EMBL-EBI

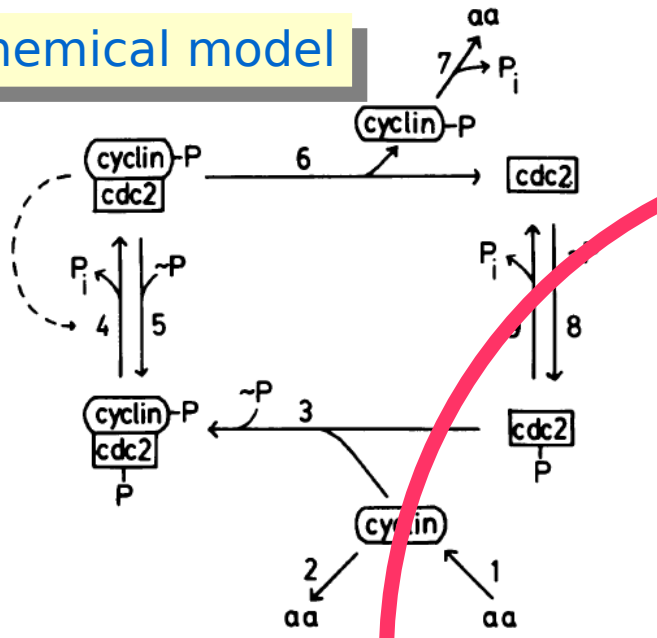
Metadata for Systems Biology

Nick Juty, EMBL-EBI

MLSB09 Ljubljana, Slovenia 5-6th September, 2009

Which models do we mean?

biochemical model

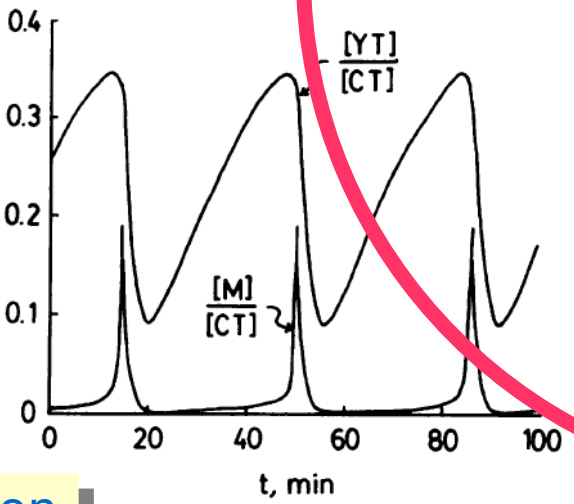


mathematical model

$$d[C2]/dt = k_6[M] - k_8[\sim P][C2] + k_9[CP]$$
$$d[CP]/dt = -k_3[CP][Y] + k_8[\sim P][C2] - k_9[CP]$$
$$d[pM]/dt = k_3[CP][Y] - [pM]F([M]) + k_5[\sim P][M]$$
$$d[M]/dt = [pM]F([M]) - k_5[\sim P][M] - k_6[M]$$
$$d[Y]/dt = k_1[aa] - k_2[Y] - k_3[CP][Y]$$
$$d[YP]/dt = k_6[M] - k_7[YP]$$

Parameter	Value	Notes
$k_1[aa]/[CT]$	0.015 min^{-1}	*
k_2	0	†
$k_3[CT]$	200 min^{-1}	*
k_4	10–1000 min^{-1} (adjustable)	
k_4'	0.018 min^{-1}	
$k_5[\sim P]$	0	‡
k_6	0.1–10 min^{-1} (adjustable)	
k_7	0.6 min^{-1}	†
$k_8[\sim P]$	$>> k_9$	§
k_9	$>> k_6$	§

simulation



computational model

Tyson et al (1991) PNAS 88(1):7328-32

What is a model

- A model is a mathematical description of the components of a system, their relationships, and the evolution of both.
  - ordinary differential equations (system evolution) $dX/dt = f(X)$
  - partial differential equation (system description) $\nabla X = g(X)$
  - algebraic equations (conservation laws) $h(X) = 0$
  - probability distributions $PX = i(X)$
  - master equation $dPX/dt = j(PX)$
  - ..

Common formats for model storage and exchange

SBML	http://sbml.org/
CellML	http://www.cellml.org/
BioPAX	http://www.biopax.org/

Common formats for model storage and exchange

SBML	http://sbml.org/	quantitative
CellML	http://www.cellml.org/	quantitative
BioPAX	http://www.biopax.org/	kinetic unsupported

General and abstract – range of problems

- inference of components
- automated conversion


Main Page - SBML.org - SeaMonkey
SBML.org
The Systems Biology Markup Language
News Documents Downloads Forums Facilities Community Events About
Google Site Search.
The Systems Biology Markup Language (SBML) is a computer-readable format for representing models of biological processes. It's applicable to simulations of metabolism, cell-signaling, and many other topics. SBML has been evolving since mid-2000 thanks to an international community of software developers and users. This website is the global portal for the SBML effort; here you can find information about all aspects of SBML.
For the curious
What is SBML? Read our basic introduction and then perhaps browse the mailing lists to get a sense for what's currently going on in the world of SBML.
For modelers
Are you looking for ready-to-run software that supports SBML? The SBML Software Guide lists over 170 systems today. Are you instead looking for ready-to-use models? Visit the BioModels Database, where you can find hundreds!
For software developers
Are you interested in developing SBML support for 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, an API library supporting many programming languages.
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 call attention to your project's support of SBML by displaying the SBML logo.
None of this would be possible without the support of multiple agencies and organizations. Visit our acknowledgements page to learn about the visionary funding agencies that have backed SBML over the years.
SBML News
More errata for L2v4r1
(11 July '09) A few more issues have been noted recently on the SBML Level 2 Version 4 errata list.
SBML Level 3 plans updated
(8 Jul. '09) The SBML Editors have updated the plans for SBML Level 3 Core and are seeking feedback.
SBML FAQ updated
(1 Jul. '09) FAQ item #3.6 now features an expanded explanation of SBML Level 2 annotations.
Older news ...
Community News
BioModels Database rel. 14
(16 Jun.'09) BioModels Database, a free database of published models, now contains 216 curated and 196 uncurated models.
MCSim supports SBML
(25 Apr.'09) GNU MCSim lets you design statistical or simulation models & perform Bayesian inference.
Pathway Tools workshop
(24 Apr.'09) Tutorial & workshop on SRI's Pathway Tools and BioCyc database in August '09.
Older news ...
Done

- machine readable XML format for encoding models
- software neutral
- developed since 2000 as an exchange format
  levels / versions

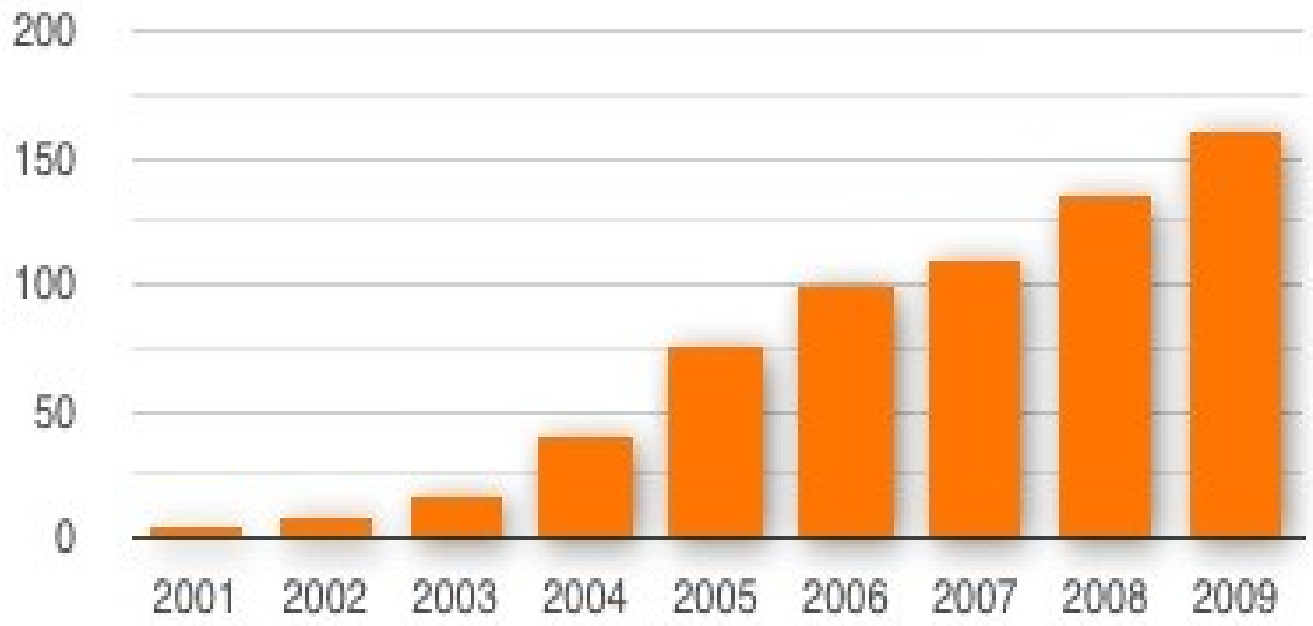
- commonly used format allows:

  sharing of models
  sharing of tools

SBML : A → B

```
<?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>cell</ci>
            </apply>
          </math>
        </kineticLaw>
      </reaction>
    </listOfReactions>
  </model>
</sbml>
```


SBML-compatible software



>170 tools supporting SBML

SBML Software Matrix

This matrix provides an at-a-glance summary of software known to us to provide some degree of support for reading, writing, or otherwise working with SBML. The columns' meanings are explained below. For a list of longer descriptions grouped into themes, please see our **SBML Software Summary** page.

	Capabilities					Frameworks							API	Dep.	Platforms	SBML		Availabil.		
	Creation	Simulation	Analysis	Database	Utility	ODE	DAE	PDE	Stochastic	Events	Logical	Other				Import	Export	Open source	Academic use	Commercial use
acslXtreme	•														W	•			$	$
ALC	•					•	•		•			•			L, W, M, B		•	•	F	F
Asmparts	•				•	•									L,W	•	•	•	F	F
Antimony	•				•								C, C++		L, W, M	•	•	•	F	F
AutoSBW			•			•							SBW	SBW	L, W, M	•	•	•	F	F
AVIS												•		various	L	•		•	F	F
BALSA	•													Sigtran						
BASIS	•	•		•					•	•			WS		B	•	•	•	F	F
BetaWB	•	•	•						•	•					L,W,M		•		F	F
BiNoM	•		•		•							•			L, W, M	•	•	•	F	F
BiNoM Cytoscape Plugin	•		•		•							•		Cytoscape	L, W, M	•	•	•	F	F
BIOCHAM		•			•	•									L,W,M		•	•	F	F
BioCharon	•	•	•		•	•								CHARON						
Biological Networks	•		•		•										L,W,M	•	•		F	$
BioCyc				•													•		F	$
BioGrid																				

The columns of this table should be read in the following way:

- *Capabilities* summarizes the facilities that a package provides by itself (i.e., without invoking another package) for working with SBML: "Creation" = creating/editing models, "Simulation" = performing time-series simulation of models, "Analysis" = analyzing models (e.g., sensitivity analysis, flux-balance analysis, etc.), "Database" = providing a database of models, and "Utility" = providing other utility functions (e.g., translating SBML to/from other formats).
- *Frameworks* summarizes the modeling frameworks supported by a package, regardless of whether the package

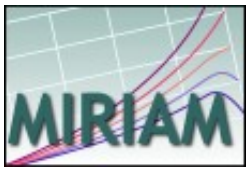
Is SBML enough? What's missing?

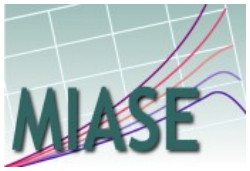
- lists participants, but does not identify them
- contains mathematical expressions but does describe meaning
- model constructed for a modelling approach cannot be easily cast

- **SBML models not easily converted**
- **SBML models not easily merged**

layer of semantics necessary to capture missing knowledge

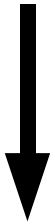
Minimal requirements





?

implemented by

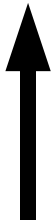


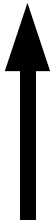
Example Data-model

SED-ML

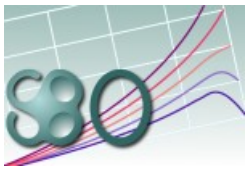
SBRML

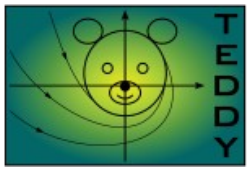
adds meaning to





Ontology





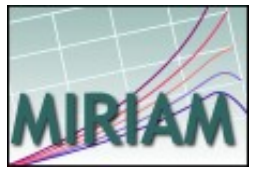
Minimal requirements

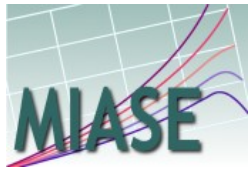
implemented by

Example Data-model

adds meaning to

Ontology

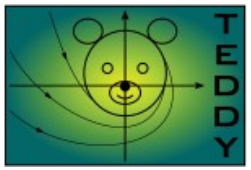




?

SED-ML

SBRML



Annotations to provide missing information

Annotation of model components are essential at 2 levels:

- unambiguously identify model components
  - improve understanding of model structure
  - 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

_computational BIOLOGY

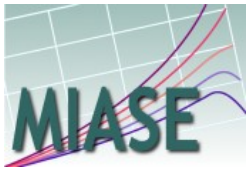
COMMENTARY

Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project

Chris F Taylor[*1,2], Dawn Field[2,3], Susanna-Assunta Sansone[1,2], Jan Aerts[4], Rolf Apweiler[1], Michael Ashburner[5], Catherine A Ball[6], Pierre-Alain Binz[7,8], Molly Bogue[9], Tim Booth[2], Alvis Brazma[1], Ryan R Brinkman[10], Adam Michael Clark[11], Eric W Deutsch[12], Oliver Fiehn[13], Jennifer Fostel[14], Peter Ghazal[15], Frank Gibson[16], Tanya Gray[2,3], Graeme Grimes[15], John M Hancock[17], Nigel W Hardy[18], Henning Hermjakob[1], Randall K Julian Jr[19], Matthew Kane[20], Carsten Kettner[21], Christopher Kinsinger[22], Eugene Kolker[23,24], Martin Kuiper[25], Nicolas Le Novère[1], Jim Leebens-Mack[26], Suzanna E Lewis[27], Phillip Lord[16], Ann-Marie Mallon[17], Nishanth Marthandan[28], Hiroshi Masuya[29], Ruth McNally[30], Alexander Mehrle[31], Norman Morrison[2,32], Sandra Orchard[1], John Quackenbush[33], James M Reecy[34], Donald G Robertson[35], Philippe Rocca-Serra[1,36], Henry Rodriguez[22], Heiko Rosenfelder[31], Javier Santoyo-Lopez[15], Richard H Scheuermann[28], Daniel Schober[1], Barry Smith[37], Jason Snape[38], Christian J Stoeckert Jr[39], Keith Tipton[40], Peter Sterk[1], Andreas Untergasser[41], Jo Vandesompele[42] & Stefan Wiemann[31]

The Minimum Information for Biological and Biomedical Investigations (MIBBI) project provides a resource for those exploring the range of extant minimum information checklists and fosters coordinated development of such checklists.

Minimal requirements



?

MIBBI

implemented by

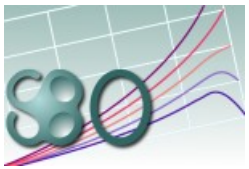
Example Data-model

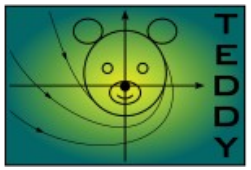
SED-ML

SBRML

adds meaning to

Ontology





What are standard reporting guidelines?

MI standards provide standard reporting guidelines.

List of the core set of information that has to be provided with a data-set, so that a user is able to make sensible use of it.

- “Minimal” - only the essential information is given
- “Standard” - information should be provided in a form that can be fully interpreted by community

MIBBI: Minimum Information for Biological and Biomedical Investigations

Project News

- **BMC journals recommend MIBBI** in their 'Instructions to Authors' (example)
- **Free download:** The MIBBI paper (*Nature Biotechnology*) & supplementary information (additional figures)

Site navigation

The MIBBI Portal

Access to Minimum Information guidelines for diverse bioscience domains

About us

A contextualisation of the project, our rules and regulations, and our publications and talks.

The MIBBI Foundry

Towards the next generation of MI guidelines for the biosciences

Project news

Announcements relating to the project, such as new registrations, meetings, *etc*.

Related resources

Links to other cross-domain projects, policy statements and sundry useful material

Discussion

How to post to the MIBBI discussion forum, or join the Foundry developers' mailing list

MIBBI search

A Google™ Custom

http://www.mibbi.org

links

- Home
- The MIBBI Portal
- The MIBBI Foundry
- Related resources
- About us
- Project news
- Discussion
- MIBBI search

search

Go | Search

toolbox

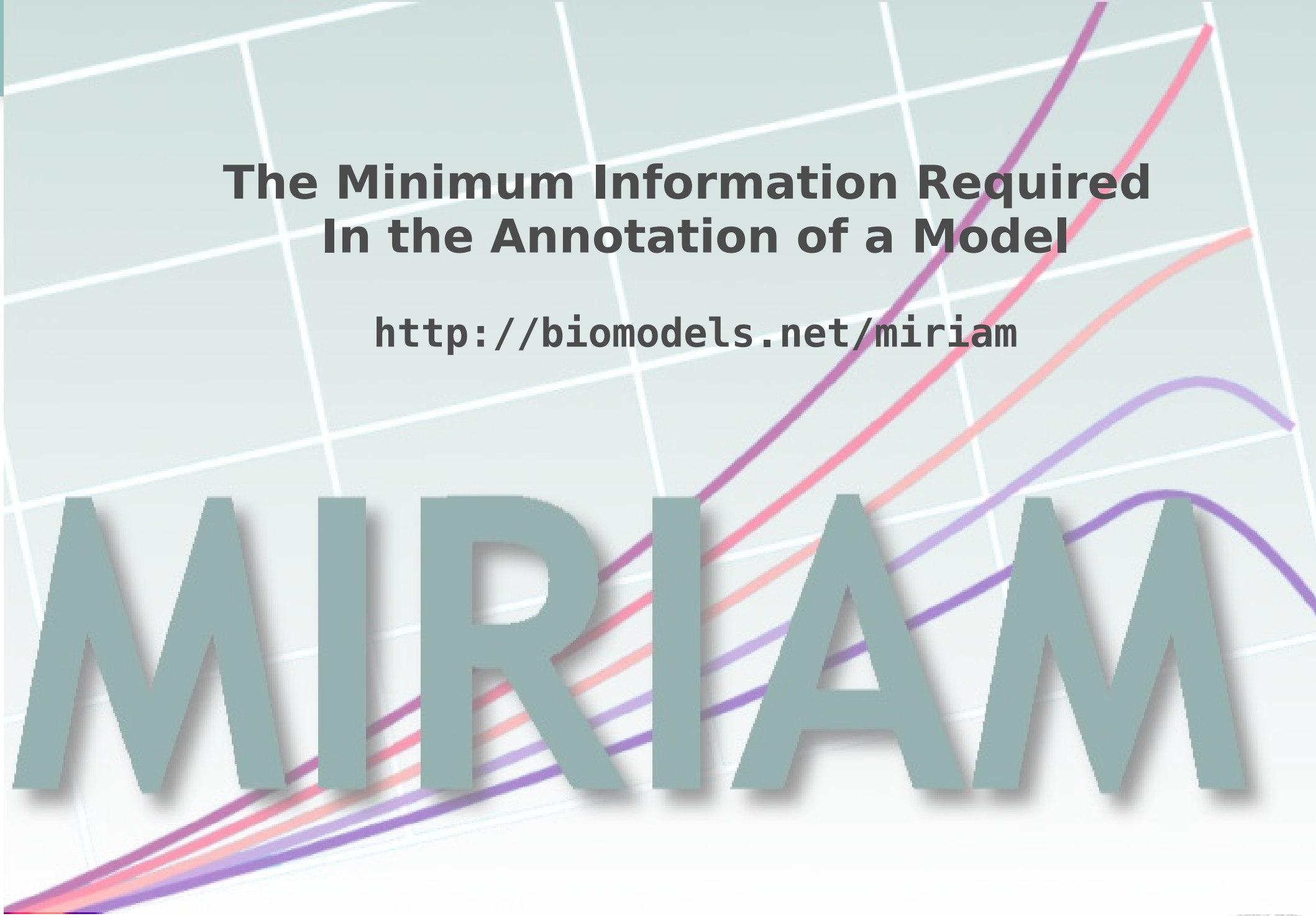
- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

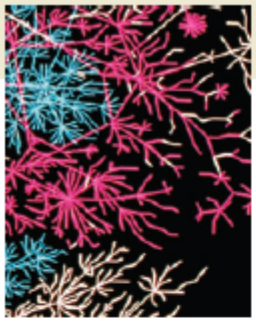
- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

MIAME	**M**inimum **I**nformation **A**bout a **M**icroarray **E**xperiment
MIAME/Env	**MIAME / Env**ironmental transcriptomic experiment
MIAME/Nutr	**MIAME / Nutr**igenomics
MIAME/Plant	**MIAME / Plant** transcriptomics
MIAME/Tox	**MIAME / Tox**icogenomics
MIAPA	**M**inimum **I**nformation **A**bout a **P**hylogenetic **A**nalysis
MIAPAR	**M**inimum **I**nformation **A**bout a **P**rotein **A**ffinity **R**eagent
MIAPE	**M**inimum **I**nformation **A**bout a **P**roteomics **E**xperiment
MIARE	**M**inimum **I**nformation **A**bout a **R**NAi **E**xperiment
MIASE	**M**inimum **I**nformation **A**bout a **S**imulation **E**xperiment
MIENS	**M**inimum **I**nformation about an **EN**vironmental **S**equence
MIFlowCyt	**M**inimum **I**nformation for a **Flow Cyt**ometry Experiment
MIGen	**M**inimum **I**nformation about a **Gen**otyping Experiment
MIGS	**M**inimum **I**nformation about a **G**enome **S**equence
MIMIx	**M**inimum **I**nformation about a **M**olecular **I**nteraction **Ex**periment
MIMPP	**M**inimal **I**nformation for **M**ouse **P**henotyping **P**rocedures
MINI	**M**inimum **I**nformation about a **N**euroscience **I**nvestigation
MINIMESS	**Mini**mal **Me**tagenome **S**equence Analysis **S**tandard
MINSEQE	**M**inimum **I**nformation about a high-throughput **SeQ**uencing Experiment
MIPFE	**M**inimal **I**nformation for **P**rotein **F**unctional **E**valuation
MIQAS	**M**inimal **I**nformation for **Q**TLs and **A**ssociation **S**tudies
MIqPCR	**M**inimum **I**nformation about a **q**uantitative **P**olymerase **C**hain **R**eaction experiment
MIRIAM	**M**inimal **I**nformation **R**equired **I**n the **A**nnotation of biochemical **M**odels
	Minimum **I**nformation **S**pecification **F**or **I**n **S**itu **H**ybridization and

The Minimum Information Required In the Annotation of a Model

http://biomodels.net/miriam





_computational BIOLOGY

PERSPECTIVE

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

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

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

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[1,2]. 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

MIRIAM guidelines

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

Annotation - how?

Why annotations should not be raw text

- EMBL bank version 45 (04-DEC-1995):
  /db_xref="PID:g984120"
- EMBL bank version 47 (07-JUN-1996):
  /db_xref="PID:g984120"
  /db_xref="SWISS-PROT:P49581"
- EMBL bank version 60 (03-SEP-1999):
  /db_xref="SWISS-PROT:P49581"
  /protein_id="CAA58766.1"
- EMBL bank version 73 (30-NOV-2002):
  /db_xref="SWISS-PROT:P49581"
  /protein_id="CAA58766.1"
  /db_xref="GOA:P49581"
- EMBL bank version 79 (08-JUN-2004):
  /db_xref="UniProt/Swiss-Prot:P49581"
  /protein_id="CAA58766.1"
  /db_xref="GOA:P49581"
- EMBL bank version 84 (12-SEP-2005):
  /db_xref="UniProtKB/Swiss-Prot:P49581"
  /protein_id="CAA58766.1"
  /db_xref="GOA:P49581"

Why annotations should not be raw text

- EMBL bank version 45 (04-DEC-1995):
  /db_xref="PID:g984120"
- EMBL bank version 47 (07-JUN-1996):
  /db_xref="PID:g984120"
  /db_xref="SWISS-PROT:P49581"
- EMBL bank version 60 (03-SEP-1999):
  /db_xref="SWISS-PROT:P49581"
  /protein_id="CAA58766.1"
- EMBL bank version 73 (30-NOV-2002):
  /db_xref="SWISS-PROT:P49581"
  /protein_id="CAA58766.1"
  /db_xref="GOA:P49581"
- EMBL bank version 79 (08-JUN-2004):
  /db_xref="UniProt/Swiss-Prot:P49581"
  /protein_id="CAA58766.1"
  /db_xref="GOA:P49581"
- EMBL bank version 84 (12-SEP-2005):
  /db_xref="UniProtKB/Swiss-Prot:P49581"
  /protein_id="CAA58766.1"
  /db_xref="GOA:P49581"

vanishes

Why annotations should not be raw text

- EMBL bank version 45 (04-DEC-1995):
  /db_xref="PID:g984120"
- EMBL bank version 47 (07-JUN-1996):
  /db_xref="PID:g984120"
  /db_xref="SWISS-PROT:P49581"
- EMBL bank version 60 (03-SEP-1999):
  /db_xref="SWISS-PROT:P49581"
  /protein_id="CAA58766.1"
- EMBL bank version 73 (30-NOV-2002):
  /db_xref="SWISS-PROT:P49581"
  /protein_id="CAA58766.1"
  /db_xref="GOA:P49581"
- EMBL bank version 79 (08-JUN-2004):
  /db_xref="UniProt/Swiss-Prot:P49581"
  /protein_id="CAA58766.1"
  /db_xref="GOA:P49581"
- EMBL bank version 84 (12-SEP-2005):
  /db_xref="UniProtKB/Swiss-Prot:P49581"
  /protein_id="CAA58766.1"
  /db_xref="GOA:P49581"

Name change

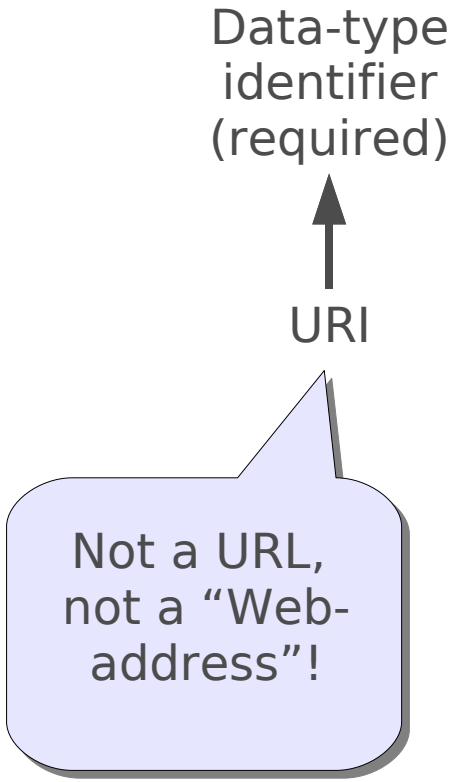
Why annotations should not be uncontrolled URLs

- http://srs6.ebi.ac.uk/srs6bin/cgi-bin/wgetz?[swissprot-AccNumber:P01133]+-e
- http://www.ebi.uniprot.org/uniprot-srv/uniProtView.do?proteinId=P01133
- http://www.ebi.uniprot.org/entry/P01133
- http://www.uniprot.org/uniprot/P01133?proteinId=P01133
- http://www.uniprot.org/uniprot/P01133

Annotation needs to be

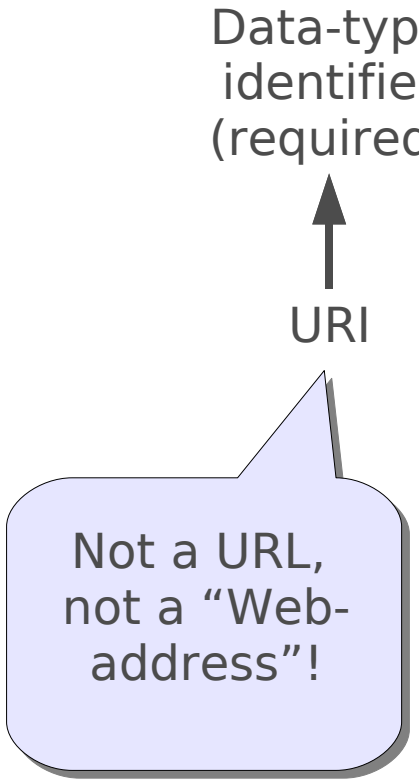
- Unambiguous
- Persistent
- Consistent or reproducible form
- Machine interpretable

MIRIAM annotation



Corresponds to a namespace

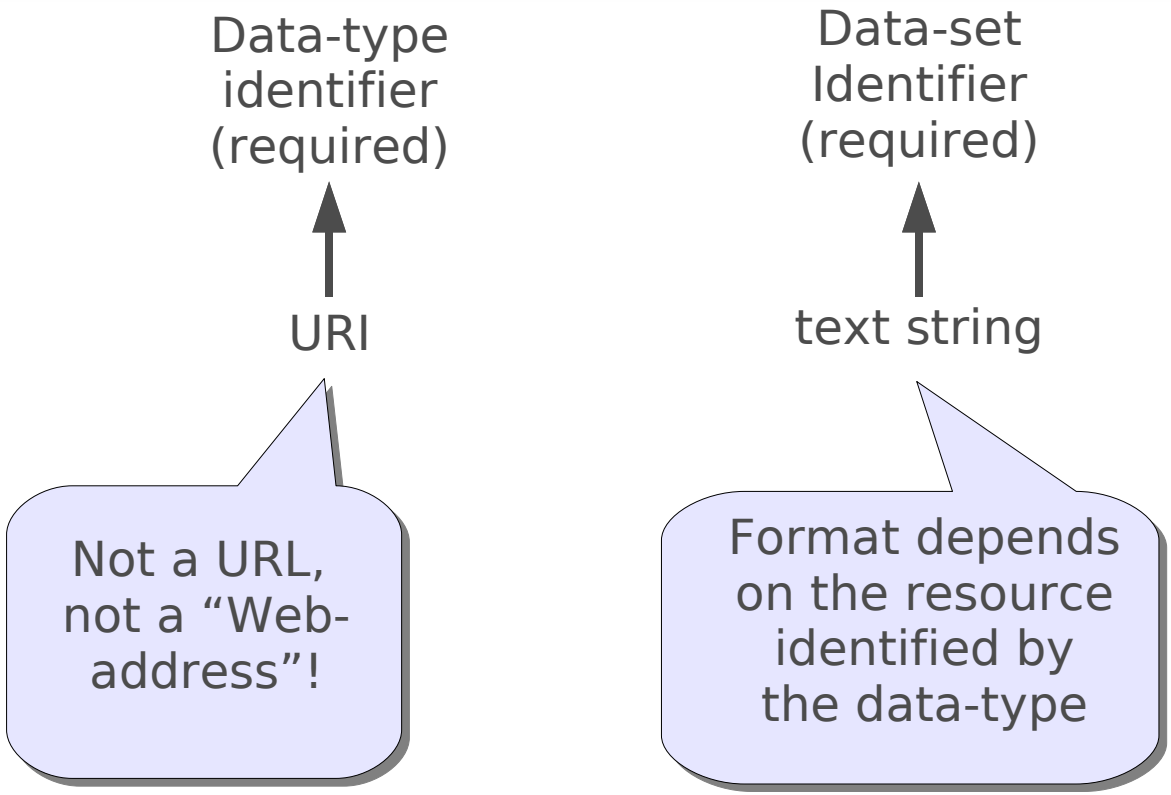
MIRIAM annotation



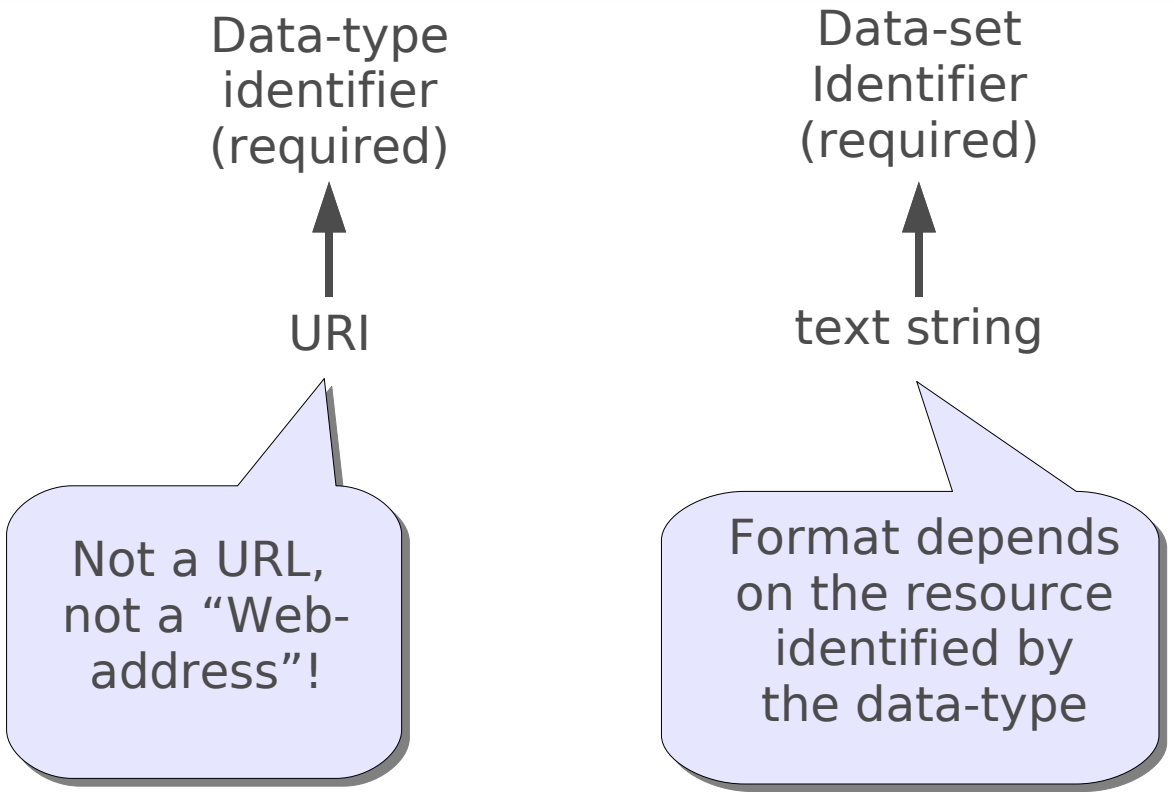
Corresponds to a namespace

```
urn:miriam:uniprot
```


MIRIAM annotation



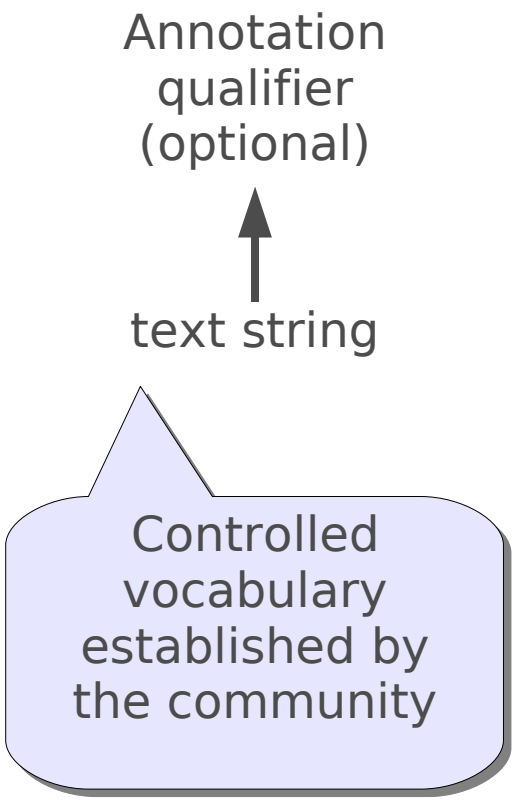
MIRIAM URI



UniProt P62158 (human calmodulin) urn:miriam:uniprot:P62158

EC code 1.1.1.1 (alcohol dehydrogenase) urn:miriam:ec-code:1.1.1.1

Qualifier



defines relationship between model entity and annotation

SBML and MIRIAM

```
<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">
        <bqbiol: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>
```


Some BioModels.net Qualifiers

- **bqmodel:is** The modelling object represented by the model component is the subject of the referenced resource.
- **bqmodel:isDescribedBy** The modelling object represented by the component of the encoded model is described by the referenced resource.
- **bqbiol:is** The biological entity represented by the model component is the subject of the referenced resource.
- **bqbiol:hasPart** The biological entity represented by the model component includes the subject of the referenced resource, either physically or logically.
- **bqbiol:isPartOf** The biological entity represented by the model component is a physical or logical part of the subject of the referenced resource
- **bqbiol:isVersionOf** The biological entity represented by the model component is a version or an instance of the subject of the referenced resource.
- **bqbiol:hasVersion** The subject of the referenced resource is a version or an instance of the biological entity represented by the model component.
- **bqbiol:isHomologTo** The biological entity represented by the model component is homolog, to the subject of the referenced resource, i.e. they share a common ancestor.
- **bqbiol:isDescribedBy** The biological entity represented by the model component is described by the referenced resource.

http://www.biomodels.net/qualifiers/

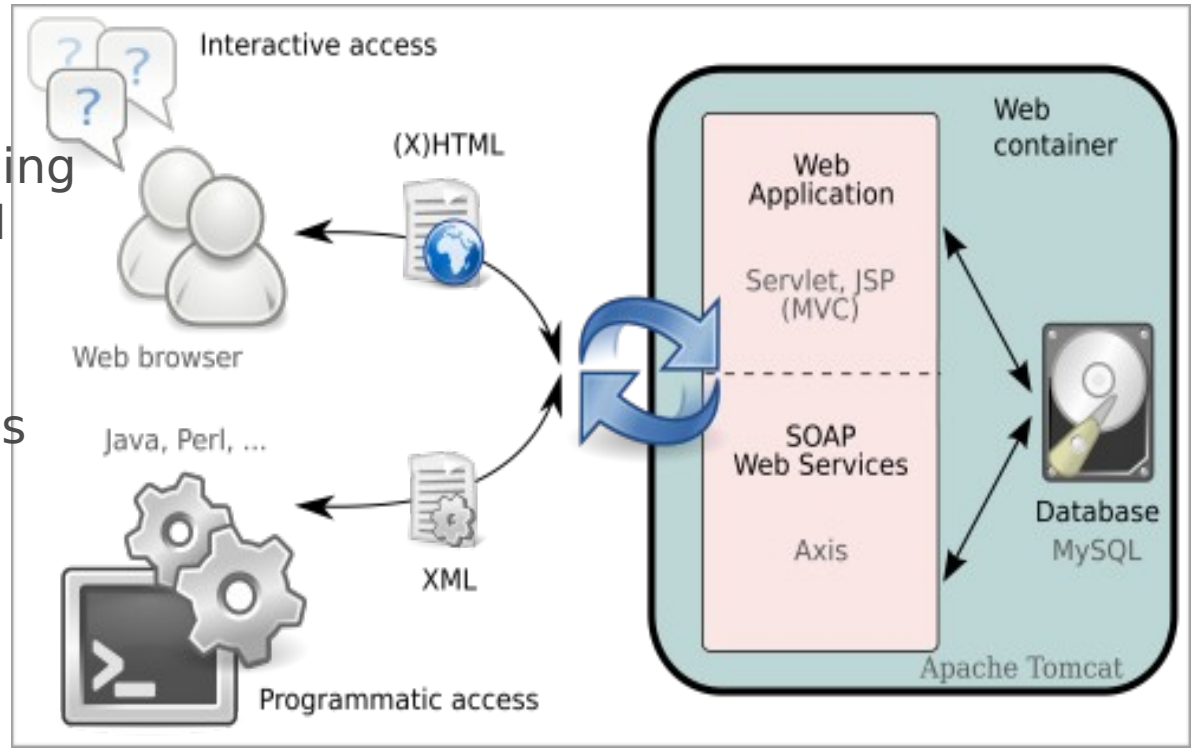
MIRIAM Resources

- **MIRIAM Database**
  Core element of the resource, storing all the information about the data-types and associated information;

- **MIRIAM Web Services**
  SOAP-based application programming interface (API) for querying MIRIAM Database

- **MIRIAM Library**
  Library to use MIRIAM Web Services

- **MIRIAM Web Application**
  Interactive web interface for browsing and querying MIRIAM Database, and submit or edit data-types.



http://www.biomodels.net/miriam/

MIRIAM Resources

Browse the data types

Brief overview of the different data types stored in *MIRIAM Database*.

Next page ⇨

Display: 10 | 20 | 30 | All

Name	URI	Definition
3DMET	urn:miriam:3dmet	3DMET is a database collecting three-dimensional structures of natural metabolites.
Aclame	urn:miriam:aclame	ACLAME is a database dedicated to the collection and classification of mobile genetic elements (MGEs) from various sources, comprising all known phage genomes, plasmids and transposons.
ArrayExpress	urn:miriam:arrayexpress	ArrayExpress is a public repository for microarray data, which is aimed at storing MIAME-compliant data in accordance with Microarray Gene Expression Data (MGED) recommendations.
arXiv	urn:miriam:arxiv	arXiv is an e-print service in the fields of physics, mathematics, non-linear science, computer science, and quantitative biology.
BioGRID	urn:miriam:biogrid	BioGRID is a database of physical and genetic interactions in Saccharomyces cerevisiae, Caenorhabditis elegans, Drosophila melanogaster, Homo sapiens, and Schizosaccharomyces pombe.
BioModels Database	urn:miriam:biomodels.db	BioModels Database is a data resource that allows biologists to store, search and retrieve published mathematical models of biological interests.
BRENDA	urn:miriam:brenda	BRENDA is a collection of enzyme functional data available to the scientific community. Data on enzyme function are extracted directly from the primary literature The database covers information on classification and nomenclature, reaction and specificity, functional parameters, occurrence, enzyme structure and stability, mutants and enzyme engineering, preparation and isolation, the application of enzymes, and ligand-related data.
ChEBI	urn:miriam:obo.chebi	Chemical Entities of Biological Interest (ChEBI) is a freely available dictionary of molecular entities focused on 'small' chemical compounds.
CluSTr	urn:miriam:clustr	The CluSTr database offers an automatic classification of UniProt Knowledgebase and IPI proteins into groups of related proteins. The clustering is based on analysis of all pairwise comparisons (Smith-Waterman) between protein sequences.
Database of Interacting Proteins	urn:miriam:dip	The database of interacting protein (DIP) database stores experimentally determined interactions between proteins. It combines information from a variety of sources to create a single, consistent set of protein-protein interactions
DOI	urn:miriam:doi	The Digital Object Identifier System is for identifying content objects in the digital environment.
Ensembl	urn:miriam:ensembl	Ensembl is a joint project between EMBL - EBI and the Sanger Institute to develop a software system which produces and maintains automatic annotation on selected eukaryotic genomes.
Entrez Gene	urn:miriam:entrez.gene	Entrez Gene is the NCBI's database for gene-specific information, focusing on completely sequenced genomes, those with an active research community to contribute gene-specific information, or those that are scheduled for intense sequence analysis.
Enzyme Nomenclature	urn:miriam:ec-code	The Enzyme Classification contains the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzyme-catalysed reactions.
Evidence Code	urn:miriam:obo.eco	Evidence codes can be used to specify the type of supporting evidence for a piece of knowledge. This allows inference of a 'level of support' between an entity and an annotation made to an entity.
FlyBase	urn:miriam:flybase	FlyBase is the database of the Drosophila Genome Projects and of associated literature.
FMA	urn:miriam:obo.fma	The Foundational Model of Anatomy Ontology (FMA) is a biomedical informatics ontology. It is concerned with the representation of classes or types and relationships necessary for the symbolic representation of the phenotypic structure of the human body. Specifically, the FMA is a domain ontology that represents a coherent body of explicit declarative knowledge about human anatomy.
Gene Ontology	urn:miriam:obo.go	The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism.

Data type: *Enzyme Nomenclature*

General | Tags | Annotation

General information about the data type

Name		
Identifier		MIR:00000004
Name		Enzyme Nomenclature
Synonyms		EC code
		Enzyme Classification
		EC
URIs		
Official URN		urn:miriam:ec-code
Deprecated		http://www.ec-code.org/
		urn:lsid:ec-code.org
		http://www.ebi.ac.uk/IntEnz/
Information		
Definition		The Enzyme Classification contains the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzyme-catalysed reactions.
Identifier Pattern		^\d+\.-\.-\.-|\d+\.\d+\.-\.-|\d+\.\d+\.\d+\.-|\d+\.\d+\.\d+\.\d+$
Physical Locations		
Resource #1	**Data Entry**	http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&ec=**$id** [Example: 1.1.1.1]
	Data Resource	http://www.ebi.ac.uk/intenz/
	Information	IntEnZ (Integrated relational Enzyme database)
	Institution	European Bioinformatics Institute, United Kingdom
Resource #2	**Data Entry**	http://www.genome.jp/dbget-bin/www_bget?ec:**$id** [Example: 1.1.1.1]
	Data Resource	http://www.genome.jp/dbget-bin/www_bfind?enzyme
	Information	KEGG Ligand Database for Enzyme Nomenclature
	Institution	Kyoto University Bioinformatics Center, Japan
Resource #3	**Data Entry**	http://us.expasy.org/cgi-bin/nicezyme.pl?**$id** [Example: 1.1.1.1]
	Data Resource	http://us.expasy.org/enzyme/
	Information	Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)
	Institution	Swiss Institute of Bioinformatics, Switzerland
Documentation		
URL(s)		http://www.chem.qmul.ac.uk/iubmb/enzyme/
		⊞ http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-view+MedlineFull+[medline-PMID:10812475]
Miscellaneous		
Date of creation		2006-08-14 19:38:06 GMT
Date of last modification		2009-05-08 14:59:31 GMT

Go back to the list of data types

Edit this data type

requirements for a MIRIAM-compliant data type

- **Open access**

Anybody can access any public data without restriction (no commercial licence; no login page etc.)

- **Atomicity**

The granularity of the data distributed has to be appropriately selected (A database of “reactions” distributes reactions and not pathways) and consistent (e.g. classes or instances but not classes AND instances)

- **Identifier**

An atomic data is associated to a unique and perennial identifier

- **Community recognition**

The resource has to be “recognised” by the corresponding experimental community, be reasonably supported etc

Tools developing support for MIRIAM identifiers

- Data resources
  - BioModels Database (kinetic models)
  - PSI consortium (protein interactions)
  - Reactome (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)

Are we almost there yet?

SBML gives us :

$$
\begin{aligned}
d[\mathrm{C2}]/dt &= k_6[\mathrm{M}] - k_8[\sim P][\mathrm{C2}] + k_9[\mathrm{CP}] \\
d[\mathrm{CP}]/dt &= -k_3[\mathrm{CP}][\mathrm{Y}] + k_8[\sim P][\mathrm{C2}] - k_9[\mathrm{CP}] \\
d[\mathrm{pM}]/dt &= k_3[\mathrm{CP}][\mathrm{Y}] - [\mathrm{pM}]F([\mathrm{M}]) + k_5[\sim P][\mathrm{M}] \\
d[\mathrm{M}]/dt &= [\mathrm{pM}]F([\mathrm{M}]) - k_5[\sim P][\mathrm{M}] - k_6[\mathrm{M}] \\
d[\mathrm{Y}]/dt &= k_1[\mathrm{aa}] - k_2[\mathrm{Y}] - k_3[\mathrm{CP}][\mathrm{Y}] \\
d[\mathrm{YP}]/dt &= k_6[\mathrm{M}] - k_7[\mathrm{YP}]
\end{aligned}
$$

The model revealed

bqbiol:isVersionOf
Gene Ontology GO:0006468
EC 3.1.3.16

bqbiol:isVersionOf
Gene Ontology
GO:0000079

bqbiol:isVersionOf
UniProt
CDC2_SCHPO

$$
\begin{aligned}
d[\mathrm{C2}]/dt &= k_6[\mathrm{M}] - k_8[\sim P][\mathrm{C2}] + k_9[\mathrm{CP}] \\
d[\mathrm{CP}]/dt &= -k_3[\mathrm{CP}][\mathrm{Y}] + k_8[\sim P][\mathrm{C2}] - k_9[\mathrm{CP}] \\
d[\mathrm{pM}]/dt &= k_3[\mathrm{CP}][\mathrm{Y}] - [\mathrm{pM}]F([\mathrm{M}]) + k_5[\sim P][\mathrm{M}] \\
d[\mathrm{M}]/dt &= [\mathrm{pM}]F([\mathrm{M}]) - k_5[\sim P][\mathrm{M}] - k_6[\mathrm{M}] \\
d[\mathrm{Y}]/dt &= k_1[\mathrm{aa}] - k_2[\mathrm{Y}] - k_3[\mathrm{CP}][\mathrm{Y}] \\
d[\mathrm{YP}]/dt &= k_6[\mathrm{M}] - k_7[\mathrm{YP}]
\end{aligned}
$$

bqbiol:hasPart
UniProt CDC2_SCHPO
InterPro IPR006670

bqbiol:isVersionOf
Gene Ontology GO:0006470
EC 2.7.11.1

bqbiol:isVersionOf
InterPro
IPR006670

bqbiol:hasVersion
Reactome REACT_6216
Reactome REACT_6308

Lack of biological semantics in SBML

```
<listOfCompartments>
  <compartment id="C">
</listOfCompartments>
<listOfSpecies>
  <species id="A"/>
  <species id="B"/>
  <species id="C"/>
</listOfSpecies>
<listOfReactions>
  <reaction>
    <listOfReactants>
      <speciesReference species="A"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C"/>
    </listOfModifiers>
    <kineticLaw>
      <math></math>
      <listOfParameters>
        <parameter id="U"/>
        <parameter id="V"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

"what" has been answered by MIRIAM annotations

Roles and inter-relationships unknown

Adding ontologies

Perspective

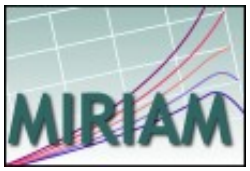
Nature Biotechnology **25**, 1251 - 1255 (2007)
Published online: 7 November 2007 | doi:10.1038/nbt1346

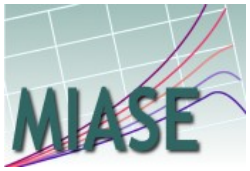
The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration

Barry Smith[1], Michael Ashburner[2], Cornelius Rosse[3], Jonathan Bard[4], William Bug[5], Werner Ceusters[6], Louis J Goldberg[7], Karen Eilbeck[8], Amelia Ireland[9], Christopher J Mungall[10], The OBI Consortium[11], Neocles Leontis[12], Philippe Rocca-Serra[9], Alan Ruttenberg[13], Susanna-Assunta Sansone[9], Richard H Scheuermann[14], Nigam Shah[15], Patricia L Whetzel[16] & Suzanna Lewis[10]

The value of any kind of data is greatly enhanced when it exists in a form that allows it to be integrated with other data. One approach to integration is through the annotation of multiple bodies of data using common controlled vocabularies or 'ontologies'. Unfortunately, the very success of this approach has led to a proliferation of ontologies, which itself creates obstacles to integration. The Open Biomedical Ontologies (OBO) consortium is pursuing a strategy to overcome this problem. Existing OBO ontologies, including the Gene Ontology, are undergoing coordinated reform, and new ontologies are being created on the basis of an evolving set of shared principles governing ontology development. The result is an expanding family of ontologies designed to be interoperable and logically well formed and to incorporate accurate representations of biological reality. We describe this OBO Foundry initiative and provide guidelines for those who might wish to become involved.

Minimal requirements





?

implemented by

Example Data-model

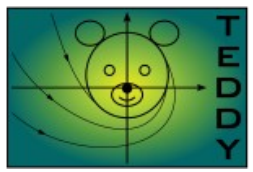
SED-ML

SBRML

adds meaning to

OBO

Ontology



OBO –> guidelines and best practices

>60 ontologies library members

- open - terms should be available for use without restriction
- changeable – modified by community effort
- orthogonal – additivity and modularity
- machine friendly – interpretable and syntactically sound
- identifiers – backward compatible

The Open Biomedical Ontologies

Ontologies | Resources | Participate | About

The OBO Foundry is a collaborative experiment involving developers of science-based ontologies who are establishing a set of principles for ontology development with the goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain. The groups developing ontologies who have expressed an interest in this goal are listed below, followed by other relevant efforts in this domain.

In addition to a listing of OBO ontologies, this site also provides a statement of the OBO Foundry principles, discussion fora, technical infrastructure, and other services to facilitate ontology development. We welcome feedback and encourage participation.

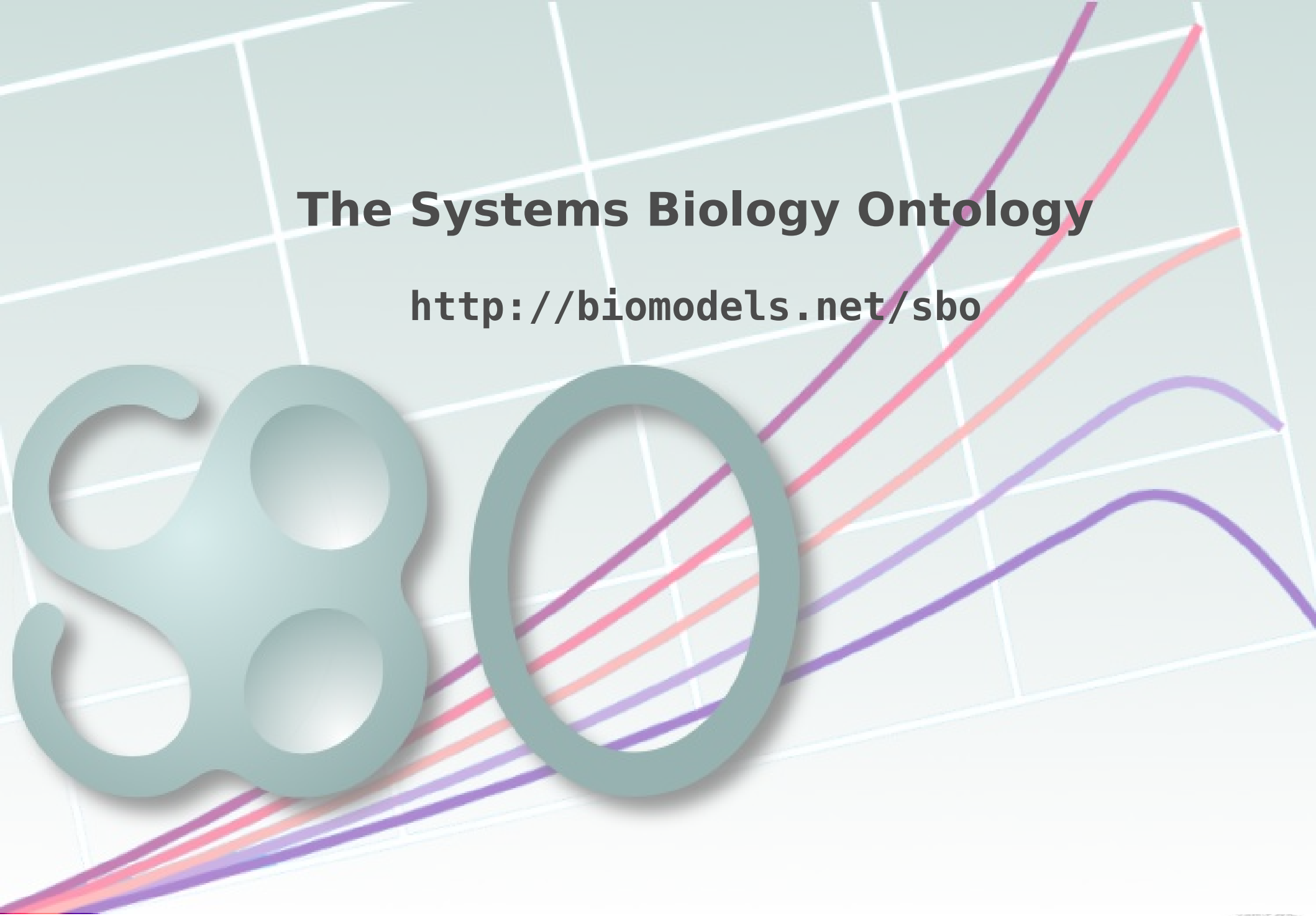
Click any column header to sort the table by that column. The s link to the term request trackers for the listed ontologies.

OBO Foundry candidate ontologies

Title	Domain	Prefix	File	Last changed
Amphibian gross anatomy	anatomy	AAO	amphibian_anatomy.obo	2008/06/19
Amphibian taxonomy	anatomy	ATO	amphibian_taxonomy.obo	
Ascomycete phenotype ontology	phenotype	APO	ascomycete_phenotype.obo	2009/07/10
Biological process	biological process	GO	gene_ontology_edit.obo	2009/08/21
C. elegans development	anatomy	WBls	worm_development.obo	2008/01/31
C. elegans gross anatomy	anatomy	WBbt	WBbt.obo	2009/08/19
C. elegans phenotype	phenotype	WBPhenotype	worm_phenotype.obo	2009/08/19
Systems Biology	biochemistry	SBO	SBO_OBO.obo	

The Systems Biology Ontology

http://biomodels.net/sbo



What is SBO?

- collaborative
- Provide a strictly defined relational vocabulary of terms for use in Systems Biology
- A navigable taxonomic structure of terms that has 'parents', 'children'

Updated SBO structure

6 orthogonal vocabularies:

- entity (macromolecule)
- interaction (transport, reactions)
- mathematical expressions (mass action rate law)
- modelling framework (discrete)
- participant roles (S, P, M)
- quantitative parameters (Hill coefficient)

EBI > SBO > Browsing

Systems Biology Ontology

- SBO:0000000 - sbo term
  - SBO:0000236 - entity
  - SBO:0000231 - interaction
  - SBO:0000064 - mathematical expression
    - SBO:0000355 - conservation law
    - SBO:0000001 - rate law
      - SBO:0000268 - enzymatic rate law
        - SBO:0000150 - enzymatic rate law for irrevers
          - SBO:0000151 - enzymatic rate law for i
          - SBO:0000152 - enzymatic rate law for i
          - SBO:0000028 - enzymatic rate law for i
            - SBO:0000031 - Briggs-Haldane
            - SBO:0000029 - Henri-Michaelis-
            - SBO:0000199 - normalised enzy
            - SBO:0000030 - Van Slyke-Culle
        - SBO:0000269 - enzymatic rate law for unireac
      - SBO:0000192 - Hill-type rate law, generalised form
      - SBO:0000012 - mass action rate law
    - SBO:0000391 - steady state expression
  - SBO:0000004 - modelling framework
  - SBO:0000003 - participant role
  - SBO:0000002 - quantitative parameter

Legend

"is a" relationship

Contact EBI | ©European Bioinformatics Institute 2009. EBI is an Outstation of the E

http://www.ebi.ac.uk - Systems Biology Ontology - SeaMonkey

SBO:0000031

Name
Briggs-Haldane rate law

Definition
Rate-law presented in "G.E. Briggs and J.B.S. Haldane (1925) A note on the kinetics of enzyme action, Biochem. J., 19: 338-339". It is a general rate equation that does not require the restriction of equilibrium of Henri-Michaelis-Menten or irreversible reactions of Van Slyke, but instead make the hypothesis that the complex enzyme-substrate is in quasi-steady-state. Although of the same form than the Henri-Michaelis-Menten equation, it is semantically different since Km now represents a pseudo-equilibrium constant, and is equal to the ratio between the rate of consumption of the complex (sum of dissociation of substrate and generation of product) and the association rate of the enzyme and the substrate.

MathML

```
<math xmlns="http://www.w3.org/1998/Math/MathML">
<semantics definitionURL="http://biomodels.net/SBO/#SBO:0000062">
  <lambda>
    <bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000025">kcat</ci></bva
    <bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000014">Et</ci></bvar>
    <bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000015">S</ci></bvar>
    <bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000371">Km</ci></bvar>
    <apply>
      <divide/>
      <apply>
        <times/>
        <ci>kcat</ci>
        <ci>Et</ci>
        <ci>S</ci>
      </apply>
      <apply>
        <plus/>
```

Rendered equation

$$\lambda(\text{kcat}, \text{Et}, S, \text{Km}) = \frac{\text{kcat} \times \text{Et} \times S}{\text{Km} + S}$$

Miscellaneous
Date of creation:
23 February 2006, 14:00
Date of last modification:
25 November 2008, 16:27

Parent(s)
SBO:0000028 *enzymatic rate law for irreversible non-modulated non-interacting unireactant enzymes* (is a)

Children
This term has no child.

History [+]

Done

Continuous framework

Quantitative parameter

Entity

SBO exports

OBO

```
format-version: 1.2
date: 28:03:2009 07:00
data-version: 26:03:2009 12:18
saved-by: SBO community
auto-generated-by: SBO Browser (http://www.ebi.ac.uk/sbo/)
default-namespace: sbo
```

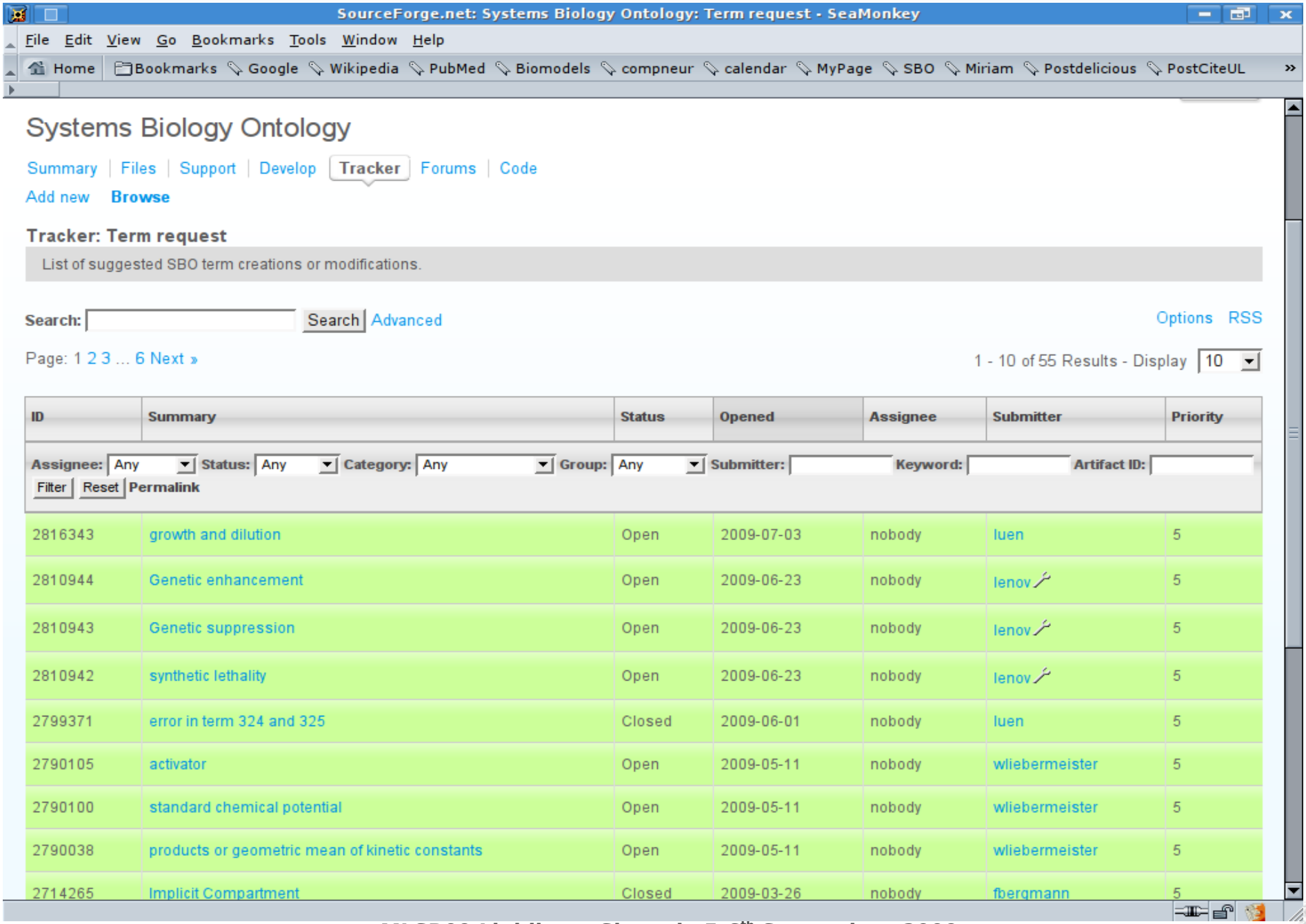
OWL

```
<owl:Ontology rdf:about="">
  <rdfs:comment xml:lang="EN">Systems Biology Ontology. OWL export generated by
  <owl:versionInfo>26:03:2009 12:18</owl:versionInfo>
  <rdfs:label xml:lang="EN">Generated: 28:03:2009 07:00</rdfs:label>
</owl:Ontology>
```

XML

```
<?xml version="1.0" encoding="UTF-8"?>
<sbo xmlns="http://www.biomodels.net/sbo"
    date="2009-03-28T07:00:31.105Z" data-version="2009-03-26T12:18:33.000Z">
    <Term>
        <id>SBO:0000000</id>
```


https://sourceforge.net/projects/sbo



Semantics

Sbo provides essential semantic layer

Semantic layer:

- software interpretation of entity, without intervention
- link between models encoded in SBML and graphical notations (such as SBGN)
- conversion to semantically enriched computing formats (such as BioPAX)
- translation of models between *continuous deterministic frameworks* and *discrete stochastic framework*
- merging/integration of models

SBML and SBO

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" /
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000027"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```


SBML and SBO

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" />
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000027"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

functional compartment

simple chemical

simple chemical

enzyme

catalysis

substrate

product

catalyst

Briggs-Haldane equation

Km

kcat

SBML to BioPAX conversion using SBO

```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" />
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```

GO annotation

Small molecule

Small molecule

Protein

catalysis

physicalEntityParticipant

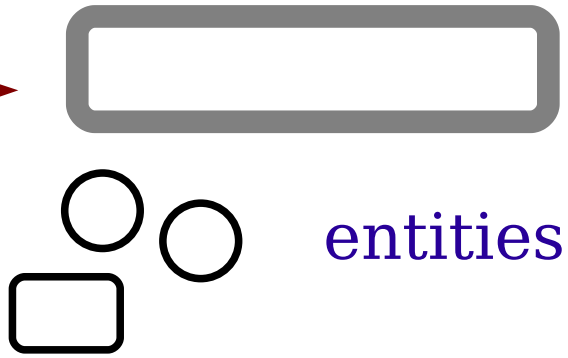
physicalEntityParticipant

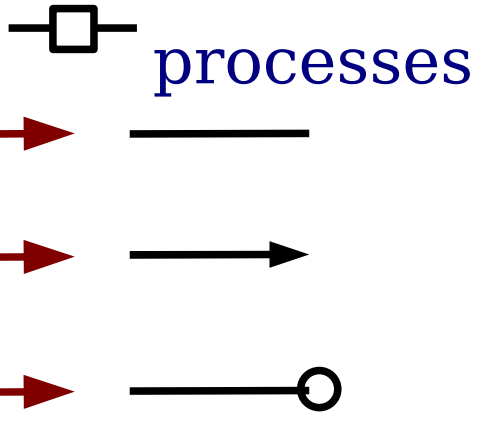
physicalEntityParticipant

Generic vs specific class mapping

SBML to SBGN conversion using SBO

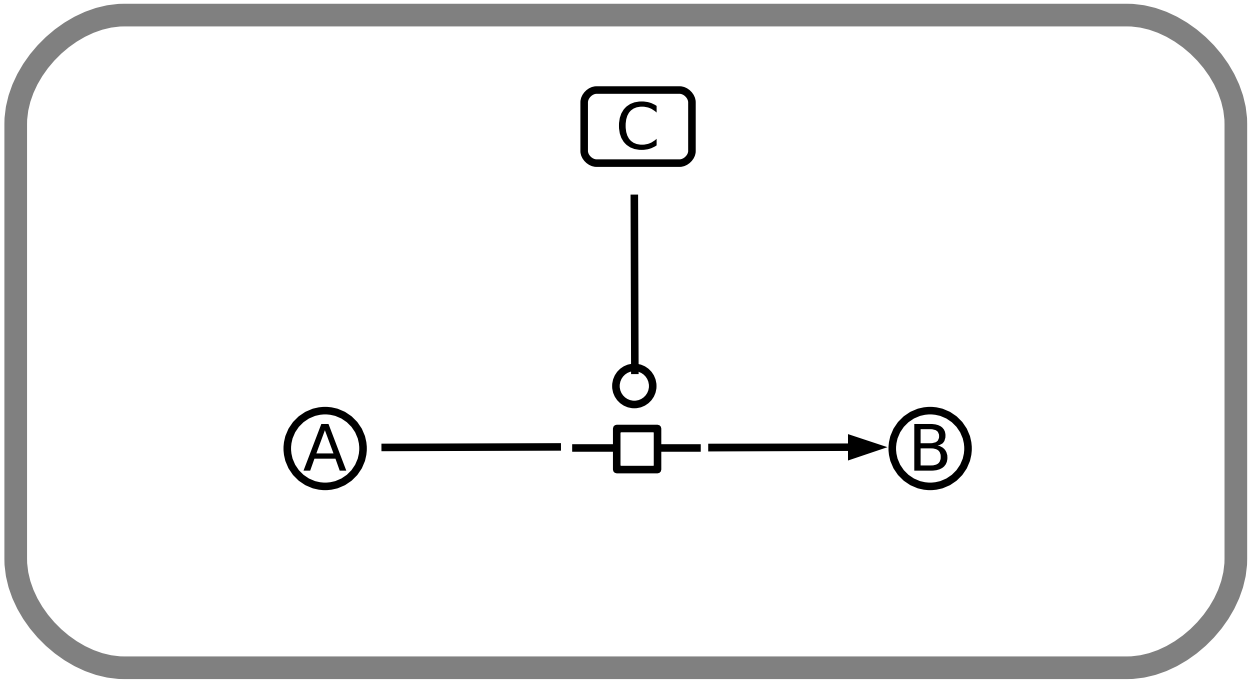
```
<listOfCompartments>
  <compartment id="C" sboTerm="SBO:0000289">
</listOfCompartments>
<listOfSpecies>
  <species id="A" sboTerm="SBO:0000247" />
  <species id="B" sboTerm="SBO:0000247" />
  <species id="C" sboTerm="SBO:0000014" />
</listOfSpecies>
<listOfReactions>
  <reaction sboTerm="SBO:0000172">
    <listOfReactants>
      <speciesReference species="A" sboTerm="SBO:0000015"/>
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="B" sboTerm="SBO:0000011"/>
    </listOfProducts>
    <listOfModifiers>
      <speciesReference species="C" sboTerm="SBO:0000014"/>
    </listOfModifiers>
    <kineticLaw sboTerm="SBO:0000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>
```





http://www.sbgn.org/

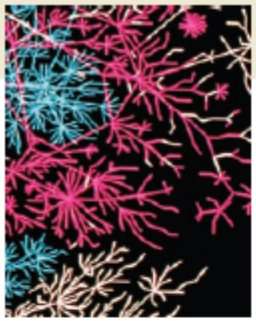
SBML to SBGN conversion using SBO



Simulation

What is a simulation

- A simulation is the instantiation of a model over time, using a given algorithmic approach, and a particular software: A model can beget simulations giving different results!
  - Logical (boolean or discrete) approach
  - Deterministic approach
  - Stochastic approach
  - Fixed timesteps
  - Adaptative timesteps
  - ...
- Plus ... range of simulations
  - parameter scan
  - parameter search/optimisation
  - phase-plane analysis
  - bifurcation analysis
  - ...



_computational BIOLOGY

PERSPECTIVE

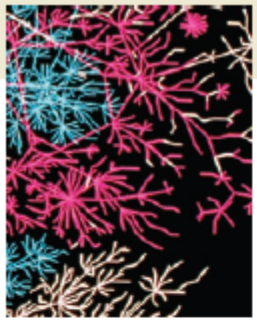
Publishing Group http://www.nature.com/naturebiotechnology

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

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

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[1,2]. 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



_computational BIOLOGY

PERSPECTIVE

6. The model, when instantiated within a suitable simulation environment, must be able to reproduce all relevant results given in the reference description that can readily be simulated. Not only does the simulation have to provide results qualitatively similar to the reference description, such as oscillation, bistability, chaos, but the quantitative values of variables, and their relationships (e.g., the shape of the phase portrait) must be reproduced within some epsilon, the difference being attributable to the algorithms used to run the simulation, and the

Minimu tion of biochem

Nicolas Le Novèr
Julio Collado-Vid en[7],
Herbert Sauro[10],

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

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[1,2]. 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

Minimum Information About a Simulation Experiment (MIASE)

MIASE describes the information needed to run and repeat a numerical simulation experiment derived from a given quantitative model. The project is divided into three parts:

- MIASE - The list of requested information to repeat a simulation result
  - base model & modifications applied
  - simulation task to run on models (algorithms, see KiSAO; simulation parameters)
  - How to post-process the numerical results and to present them

https://sourceforge.net/projects/miase

KiSAO and SED-ML

- KiSAO - Kinetic Simulation Algorithm Ontology
  - Classification of simulation algorithms & methods
  - Definition, literature references
  - Relations between different simulation algorithms & methods

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

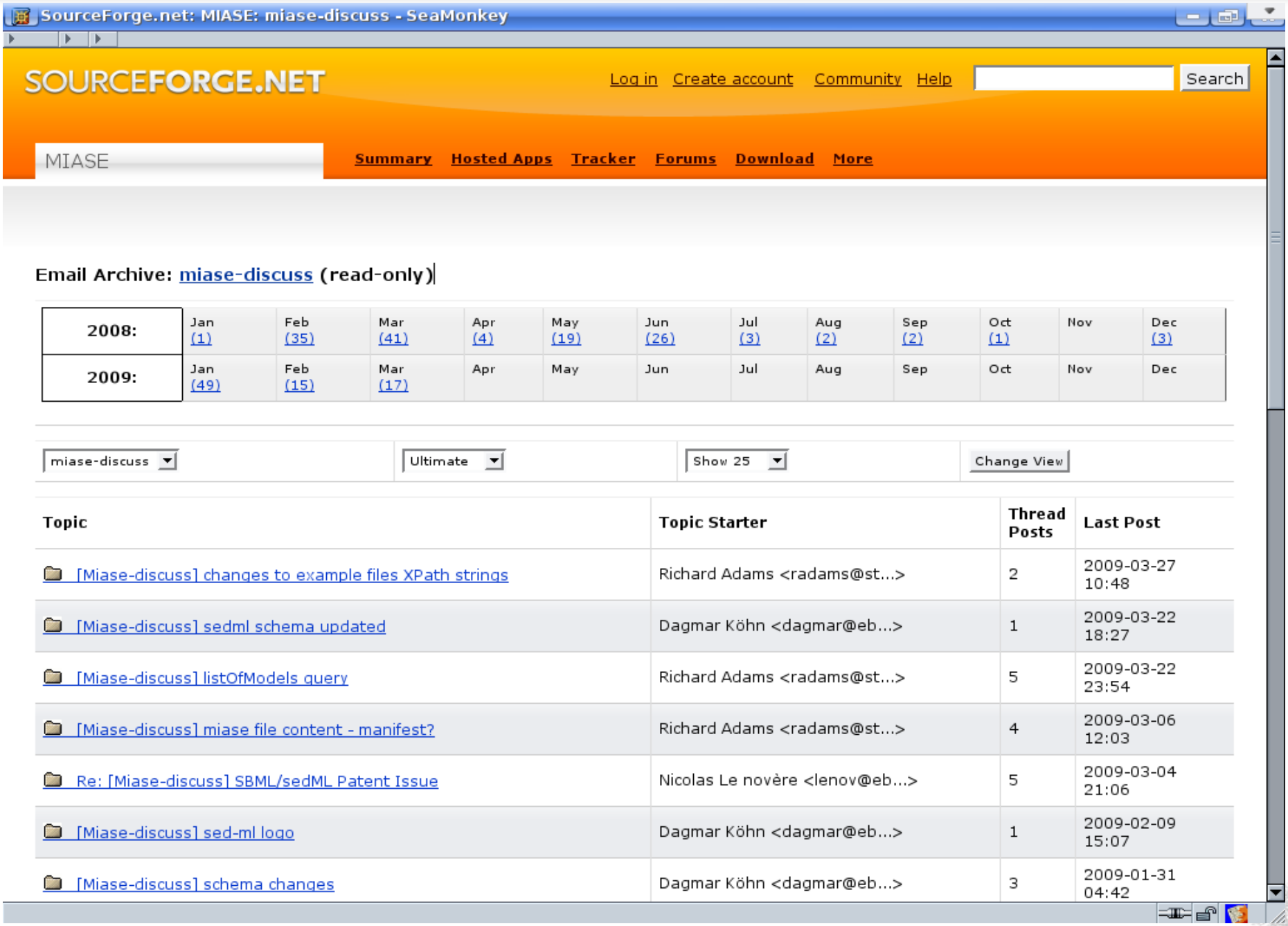
- The Simulation Experiment Description Markup Language (SED-ML)
  - Formal encoding of a subset of MIASE guidelines

http://www.ebi.ac.uk/compneur-srv/sed-ml/

MIASE – Motivation

- Aims
  - researchers should be able to exchange and share their results
- Problem
  - most models plain SBML/CellML/BioPAX
  - curator lacks information about validation simulation settings
  - after curation, the model is made available on biomodels.net
  - the next consumer of model must re-identify the parameters
- Solution:
  - Provide information about the simulation settings along with the model

https://sourceforge.net/projects/miase



More KiSAO motivation

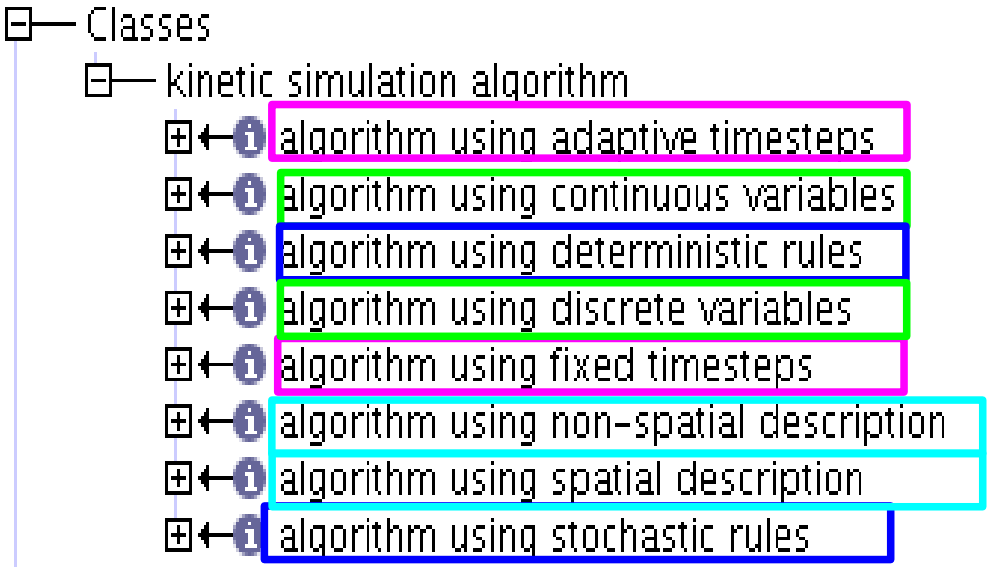
- Users often have to decipher free-text explanations accompanying models
- Authors may provide simulation scripts, which are often software-restricted
- Many algorithms available, with many variants
- Algorithms may be restricted to a limited number of simulation tools (or 1)

Precise identification of simulation approach is required.

Classifies approaches by :

- Model characteristics
  - discrete or continuous variables used in simulation
  - spatial resolution
- Numerical characteristics
  - deterministic or stochastic system behaviour
  - fixed or adaptive time steps

KiSAO branches



Model characteristics

Spatial resolution

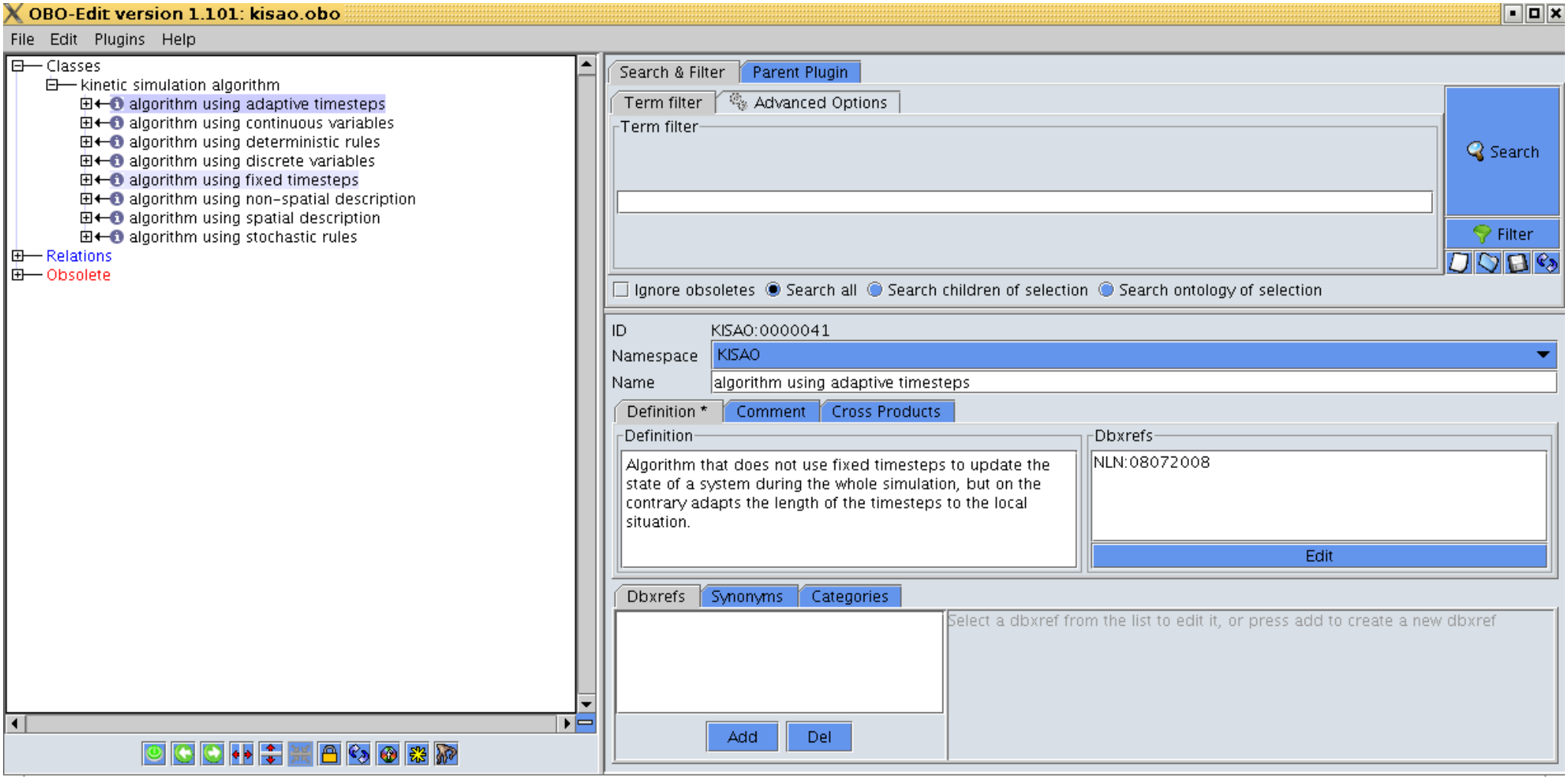
Deterministic / stochastic

Numerical characteristics

Adaptive / fixed timesteps

Continuous / discrete variables

KiSAO term info



SED-ML : Description of models

- MIASE guidelines implemented in SED-ML

```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```


SED-ML : Description of models

```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```

Any model description in XML such as SBML, CellML VCML etc.

SED-ML : Description of models

```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```


BIOMD0000000021 - Leloup1999_CircClock

SBML L2 V1 | Other formats | Actions | Submit Model Comment/Bug

Overview | **Model** | Math | Physical entities | Parameters

Reference Publication

Publication ID: 10366496

J Theor Biol 1999 Jun;198(3):445-59.
Chaos and birhythmicity in a model for circadian oscillations of the PER and TIM proteins in drosophila
Leloup JC, Goldbeter A.
Unite de Chronobiologie Theorique, Faculte des Sciences, Universite Libre de Bruxelles, Campus Plaine, C.P. 231, B-1050 Brussels, Belgium. [more]

Model

Original Model: *Unspecified*

Submitter: Nicolas Le Novère

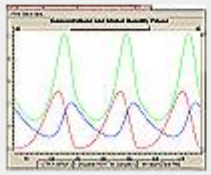
Submission Date: 2005-09-13T13:24:15+00:00

Last Modification Date: 2007-09-25T09:32:00+00:00

Creation Date: 2005-06-29T10:27:52+00:00

Creators: Nicolas Le Novère
Bruce Shapiro

Curation Result:



set #1	bqbiol:is	**Taxonomy** Drosophila melanogaster **KEGG Pathway** dme04710
set #2	bqbiol:isVersionOf	**Gene Ontology** circadian rhythm

http://biomodels.net/database

Notes

Description of models

```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```


Description of models

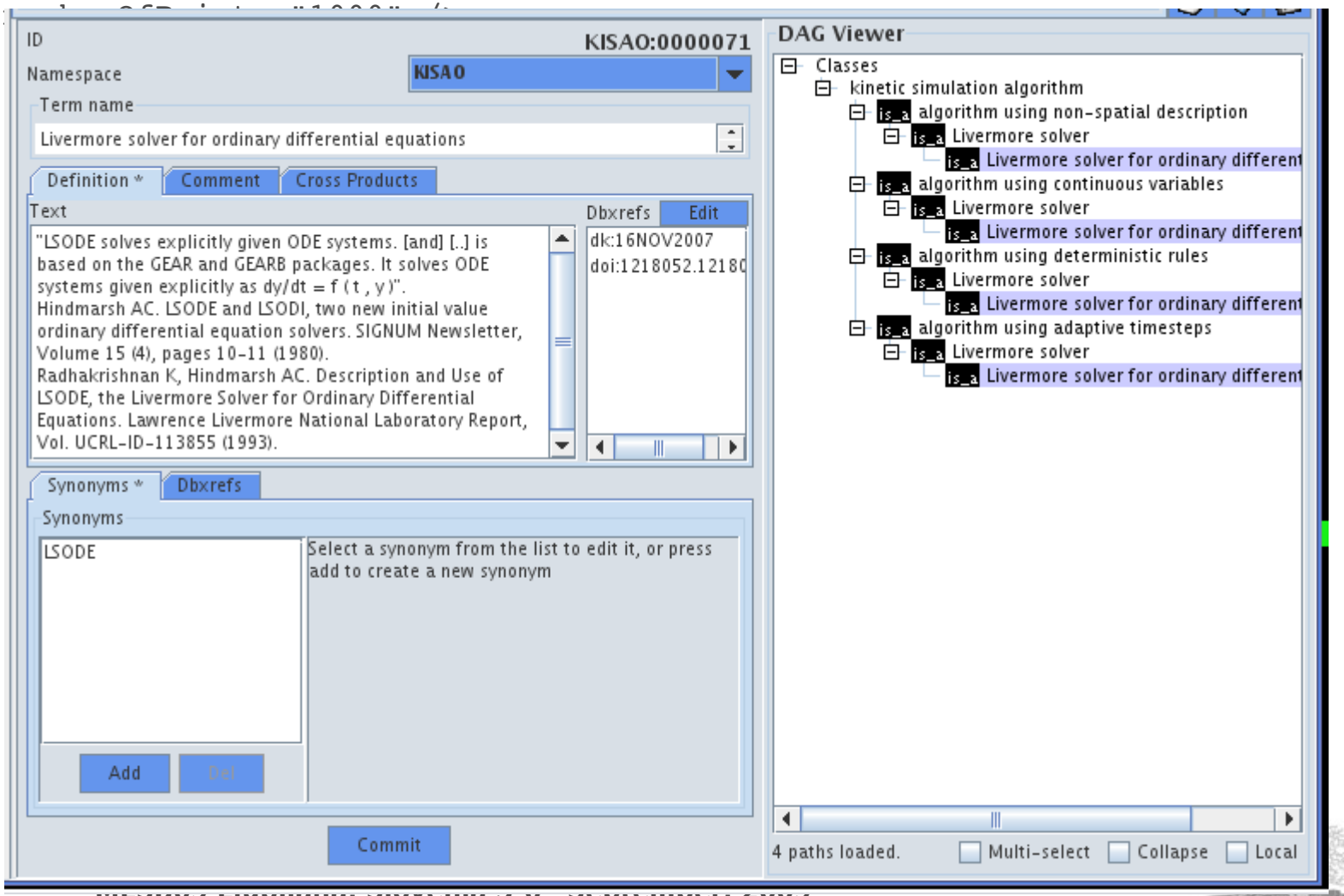
```
<?xml version="1.0" encoding="utf-8"?>
<sedML version="1.0" xmlns="http://www.miase.org/">
  <notes>Changing a system from oscillation to chaos</notes>
  <listOfModels>
    <model id="model1"
           name="Circadian Oscillations"
           type="SBML"
           source="urn:miriam:biomodels.db:BIOMD0000000021" />
    <model id="model2"
           name="Circadian Chaos"
           type="SBML"
           source="model1">
      <listOfChanges>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_mT']/@value" newValue="0.28">
        </changeAttribute>
        <changeAttribute target=
             "/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
        </changeAttribute>
      </listOfChanges>
    </model>
  </listOfModels>
```


Simulation approach

```
<listOfSimulations>
  <uniformTimeCourse id="simulation1"
                     algorithm="KiSAO:0000071"
                     initialTime="0"
                     outputStartTime="50"
                     outputEndTime="1000"
                     numberOfPoints="1000" />
</listOfSimulations>
```


Simulation approach

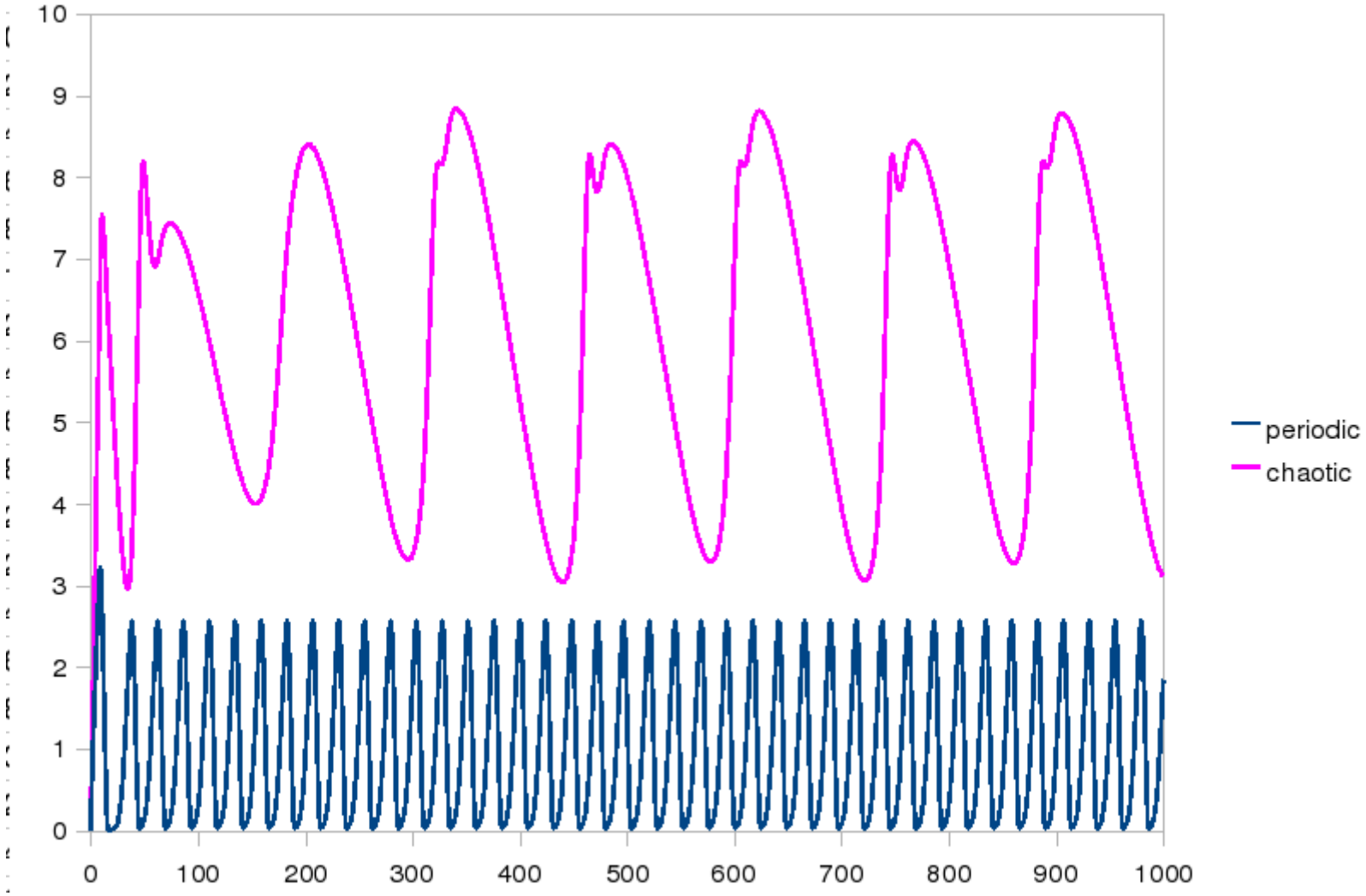
```
<listOfSimulations>
  <uniformTimeCourse id="simulation1"
                     algorithm="KiSAO:0000071"
                     initialTime="0"
                     outputStartTime="50"
                     outputEndTime="1000"
</listOfSimulations>
```



Simulation task

```
<listOfTasks>
  <task id="task1"
        name="Baseline"
        modelReference="model1"
        simulationReference="simulation1">
  </task>
  <task id="task2"
        name="Modified parameters"
        modelReference="model2"
        simulationReference="simulation1">
  </task>
</listOfTasks>
```


Production of results



Acknowledgements

SBO + MIRIAM

- Mélanie Courtot
- Camille Laibe
- Nicolas Le Novère
- Lukas Endler

SBML team

- Michael Hucka
- Sarah Keating

MIASE

- Dagmar Koehn

- BioModels Database developers and curators

The Systems Biology community for their contributions, software support and their comments.