantics**

	Models	Simulation	Results
Minimal requiremer	MIRIAM	MASE	
Data-models	SIML SIGN	SED ML	SBRML
Ontologies	<b>S30</b>	KISAO	TEDDY





\_computational BIOLOGY

#### PERSPECTIVE

# Minimum information requested in the annotation of biochemical models (MIRIAM)

Nicolas Le Novère<sup>1,15</sup>, Andrew Finney<sup>2,15</sup>, Michael Hucka<sup>3</sup>, Upinder S Bhalla<sup>4</sup>, Fabien Campagne<sup>5</sup>, Julio Collado-Vides<sup>6</sup>, Edmund J Crampin<sup>7</sup>, Matt Halstead<sup>7</sup>, Edda Klipp<sup>8</sup>, Pedro Mendes<sup>9</sup>, Poul Nielsen<sup>7</sup>, Herbert Sauro<sup>10</sup>, Bruce Shapiro<sup>11</sup>, Jacky L Snoep<sup>12</sup>, Hugh D Spence<sup>13</sup> & Barry L Wanner<sup>14</sup>

Most of the published quantitative models in biology are lost for the community because they are either not made available or they are insufficiently characterized to allow them to be reused. The lack of a standard description format, lack of stringent reviewing and authors' carelessness are the main causes for incomplete model descriptions. With today's increased interest in detailed biochemical models, it is necessary to define a minimum quality standard for the encoding of those models. We propose a set of rules for curating quantitative models of biological systems. These rules define procedures for encoding and annotating models represented in machine-readable form. We believe their

During the genomic era we have witnessed a vast increase in availability of large amounts of quantitative data. This is motivating a shift in the focus of molecular and cellular research from qualitative descriptions of biochemical interactions towards the quantification of such interactions and their dynamics. One of the tenets of systems biology is the use of quantitative models (see Box 1 for definitions) as a mechanism for capturing precise hypotheses and making predictions <sup>1,2</sup>. Many specialized models exist that attempt to explain aspects of the cellular machinery. However, as has happened with other types of biological information, such as sequences, macromolecular structures or

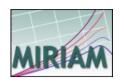
ublishing Group http://www.nature.com/naturebiotechnology



#### **MIRIAM** compliance

#### Models must:

- be encoded in a public machine-readable format
- be clearly linked to a single reference description
- reflect the structure of the biological processes described in the reference paper (list of reactions etc.)
- be instantiable in a simulation (possess initial conditions etc.)
- be able to reproduce the results given in the reference paper
- contain creator's contact details
- annotation to unambiguously identify each model constituent



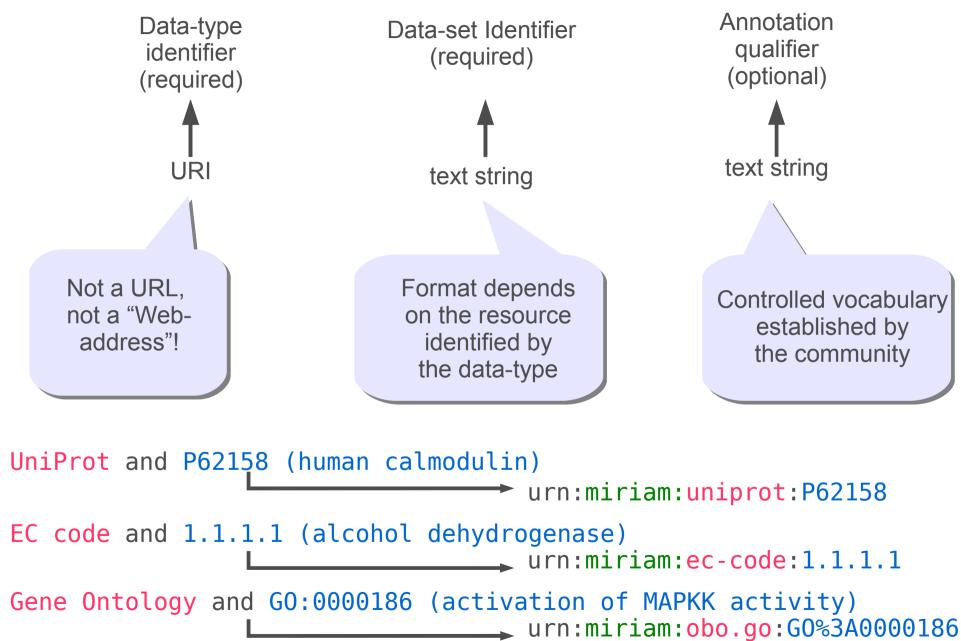
### Why are annotations important?

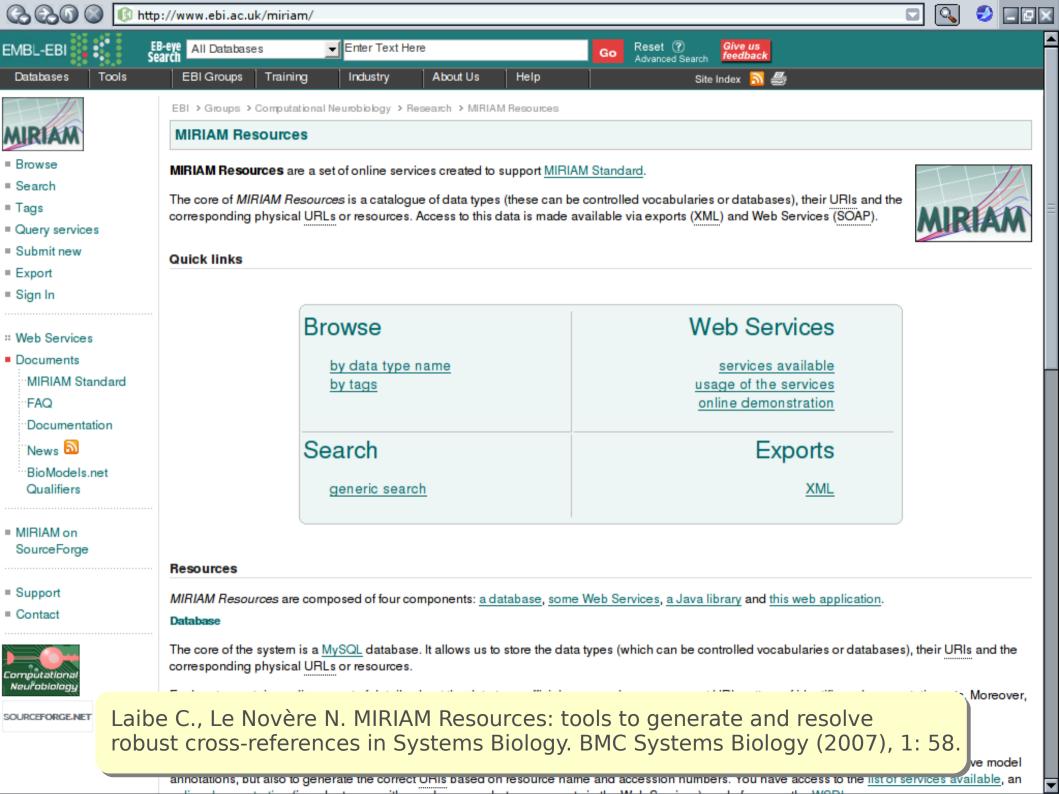
#### Annotation of model components are essential to:

- allow efficient search strategies
- unambiguously identify model components
  - improve understanding the structure of the model
  - allow easier comparison of different models
  - ease the integration of models
- add a semantic layer to the model
  - improve understanding of the biology behind the model
  - allow conversion and reuse of the model
  - ease the integration of model and biological knowledge



#### **MIRIAM cross-references**



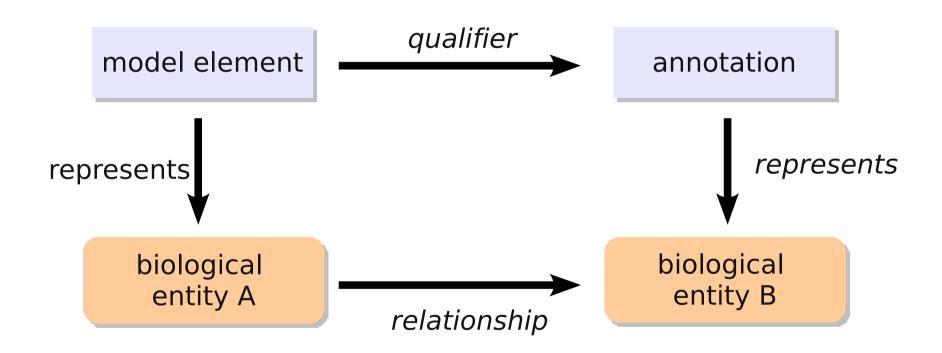


Browse Data type: Enzyme Nomenclature ■ Search ■ Tags General Tags Annotation Query services General information about the data type Submit new Export Name Sign In MIR:00000004 **Identifier** Enzyme Nomenclature Name EC code :: Web Services Synonyms Enzyme Classification Documents EC MIRIAM Standard URIS urn:miriam:ec-code FAQ Official URN http://www.ec-code.org/ Documentation Deprecated urn:Isid:ec-code.org News 🔊 http://www.ebi.ac.uk/IntEnz/ Information ·BioModels.net The Enzyme Classification contains the recommendations of the Nomenclature Committee of the International Union of Qualifiers **Definition** Biochemistry and Molecular Biology on the nomenclature and classification of enzyme-catalysed reactions. Identifier Pattern MIRIAM on **Physical Locations** SourceForge Data Entry http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&ec=\$id [Example: 1.1.1.1 @] http://www.ebi.ac.uk/intenz/ Data Resource Resource #1 Support Information IntEnZ (Integrated relational Enzyme database) European Bioinformatics Institute, United Kingdom Institution Contact Data Entry http://www.genome.jp/dbget-bin/www\_bget?ec:**\$id** [Example: 1.1.1.1 위] Data Resource http://www.genome.jp/dbget-bin/www\_bfind?enzyme Resource #2 Information KEGG Ligand Database for Enzyme Nomenclature Institution Kyoto University Bioinformatics Center, Japan Neurobiologu Data Entry http://us.expasy.org/cgi-bin/nicezyme.pl?\$id [Example: 1.1.1.1 6] Data Resource http://us.expasy.org/enzyme/ SOURCEFORGE NET Resource #3 Enzyme nomenclature database, ExPASy (Expert Protein Analysis System) Information Institution Swiss Institute of Bioinformatics, Switzerland **Documentation** http://www.chem.qmul.ac.uk/iubmb/enzyme/ URL(s) http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-view+MedlineFull+[medline-PMID:10812475] Miscellaneous 2006-08-14 19:38:06 GMT Date of creation 2009-05-08 14:59:31 GMT Date of last modification Go back to the list of data types



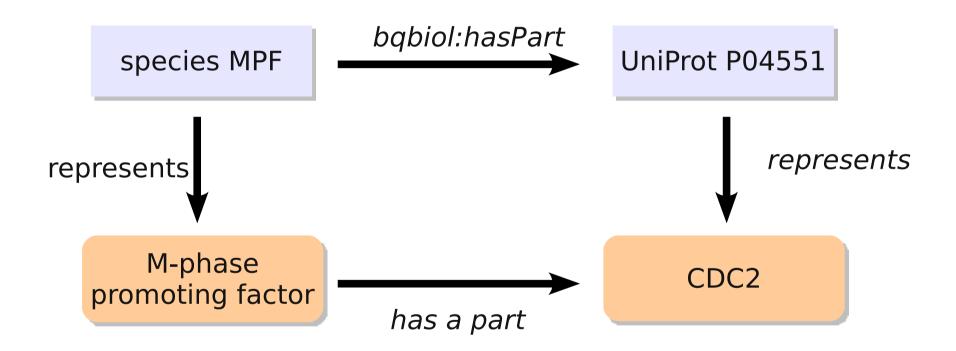


#### **Qualification of annotation**





#### **Qualification of annotation**





#### SBML and MIRIAM cross-references

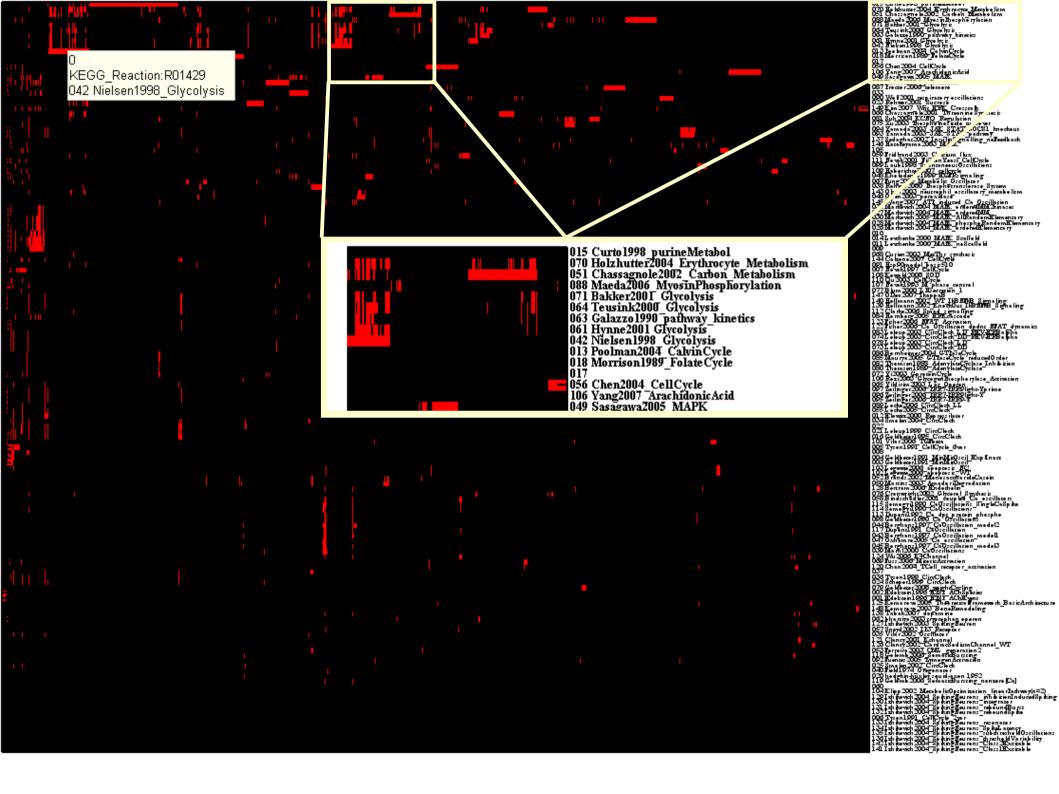
```
<species id="Ca_calmodulin" metaid="cacam">
  <annotation>
    <rdf:RDF
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:bqbiol="http://biomodels.net/biology-qualifiers/">
      <rdf:Description rdf:about="#cacam">
        <bgbiol:hasPart>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:uniprot:P62158"/>
            <rdf:li rdf:resource="urn:miriam:obo.chebi:CHEBI%3A29108"/>
          </rdf:Bag>
        </bqbiol:hasPart>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```



### Tools developing support for MIRIAM identifiers

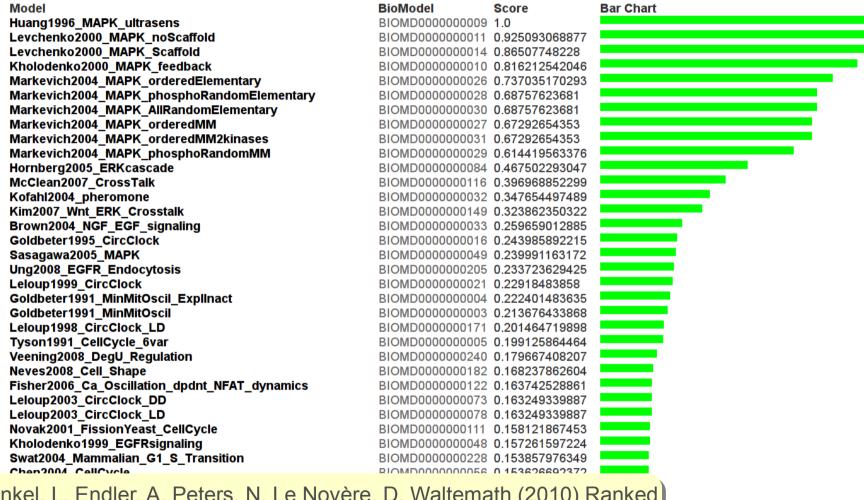
- Data resources
  - BioModels Database (kinetic models)
  - PSI consortium (protein interactions)
  - Reactome (pathways)
  - Pathway commons (pathways)
  - SABIO-RK (reaction kinetics)
  - Yeast consensus model database
  - Human consensus model database
  - E-MeP (structural genomics)
- MIRIAM Resources statistics
  - ~5000 web page requests per month
  - ~550000 web service requests per month

- Application software
  - ARCADIA (graph editor)
  - BIOUML (modeling and simulation)
  - COPASI (Simulation)
  - libAnnotationSBML
  - libSBML
  - SAINT (semantic annotation)
  - SBML2BioPAX
  - SBML2LaTeX
  - SBMLeditor (model editor)
  - SemanticSBML (annotation and merging)
  - Snazer (Network analysis, Simulations)
  - Systems Biology Workbench (model design and simulation)
  - The Virtual Cell (Simulation)



semanticSBML 2.0

#### BioModels Similar to BIOMD000000009.xml



Q

Login

**∌** □ □ ×

R. Henkel, L. Endler, A. Peters, N. Le Novère, D. Waltemath (2010) Ranked Retrieval of Computational Biology Models. BMC Bioinformatics, 11:423

M Schulz, F Krause, N Le Novère, E Klipp, and W Liebermeister: Comparison and clustering of biochemical network models based on semantic annotations, submitted

## Visual representation of models

	Models	Simulation	Results
Minimal requirements	MIRIAM	MIASE	
Data-mode	SSGN	ML	SBRML
Ontologies	S30	KISAO	TEDDY

Home

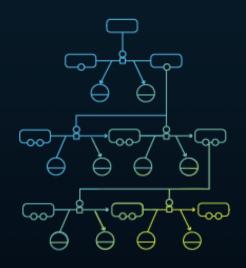
News Documents

Lists

Community

Events About

Q Google Site Search.



#### A Visual Notation for Network Diagrams in Biology

SBGN.org is the global portal for documentation, news, and other information about the Systems Biology Graphical Notation (SBGN) project, an effort to standardize the graphical notation used in maps of biochemical and cellular processes studied in systems biology.

Standardizing the visual representation is crucial for more efficient and accurate transmission of biological knowledge between different communities in research, education, publishing, and more. When biologists are as familiar with the notation as electronics engineers are familiar with the notation of circuit schematics, they can save the time and effort required to familiarize themselves with different notations, and instead spend more time thinking about the biology being depicted.

SBGN is made up of <u>three orthogonal languages</u>, representing different visions of biological systems. Each language defines a comprehensive set of symbols with precise semantics, together with detailed syntactic rules how maps are to be interpreted.

On this site, you can browse some <u>example maps</u> to get a feeling for SBGN, read the SBGN <u>specification documents</u>, <u>software supporting SBGN</u>, get answers to <u>frequent questions about SBGN</u>, access join <u>online discussions</u>, see current working documents in the SBGN wiki, and much more.

SBGN is the work of many people. It would not have been possible without the generous <u>support of multiple organizations</u> over the years, for which we are very thankful.

To quote SBGN as a whole, please use:

Le Novère N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, Bergman FT, Gauges R, Ghazal P, Kawaji H, Li L,

N. Le Novère, M. Hucka, H. Mi, S. Moodie, F. Shreiber, A. Sorokin, et al (2009). The Systems Biology Graphical Notation. *Nature Biotechnology*, 27: 735-741.

SBGN News

(23 Apr.'10) The first <u>annual</u> competition is opened, with categories such as Best Software, Best Map and Best Outreach.



#### What is SBGN?

- An unambiguous way of graphically describing and interpreting biochemical and cellular events
- Limited amount of symbols
   Re-use existing symbols

Smooth learning curve

- Can represent logical or mechanistic models, biochemical pathways, at different levels of granularity
- Detailed technical specification, precise data-models and growing software support
- Developed over four years by a diverse community, including biologists, modellers, computer scientists etc.

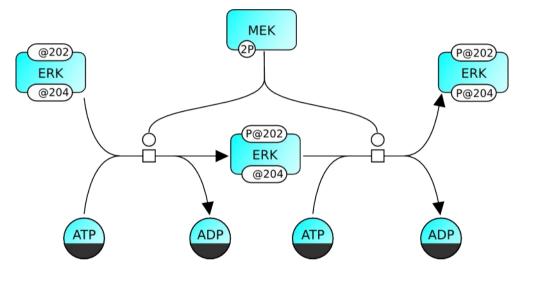


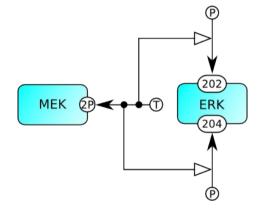
### Graph trinity: three languages in one notation

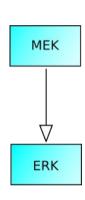
#### <u>Process Descriptions</u>

Entity Relationships

**Activity Flows** 







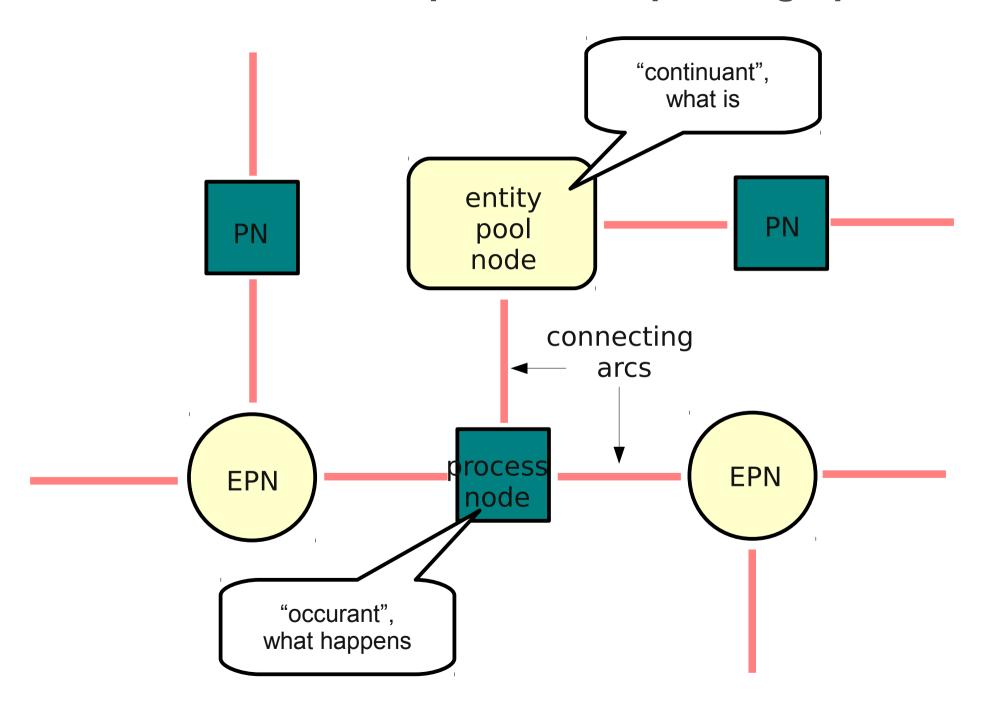
- Unambiguous
- Mechanistic
- Sequential
- Combinatorial explosion

- Unambiguous
- Mechanistic
- Non-sequential
- Independence of relationships

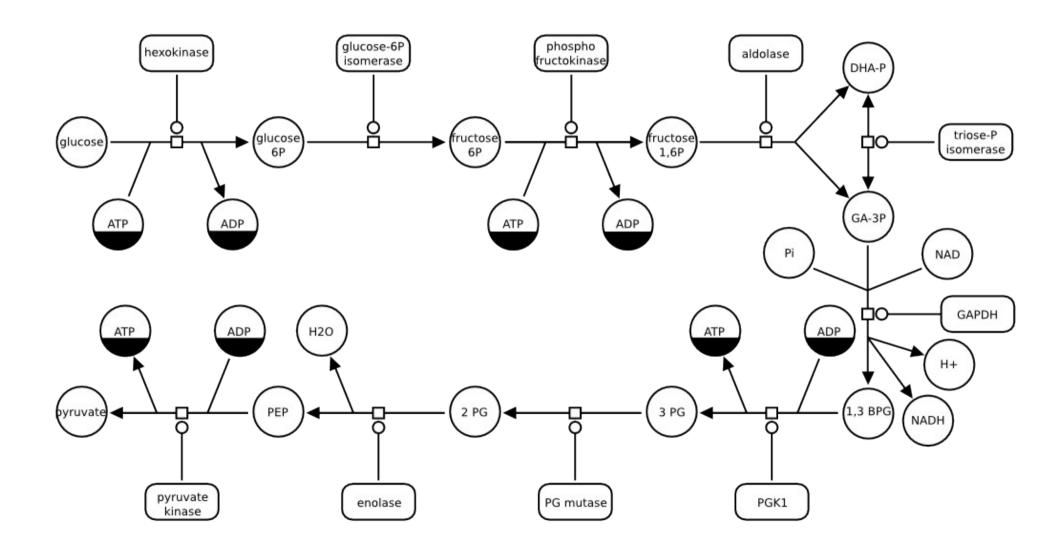
- Ambiguous
- Conceptual
- Sequential



#### **Process Descriptions are bipartite graphs**

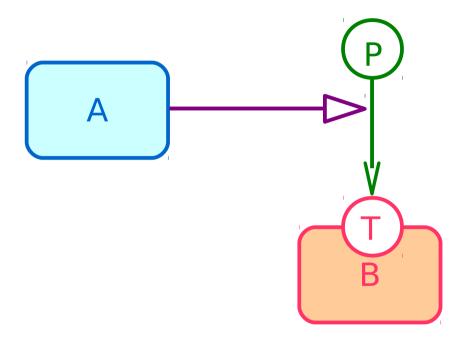


#### Metabolic network in Process Description Language





#### Entity Relationships can be viewed as rules

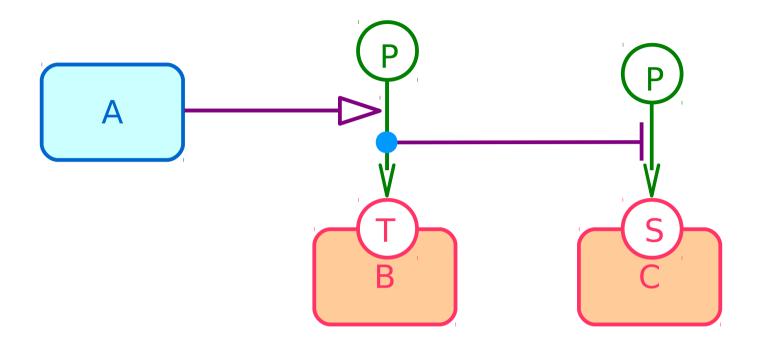


If A exists, the assignment of the value P to the state variable T of B is increased

(A stimulates the phosphorylation of B on the threonine)



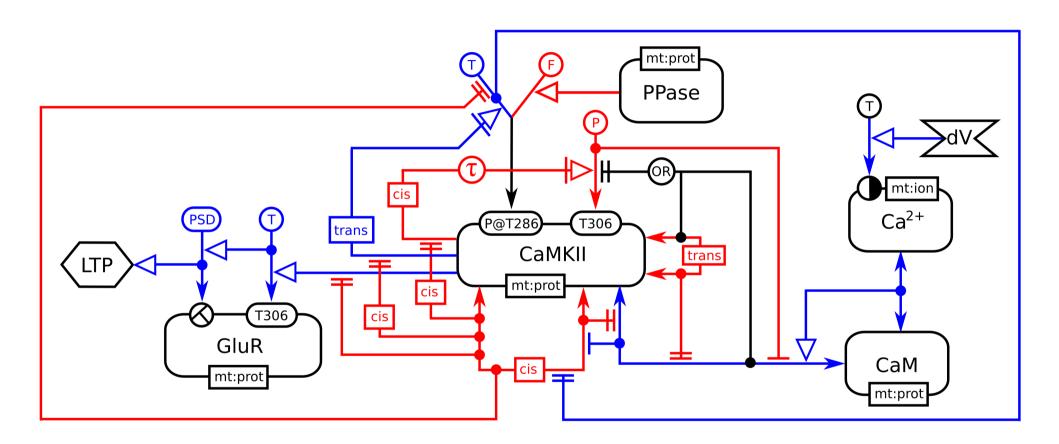
#### Entity Relationships can be viewed as rules



If A exists, the assignment of the value P to the state variable T of B is increased

If P is assigned to the state variable T of B, the assignment of the value P to the state variable S of B is decreased

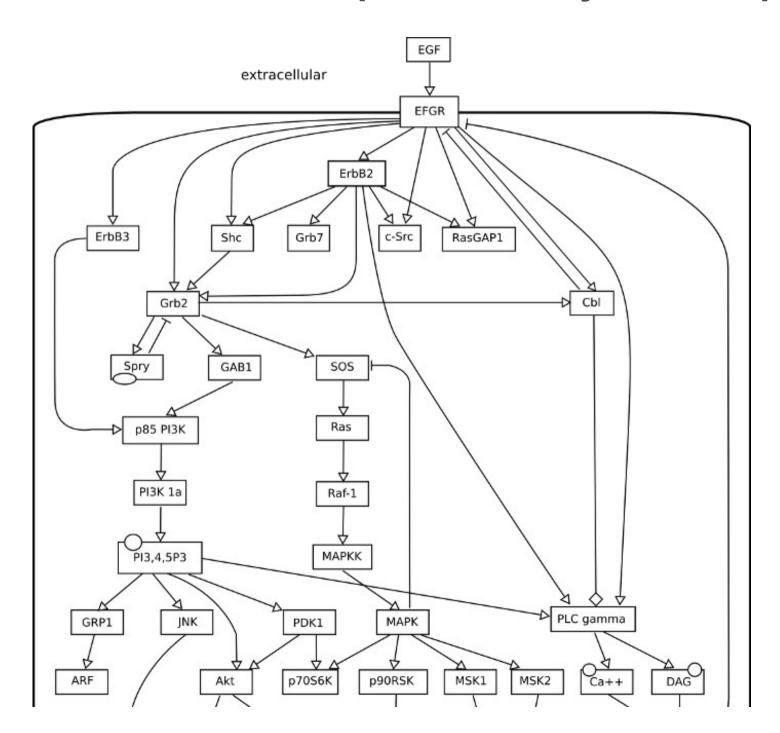
# GNER map of calcium-regulated synaptic plasticity



increases synaptic weight decreases synaptic weight

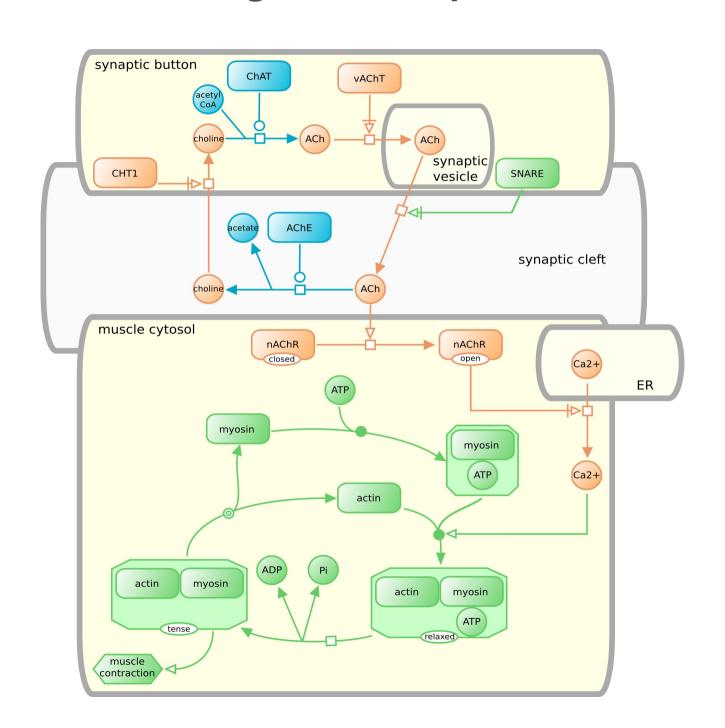


### **Example of Activity Flow map**





### **Linking SBGN maps to external information**



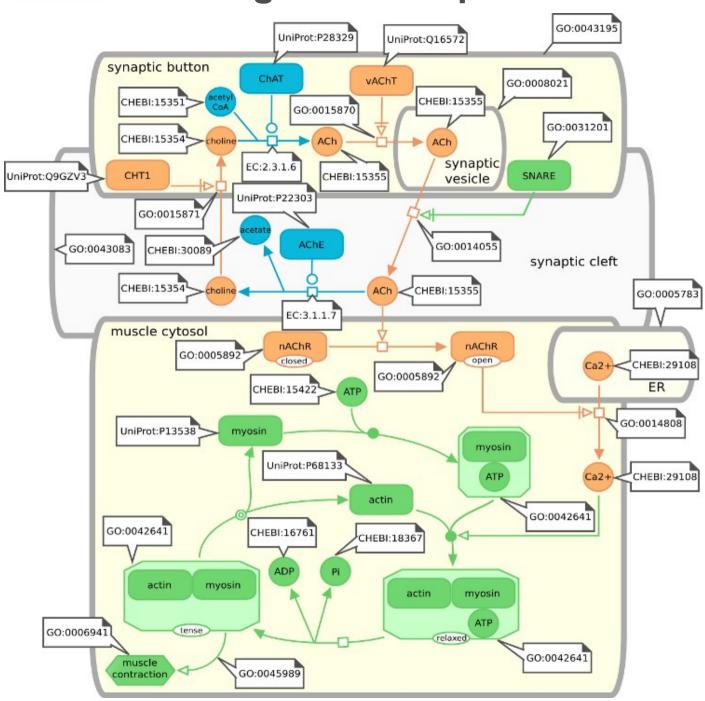
catalytic processes

transport processes

contractile proteins

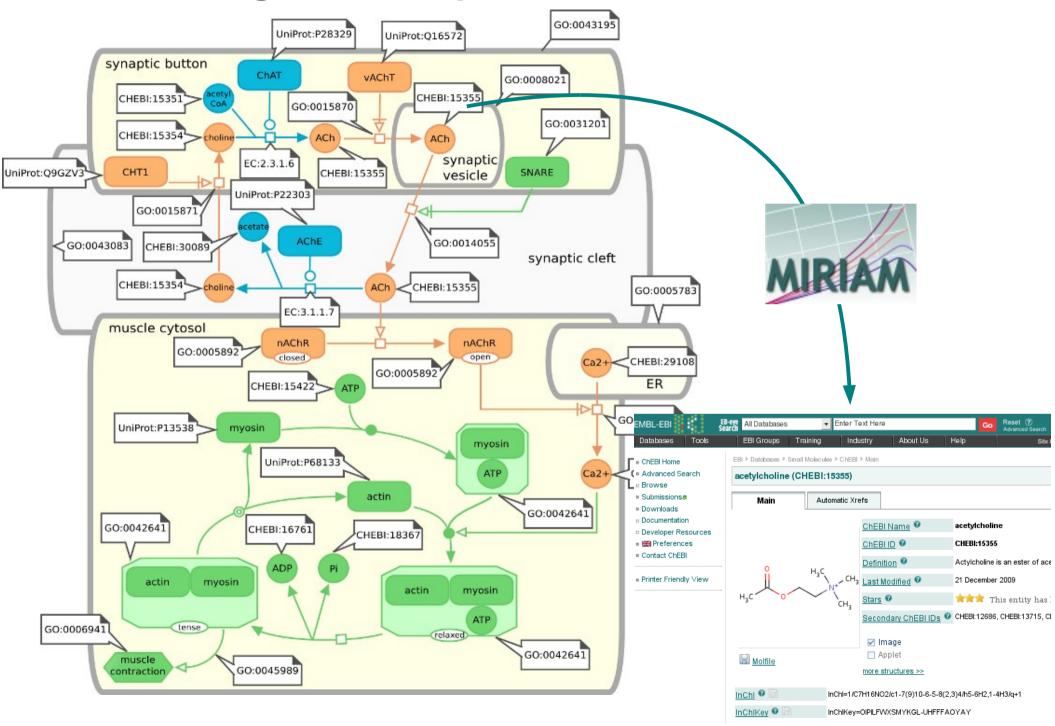


# GIN Linking SBGN maps to external information



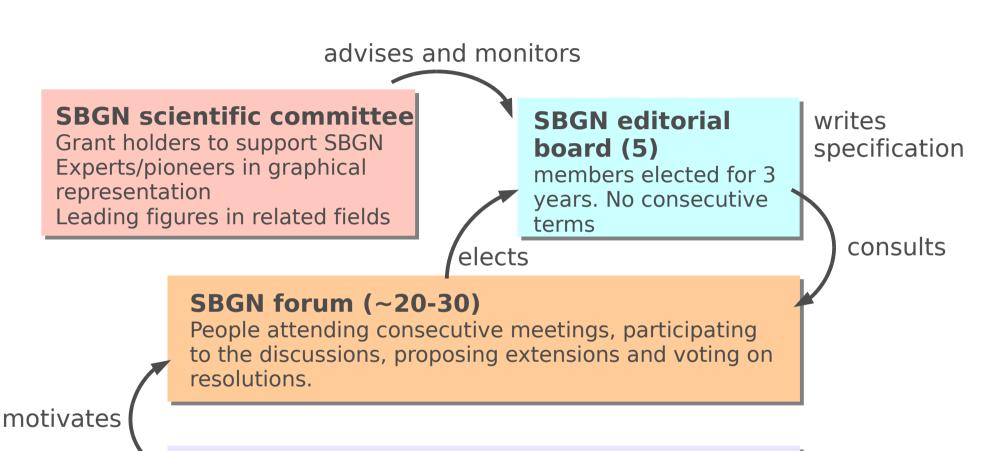


# GN Linking SBGN maps to external information





#### **Current structure of the SBGN community**



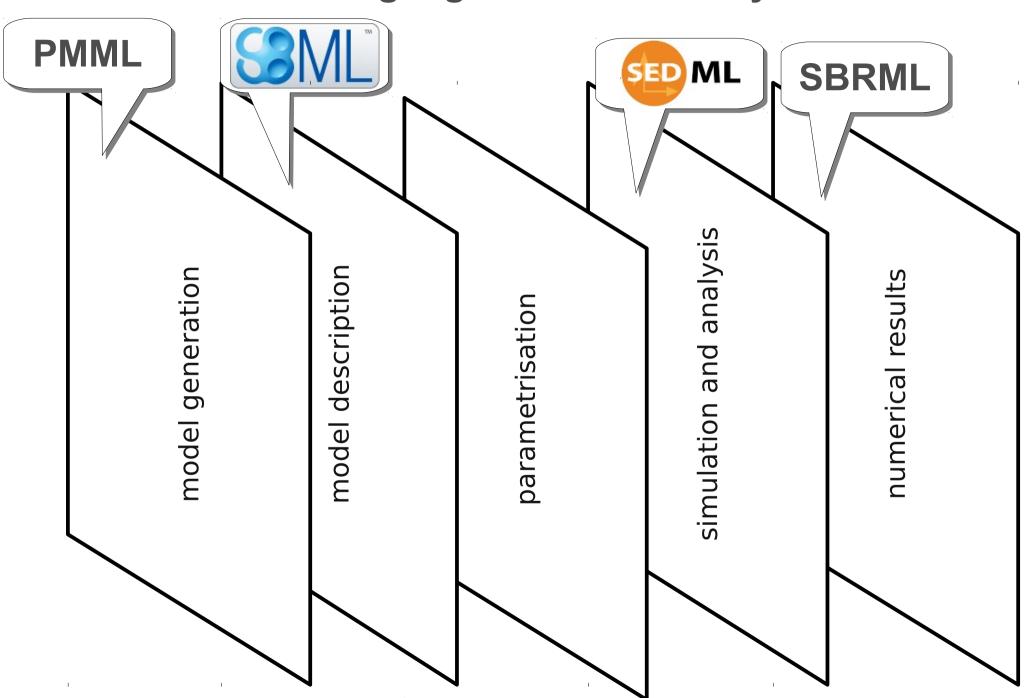
#### **SBGN** community

People who attended a meeting, participated to a discussion, requested a feature, implemented support for SBGN, or simply use SBGN

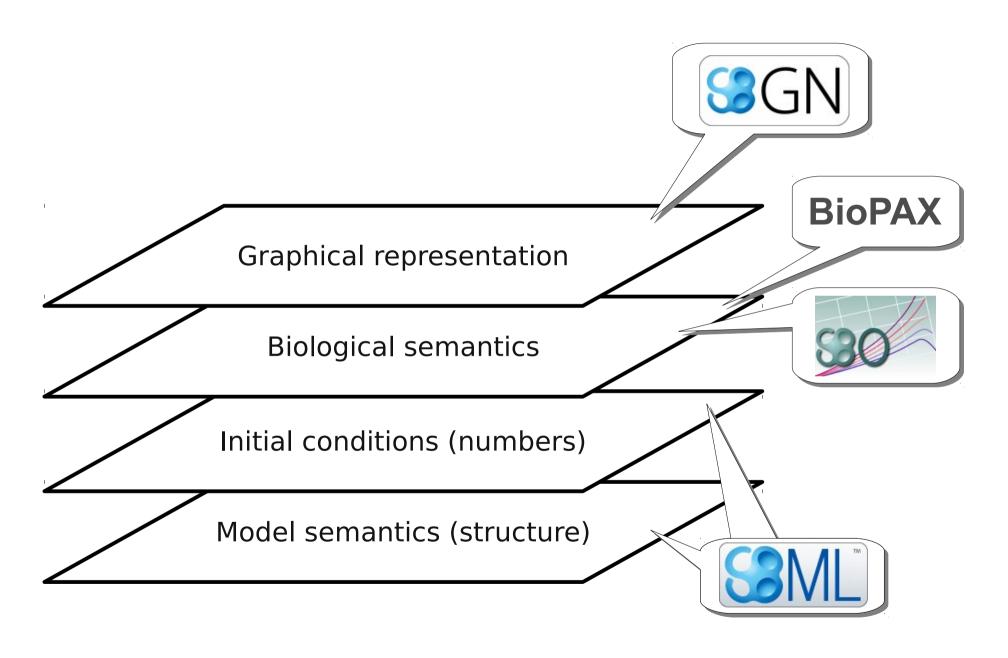
### Is the mosaic of standards complete?

	Models	Simulation	Results
Minimal requirements	MIRIAM	MIASE	
Data-models	SML SGN	SED ML	SBRML
Ontologies	<b>S30</b>	KISAO	Y E D D Y

### Disentangling the model life-cycle



### Disentangling the level of discourse



#### **Biological semantics**

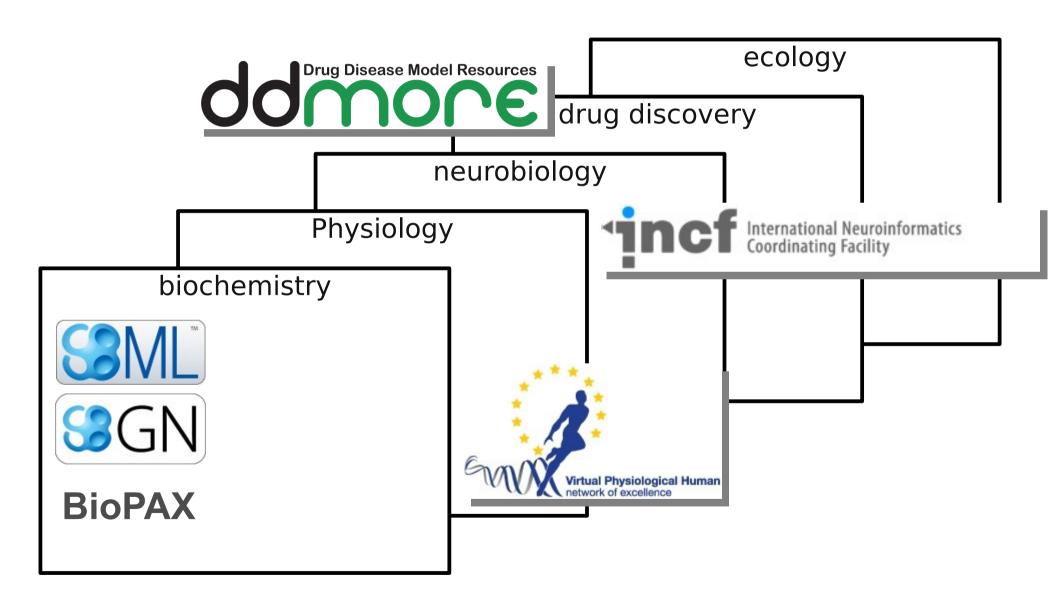
**Systems Biology Ontology (SBO)** Set of linked controlled vocabularies, defining and relating concepts used in computational modelling in biology. Used in SBML, SBGN, NeuroML

Le Novère N., Courtot M., Laibe C. Adding semantics in kinetics models of biochemical pathways. *Proc 2<sup>nd</sup> Intl Symp on experimental standard conditions of enzyme characterizations* (2007), 137-153.

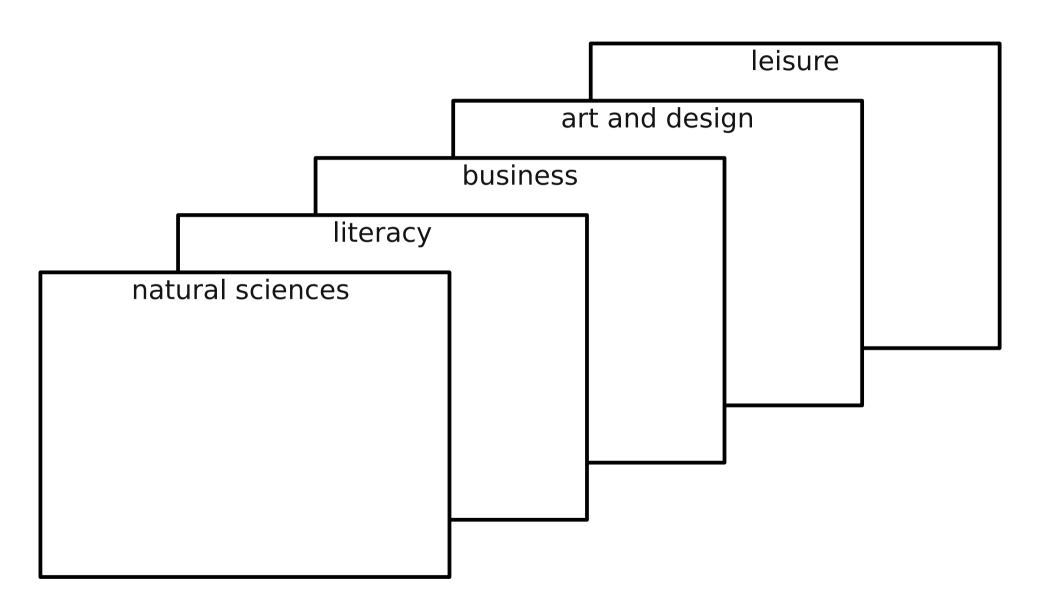
**Biological Pathway Exchange (BioPAX)** Standard language to represent biological pathways at the molecular and cellular level and to facilitate the exchange of pathway data. Used by most major pathways databases

Demir E, Cary MP, Paley S, Fukuda K, Lemer C, Vastrik I, Wu G, D'Eustachio P, Schaefer C, Luciano J, Schacherer F, Martinez-Flores I, Hu Z, Jimenez-Jacinto V, Joshi-Tope G, Kandasamy K, Lopez-Fuentes AC, Mi H, Pichler E, Rodchenkov I, Splendiani A, Tkachev S, Zucker J, Gopinathrao G, Rajasimha H, Ramakrishnan R, Shah I, Syed M, Anwar N, Babur O, Blinov M, Brauner E, Corwin D, Donaldson S, Gibbons F, Goldberg R, Hornbeck P, Luna A, Murray-Rust P, Neumann E, Reubenacker O, Samwald M, van Iersel M, WimalaratneS, Allen K, Braun B, Carrillo M, Cheung KH, Dahlquist K, Finney A, Gillespie M, Glass E, Gong L, Haw R, Honig M, Hubaut O, Kane D, Krupa S, Kutmon M, Leonard J, Marks D, Merberg D, Petri V, Pico A, Ravenscroft D, Ren L, Shah N, Sunshine M, Tang R, Whaley R, Letovksy S, Buetow KH, Rzhetsky A, Schachter V, Sobral BS, Dogrusoz U, McWeeney S, Aladjem M, Birney E, Collado-Vides J, Goto S, Hucka M, Le Novère N, Maltsev N, Pandey A, Thomas P, Wingender E, Karp PD, Sander C, Bader GD. BioPAX – A Community Standard for Pathway Data Sharing. *Nat Biotechnol* (2010), 28: 935–942

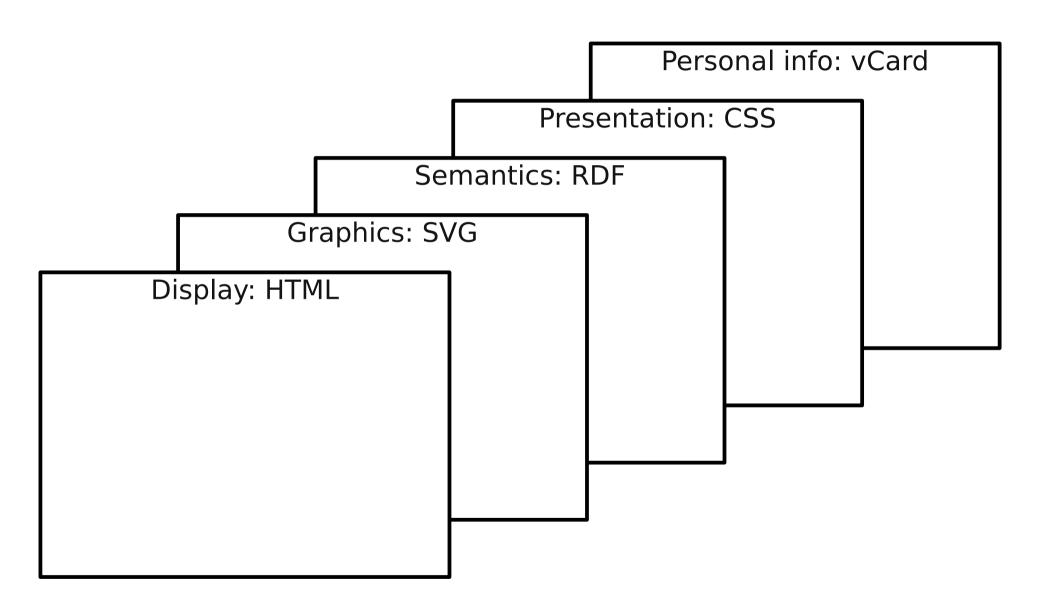
#### Parallel and redundant efforts



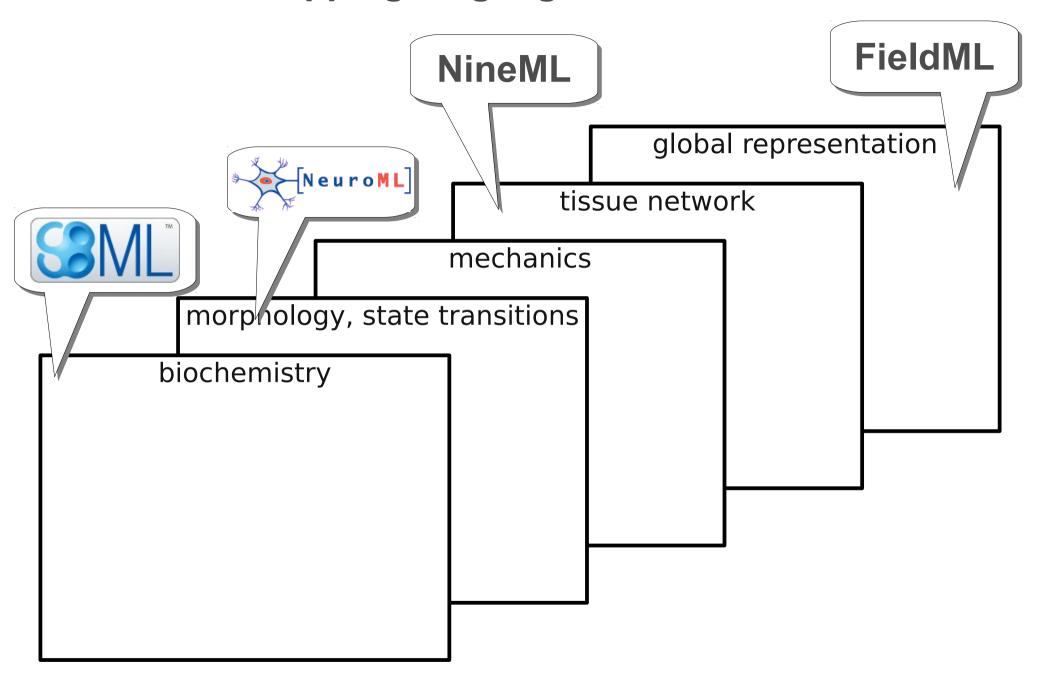
#### What if the world-wide web was built like this?



### The correct way to do it



## Non-overlapping languages to cover all models



## Standards interoperability along the 3 dimensions

- SBML to BioPAX: conversion using metadata (MIRIAM annotations and SBO terms), e.g.
  - mapping between Species and PhysicalEntity
  - mapping between Reactions and PhysicalInteraction
- Conversion of SBML or BioPAX descriptions to SBGN map
- Usage of SBML descriptions (or CellML or VCML) in SED-ML: Identification of variables using XPath
- Descriptions using SBML and NeuroML: Interface based on XPath

# Requirements for a overarching standardisation structure

- What?
  - Set of interoperable description languages
  - Cover all aspects of modelling and simulation, all types of descriptions / views of the real
  - Role of community-maintained ontologies.
- How?
  - Independence towards Institutions, funders and individuals
  - Role of European Research Infrastructures? (ELIXIR, ISBE)
- Who?
  - Communities developing their standards: Systems Biology, Physiology (VPH), Neuroscience (INCF), Drug discovery (DDMoRe), Clinical data (CDISC) ...
  - Other players in knowledge-representation: W3C ...
  - Academic and corporate users: Modeling platforms (MatWorks ...), Pharma (Pistoia alliance) ...

# Requirements for a overarching standardisation structure

- What?
  - Set of interoperable description languages
  - Cover all aspects of modelling and simulation, all types of descriptions / views of the real
  - Role of community-maintained ontologies.
- How?
  - Independence towards Institutions, funders and individuals
  - Role of European Research Infrastructures? (ELIXIR, ISBE)
- Who?
  - Communities developing their standards: Systems Biology, Physiology (VPH), Neuroscience (INCF), Drug discovery (DDMoRe), Clinical data (CDISC) ...
  - Other players in knowledge-representation: W3C ...
  - Academic and corporate users: Modeling platforms (MatWorks ...),
     Pharma (Pistoia alliance) ...

# COmputational Modeling in Blology NEtwork (COMBINE) forthcoming: http://co.mbine.org

### **COMBINE 2010**

- 6 to 9 October 2010, Edinburgh, before the ICSB
- 77 registrations so far (forecast was 50 max ...)
- 14 sessions, plus breakouts, 42 presentations, 30 posters

Physiome standards
SED-ML
SBGN languages
libSBGN and SBGN support

Encoding graph layouts
Interactions and reactions
Semantics and metadata resources
Encoding and using semantics

Format conversion
Software support
BioPAX levels
What is not covered yet

SBML Level 3
libSBML and SBML support
followed by:

SBML 10<sup>th</sup> anniversary



Views: desktop mobile print

STANDARDS PARTICIPATE MEMBERSHIP

ABOUT W3C



Skip

W3C

Standards

Participate

Membership

About W3C

Member Home

W3C » About W3C

#### **ABOUT W3C**

The World Wide Web Consortium (W3C) is an international community where Member organizations, a full-time staff, and the public work together to develop Web standards. Led by Web inventor Tim Berners-Lee and CEO Jeffrey Jaffe, W3C's mission is to lead the Web to its full potential. Contact W3C for more information.



#### W3C Mission

Principles, vision, ...



#### Facts About W3C

People, organization, revenues, process, patent policy, history, ...



#### Press and Analysts

Press releases, requests for photos and interviews, ...



#### Donations

Support for W3C operations, validator service, ...



#### Jobs and Fellowships

▶ Employment opportunities, fellows, ...

#### Questions About W3C or the Web?

Please consult the Help and FAQ for answers to questions such as:

What does W3C do? How is W3C funded?

Is W3C sending me spam?

What is the difference between the Web and the Internet?

QUICK LINKS

Contact

Staff

Help and FAQ

Systems Status (outages)



Views: desktop mobile print

STANDARDS

PARTICIPATE

MEMBERSHIP

ABOUT W3C

Google™

Q

Skip

wac 🗏

Standards

Participate

Membership

About W3C

Member Home

### ABOUT W3C

W3C » About W3C

The World Wide Web Consortium (W3C) is an international community where Member organizations, a full-time staff, and the public work together to develop Web standards. Led by Web inventor Tim Berners-Lee and CEO Jeffrey Jaffe. W3C's mission is to lead the Web to its full potential. Contact W3C for more information.



W3C Mission

Principles, vision, ...

QUICK LINKS

Contact

Staff

Help and FAQ

A "WorldWide Web consortium" for the modeling in life-sciences



Donations

Support for W3C operations, validator service, ...



Jobs and Fellowships

Employment opportunities, fellows, ...

Questions About W3C or the Web?

Please consult the Help and FAQ for answers to questions such as:

What does W3C do?

How is W3C funded?

Is W3C sending me spam?

What is the difference between the Web and the Internet?

s (outages)

## Research perspectives

- Change of scale for the support of Systems Biology
  - Towards a coherent set of interoperable standards to represent models and simulations at all levels
  - Development of ontologies for model semantics in biology and their integration with experimental and clinical data

"W3C" for modelling in biology

- Scaling up BioModels Database to cover all mathematical models in biology
- Multi-scale models of neuronal transmission in basal ganglia
  - 4D models of post-synaptic density
  - Forest of kinase signalling linked by phosphatases
  - Integration of biochemical cascades and eletrophysiological behaviours
  - Mechanical models: spine volume, dendritic remodelling
  - Modelling the basic unit of basal ganglia

## Multi-approach, multi-scale models in neuroscience

- Models and simulations in neuroscience cover all scales and approaches: Reactiondiffusion, biochemistry, electrophysiology (cable approximation), mechanical (neurite growth), networks, field approaches (EEG)
- Multi-scale biological problems are concrete:
   Molecular basis of synaptic plasticity -> transmission of signal by neuron -> cognitive function
   mutations -> molecular aggregates -> neuronal death -> behavioural pathologies
- Versatile simulation tools exist that cover all levels.
- An expertise is building up on hybrid models: Biochemistry/electrophysiology; networks/fields
- There is a political will: role of the *International Neuroinformatics Coordination Facility*: Development of standards, ontologies and simulation tools.
- Neurobiology can make real the aims of the Virtual Physiological Human by déveloping integrated models from the molecule to the organ, and by using them to understand normal functions, emergence of pathologies, and to propose/test new treatments in silico.

### Scientific environment in Bordeaux

- Molecule-synapse: CNRS UMR 5091 Physiologie cellulaire de la synapse (dir. Christophe Mulle), soon INES, on the Neurocampus, and in particular the group of Daniel Choquet for receptor dynamics
- Neuron: INSERM U 862 Physiopathologie de la plasticité neuronale, Neurocentre Magendie (dir. Pier Vincenzo Piazza), on the Neurocampus, for the dendritic remodelling
- Networks: CNRS UMR 5228 Centre de Neurosciences Intégratives et Cognitives (dir. Georges Di Scala), for the modelling of basal ganglia
- Tissue, organism, pathology: CNRS UMR 5227 Mouvement Adaptation –
  Cognition (dir. Jean-Rene Cazalets), and in particular the group of Erwan
  Bezard/Bertrand Bloch, soon institute of neurodegenerative diseases on the
  Neurocampus
- Computational resources for Systems Biology: CBiB, Antoine de Daruvar and soon Macha Nikolskaia
- Data visualisation and manipulation: MaBioVis/Gravite, Guy Melançon
- Event driven simulations: MAGNOME, David Sherman
- Advanced numerical methods: MC2, Thierry Colin
- **High performance simulation**: HiePACS, Jean Roman
- •

## **Acknowledgements**

Visionary: Hiroaki Kitano

SBML editors: Frank Bergmann, *Andrew Finney, Stefan Hoops*, *Michael Hucka*, *Nicolas Le Novère, Sarah Keating*, Sven Sahle, *Herbert Sauro*, Jim Schaff, Lucian Smith, Darren Wilkinson

SGN editors: Emek Demir, *Michael Hucka*, Nicolas Le Novère, Huaiyu Mi, Stuart Moodie, Falk Schreiber, *Anatoly Sorokin* 

MIRIAM: Nick Juty, Camille Laibe

The whole community of Computational Systems Biology

The EBI group Computational Systems Neurobiology











