Metadata for Systems Biology

Nick Juty, EMBL-EBI
Which models do we mean?

- biochemical model
- mathematical model
- computational model
- simulation


\[
\begin{align*}
\frac{d[C2]}{dt} &= k_6[M] - k_8[\sim P][C2] + k_9[CP] \\
\frac{d[CP]}{dt} &= -k_3[CP][Y] + k_8[\sim P][C2] - k_9[CP] \\
\frac{d[pM]}{dt} &= k_{31}[CP][Y] - [pM]F([M]) + k_5[\sim P][M] \\
\frac{d[M]}{dt} &= [pM]F([M]) - k_5[\sim P][M] - k_6[M] \\
\frac{d[Y]}{dt} &= k_1[aa] - k_2[Y] - k_3[CP][Y] \\
\frac{d[YP]}{dt} &= k_6[M] - k_9[YP]
\end{align*}
\]

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

<table>
<thead>
<tr>
<th>Format</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBML</td>
<td><a href="http://sbml.org/">http://sbml.org/</a></td>
</tr>
<tr>
<td>CellML</td>
<td><a href="http://www.cellml.org/">http://www.cellml.org/</a></td>
</tr>
<tr>
<td>BioPAX</td>
<td><a href="http://www.biopax.org/">http://www.biopax.org/</a></td>
</tr>
</tbody>
</table>
### Common formats for model storage and exchange

<table>
<thead>
<tr>
<th>Format</th>
<th>Website</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBML</td>
<td><a href="http://sbml.org/">http://sbml.org/</a></td>
<td>quantitative</td>
</tr>
<tr>
<td>CellML</td>
<td><a href="http://www.cellml.org/">http://www.cellml.org/</a></td>
<td>quantitative</td>
</tr>
<tr>
<td>BioPAX</td>
<td><a href="http://www.biopax.org/">http://www.biopax.org/</a></td>
<td>kinetic unsupported</td>
</tr>
</tbody>
</table>

#### General and abstract – range of problems

- inference of components
- automated conversion
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.
- 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
<?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 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.

<table>
<thead>
<tr>
<th>Capabilities</th>
<th>Frameworks</th>
<th>API</th>
<th>Dep.</th>
<th>Platforms</th>
<th>SBML</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creation</td>
<td>SBW</td>
<td>W</td>
<td>L, W, M, B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation</td>
<td>Database</td>
<td></td>
<td>L, W</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis</td>
<td>SBW</td>
<td></td>
<td>L, W, M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDE</td>
<td>Stochastic</td>
<td></td>
<td>C, C++</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDE</td>
<td>Events</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDE</td>
<td>Logical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FDE</td>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Betawb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BinaM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BinaM Cytoscape Plugin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biocam</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biocam CHARON</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biological Networks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BioGrid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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, itself, is a framework.
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

adds meaning to

Ontology

EMBL-EBI

MLSB09 Ljubljana, Slovenia 5-6th September, 2009
Minimal requirements

implemented by

Example Data-model

adds meaning to

Ontology

Models  Simulation  Results

MIAM  MIASE  ?

SBML  SED-ML  SBRML

MLSB09 Ljubljana, Slovenia 5-6th September, 2009
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
Global coordination of reporting guidelines

COMMENTARY

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

Chris F Taylor1,2, Dawn Field2,3, Susanna-Assunta Sansone4,2, Jan Aerts4, Rolf Apweiler1, Michael Ashburner5, Catherine A Ball6, Pierre-Alain Binz7,8, Molly Bogue9, Tim Booth2, Alvis Brazma1, Ryan R Brinkman10, Adam Michael Clark11, Eric W Deutsch12, Oliver Fiehn13, Jennifer Fostel14, Peter Ghazal15, Frank Gibson16, Tanya Gray2,3, Graeme Grimes15, John M Hancock17, Nigel W Hardy18, Henning Hermjakob1, Randall K Julian Jr19, Matthew Kane20, Carsten Kettner21, Christopher Kinsinger22, Eugene Kolker23,24, Martin Kuiper25, Nicolas Le Novère1, Jim Leebens-Mack26, Suzanna E Lewis27, Phillip Lord16, Ann-Marie Mallon17, Nishanth Marthandan28, Hiroshi Masuya29, Ruth McNally30, Alexander Mehrle31, Norman Morrison2,32, Sandra Orchard1, John Quackenbush33, James M Reecy34, Donald G Robertson8, Philippe Rocca-Serra1,36, Henry Rodriguez22, Heiko Rosenfelder31, Javier Sancho-Lope15, Richard H Schuetzmann28, Daniel Schober1, Barry Smith37, Jason Snape38, Christian J Stoeckert Jr39, Keith Tipton40, Peter Sterk1, Andreas Utergasser41, Jo Vandesompele42 & Stefan Wiemann31

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.

MLSB09 Ljubljana, Slovenia 5-6th September, 2009
Minimal requirements

implemented by

Example Data-model

adds meaning to

Ontology

Models    Simulation    Results

MIRIAM    MIASE    ?

implemented by

MIBBI

SBML    SED-ML    SBRML

adds meaning to

SBIO    KISAO    TEDDY
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

The MIBBI Foundry
Towards the next generation of MI guidelines for the biosciences

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

MIBBI search
A Google™ Custom

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

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

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

http://www.mibbi.org
<table>
<thead>
<tr>
<th>MIAME</th>
<th>Minimum Information About a Microarray Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIAME/Env</td>
<td>MIAME / Environmental transcriptomic experiment</td>
</tr>
<tr>
<td>MIAME/Nutr</td>
<td>MIAME / Nutrigenomics</td>
</tr>
<tr>
<td>MIAME/Plant</td>
<td>MIAME / Plant transcriptomics</td>
</tr>
<tr>
<td>MIAME/Tox</td>
<td>MIAME / Toxicogenomics</td>
</tr>
<tr>
<td>MIAPA</td>
<td>Minimum Information About a Phylogenetic Analysis</td>
</tr>
<tr>
<td>MIAPAR</td>
<td>Minimum Information About a Protein Affinity Reagent</td>
</tr>
<tr>
<td>MIAPE</td>
<td>Minimum Information About a Proteomics Experiment</td>
</tr>
<tr>
<td>MIARE</td>
<td>Minimum Information About a RNAi Experiment</td>
</tr>
<tr>
<td>MIASE</td>
<td>Minimum Information About a Simulation Experiment</td>
</tr>
<tr>
<td>MIENS</td>
<td>Minimum Information about an ENvironmental Sequence</td>
</tr>
<tr>
<td>MIFlowCyt</td>
<td>Minimum Information for a Flow Cytometry Experiment</td>
</tr>
<tr>
<td>MiGen</td>
<td>Minimum Information about a Genotyping Experiment</td>
</tr>
<tr>
<td>MIGS</td>
<td>Minimum Information about a Genome Sequence</td>
</tr>
<tr>
<td>MIMIx</td>
<td>Minimum Information about a Molecular Interaction Experiment</td>
</tr>
<tr>
<td>MIMPP</td>
<td>Minimal Information for Mouse Phenotyping Procedures</td>
</tr>
<tr>
<td>MINI</td>
<td>Minimum Information about a Neuroscience Investigation</td>
</tr>
<tr>
<td>MINIMESS</td>
<td>Minimal Metagenome Sequence Analysis Standard</td>
</tr>
<tr>
<td>MINSEQE</td>
<td>Minimum Information about a high-throughput Sequencing Experiment</td>
</tr>
<tr>
<td>MIPFE</td>
<td>Minimal Information for Protein Functional Evaluation</td>
</tr>
<tr>
<td>MIQAS</td>
<td>Minimal Information for QTLs and Association Studies</td>
</tr>
<tr>
<td>MiqPCR</td>
<td>Minimum Information about a quantitative Polymerase Chain Reaction experiment</td>
</tr>
<tr>
<td>MIRIAM</td>
<td>Minimal Information Required In the Annotation of biochemical Models</td>
</tr>
<tr>
<td></td>
<td>Minimum Information Specification For In Situ Hybridization and</td>
</tr>
</tbody>
</table>
The Minimum Information Required In the Annotation of a Model

http://biomodels.net/miriam
Minimum information requested in the annotation of biochemical models (MIRIAM)

Nicolas Le Novère\textsuperscript{1,15}, Andrew Finney\textsuperscript{2,15}, Michael Hucka\textsuperscript{3}, Upinder S Bhalla\textsuperscript{4}, Fabien Campagne\textsuperscript{5}, Julio Collado-Vides\textsuperscript{6}, Edmund J Crampin\textsuperscript{7}, Matt Halstead\textsuperscript{7}, Edda Klipp\textsuperscript{8}, Pedro Mendes\textsuperscript{9}, Poul Nielsen\textsuperscript{7}, Herbert Sauro\textsuperscript{10}, Bruce Shapiro\textsuperscript{11}, Jacky L Snoep\textsuperscript{12}, Hugh D Spence\textsuperscript{13} & Barry L Wanner\textsuperscript{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\textsuperscript{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
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"
  /db_xref="GOA:P49581"

- 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"
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 uncontrolled

URLs

- http://srs6.ebi.ac.uk/srs6bin/cgi-bin/wgetz?[swissprot-AccNumber:P01133]+-e
- 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
Data-type identifier (required)

URI

Not a URL, not a "Web-address"!

Corresponds to a namespace
Corresponds to a namespace

```
urn:miriam:uniprot
```
Data-type identifier (required)

URI

Not a URL, not a “Web-address”!

Data-set Identifier (required)

text string

Format depends on the resource identified by the data-type
UniProt P62158 (human calmodulin)  urn:miriam:uniprot:P62158
EC code 1.1.1.1 (alcohol dehydrogenase)   urn:miriam:ec-code:1.1.1.1
defines relationship between model entity and annotation
<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 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/
<table>
<thead>
<tr>
<th>Name</th>
<th>URI</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>3DMET</td>
<td>urn:miriam:3dmet</td>
<td>3DMET is a database collecting three-dimensional structures of natural metabolites.</td>
</tr>
<tr>
<td>Aclame</td>
<td>urn:miriam:acalam</td>
<td>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.</td>
</tr>
<tr>
<td>ArrayExpress</td>
<td>urn:miriam:arrayexpress</td>
<td>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.</td>
</tr>
<tr>
<td>arXiv</td>
<td>urn:miriam:arxiv</td>
<td>arXiv is an e-print service in the fields of physics, mathematics, non-linear science, computer science, and quantitative biology.</td>
</tr>
<tr>
<td>BioModels Database</td>
<td>urn:miriam:biomodels.db</td>
<td>BioModels Database is a data resource that allows biologists to search, store and retrieve published mathematical models of biological interests.</td>
</tr>
<tr>
<td>BRENDA</td>
<td>urn:miriam:brenda</td>
<td>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.</td>
</tr>
<tr>
<td>ChEBI</td>
<td>urn:miriam:obo.chebi</td>
<td>Chemical Entities of Biological Interest (ChEBI) is a freely available dictionary of molecular entities focused on 'small' chemical compounds.</td>
</tr>
<tr>
<td>CluSTr</td>
<td>urn:miriam:clustr</td>
<td>The CluSTr database offers an automatic classification of UniProt Knowledgebase and IPD proteins into groups of related proteins. The clustering is based on analysis of all pairwise comparisons (Smith-Waterman) between protein sequences.</td>
</tr>
<tr>
<td>Database of</td>
<td>urn:miriam:dip</td>
<td>The database of interacting proteins (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.</td>
</tr>
<tr>
<td>Interacting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOI</td>
<td>urn:miriam:doi</td>
<td>The Digital Object Identifier System is for identifying content objects in the digital environment.</td>
</tr>
<tr>
<td>Ensembl</td>
<td>urn:miriam:ensembl</td>
<td>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.</td>
</tr>
<tr>
<td>Entrez Gene</td>
<td>urn:miriam:entrez.gene</td>
<td>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.</td>
</tr>
<tr>
<td>Enzyme Nomenclature</td>
<td>urn:miriam:ec-code</td>
<td>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.</td>
</tr>
<tr>
<td>Evidence Code</td>
<td>urn:miriam:obo.eco</td>
<td>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.</td>
</tr>
<tr>
<td>FlyBase</td>
<td>urn:miriam:flybase</td>
<td>FlyBase is the database of the Drosophila Genome Projects and of associated literature.</td>
</tr>
<tr>
<td>FMA</td>
<td>urn:miriam:obo.fma</td>
<td>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.</td>
</tr>
<tr>
<td>Gene Ontology</td>
<td>urn:miriam:obo.go</td>
<td>The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism.</td>
</tr>
</tbody>
</table>
## Enzyme Nomenclature

### General Information
- **Identifier**: MIR:00000004
- **Name**: Enzyme Nomenclature
- **EC code**: Enzyme Classification

### URIs
- **Official URN**: urn:miriam:ec-code
- **Deprecated URN**:
  - urn:miriam:isctec-code
  - http://www.ebi.ac.uk/IntEnz/

### 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

### Physical Locations

<table>
<thead>
<tr>
<th>Resource #1</th>
<th>Data Entry</th>
<th><a href="http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&amp;ec=EC%5C$id">http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&amp;ec=EC\$id</a> <a href="#">Example: 1.1.1.1</a></th>
<th>Data Resource</th>
<th><a href="http://www.ebi.ac.uk/intenz/">http://www.ebi.ac.uk/intenz/</a></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Information</td>
<td>IntEnZ (Integrated relational Enzyme database)</td>
<td>Institution</td>
<td>European Bioinformatics Institute, United Kingdom</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Information</td>
<td>KEGG Ligand Database for Enzyme Nomenclature</td>
<td>Institution</td>
<td>Kyoto University Bioinformatics Center, Japan</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Information</td>
<td>Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)</td>
<td>Institution</td>
<td>Swiss Institute of Bioinformatics, Switzerland</td>
</tr>
</tbody>
</table>

### Documentation
- [http://www.chem.gmu.ac.uk/ubmb/enzyme/](http://www.chem.gmu.ac.uk/ubmb/enzyme/)
- [http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-view+MedlineFull+[medline-PMD:10812475]](http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-view+MedlineFull+[medline-PMD:10812475])

### Date of Creation
- 2006-08-14 19:38:06 GMT

### Date of Last Modification
- 2009-05-08 14:59:31 GMT
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)

- **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)

- **MIRIAM Resources statistics**
  - ~5000 web page requests per month
  - ~550,000 web service requests per month

MLSB09 Ljubljana, Slovenia 5-6th September, 2009
SBML gives us:

\[
\begin{align*}
\frac{d[C2]}{dt} &= k_6[M] - k_8[\sim P][C2] + k_9[CP] \\
\frac{d[CP]}{dt} &= -k_3[CP][Y] + k_8[\sim P][C2] - k_9[CP] \\
\frac{d[pM]}{dt} &= k_3[CP][Y] - [pM]F([M]) + k_5[\sim P][M] \\
\frac{d[M]}{dt} &= [pM]F([M]) - k_5[\sim P][M] - k_6[M] \\
\frac{d[Y]}{dt} &= k_1[aa] - k_2[Y] - k_3[CP][Y] \\
\frac{d[YP]}{dt} &= k_6[M] - k_7[YP]
\end{align*}
\]
The model revealed

\[ \frac{d[C2]}{dt} = k_6[M] - k_8[\sim P][C2] + k_9[CP] \]
\[ \frac{d[CP]}{dt} = -k_3[CP][Y] + k_8[\sim P][C2] - k_9[CP] \]
\[ \frac{d[pM]}{dt} = k_3[CP][Y] - [pM]F([M]) + k_5[\sim P][M] \]
\[ \frac{d[M]}{dt} = [pM]F([M]) - k_5[\sim P][M] - k_6[M] \]
\[ \frac{d[Y]}{dt} = k_1[aa] - k_2[Y] - k_3[CP][Y] \]
\[ \frac{d[YP]}{dt} = k_6[M] - k_7[YP] \]
Lack of biological semantics in SBML

“what” has been answered by MIRIAM annotations

Roles and inter-relationships unknown
Adding ontologies
The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration


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

adds meaning to

Ontology

Models | Simulation | Results

MIRIAM | MIASE | ?

S8ML | SED-ML | SBRML

SBO | KISAO | TEDDY

OBO

MLSB09 Ljubljana, Slovenia 5-6th September, 2009
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 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.

<table>
<thead>
<tr>
<th>Title</th>
<th>Domain</th>
<th>Prefix</th>
<th>File</th>
<th>Last changed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amphibian gross anatomy</td>
<td>anatomy</td>
<td>AAO</td>
<td>amphibian_anatomy.obo</td>
<td>2009/06/19</td>
</tr>
<tr>
<td>Amphibian taxonomy</td>
<td>anatomy</td>
<td>ATO</td>
<td>amphibian_taxonomy.obo</td>
<td></td>
</tr>
<tr>
<td>Ascomycete phenotype ontology</td>
<td>phenotype</td>
<td>APO</td>
<td>ascomycete_phenotype.obo</td>
<td>2009/07/10</td>
</tr>
<tr>
<td>Biological process</td>
<td>biological process</td>
<td>GO</td>
<td>genes_ontology_edit.obo</td>
<td>2009/08/21</td>
</tr>
<tr>
<td>C. elegans development</td>
<td>anatomy</td>
<td>WBke</td>
<td>worm_development.obo</td>
<td>2009/08/31</td>
</tr>
<tr>
<td>C. elegans gross anatomy</td>
<td>anatomy</td>
<td>WBbt</td>
<td>WBBt.obo</td>
<td>2009/08/19</td>
</tr>
<tr>
<td>C. elegans phenotype</td>
<td>phenotype</td>
<td>WBP Phenotype</td>
<td>worm_phenotype.obo</td>
<td>2009/08/19</td>
</tr>
<tr>
<td>Cell cycle</td>
<td>anatomy</td>
<td>Cell</td>
<td>Cell cycle.obo</td>
<td>2009/04/02</td>
</tr>
<tr>
<td>Suggested Ontology for Pharmacogenomics</td>
<td>health</td>
<td>SOPHARM</td>
<td>pharmacogenomics.obo</td>
<td>2009/08/03</td>
</tr>
<tr>
<td>Systems Biology</td>
<td>biochemistry</td>
<td>$BO</td>
<td>$BO_OBO.obo</td>
<td></td>
</tr>
</tbody>
</table>
The Systems Biology Ontology

http://biomodels.net/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'
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)
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.biomiodels.net/sbo"

date="2009-03-28T07:00:31.105Z" data-version="2009-03-26T12:18:33.000Z">
<Term>
  <id>SBO:0000000</id>
</Term>
## Systems Biology Ontology

### Tracker: Term request

List of suggested SBO term creations or modifications.

<table>
<thead>
<tr>
<th>ID</th>
<th>Summary</th>
<th>Status</th>
<th>Opened</th>
<th>Assignee</th>
<th>Submitter</th>
<th>Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016343</td>
<td>growth and dilution</td>
<td>Open</td>
<td>2009-07-03</td>
<td>nobody</td>
<td>luen</td>
<td>5</td>
</tr>
<tr>
<td>2610944</td>
<td>Genetic enhancement</td>
<td>Open</td>
<td>2009-06-23</td>
<td>nobody</td>
<td>lenov</td>
<td>5</td>
</tr>
<tr>
<td>2610943</td>
<td>Genetic suppression</td>
<td>Open</td>
<td>2009-06-23</td>
<td>nobody</td>
<td>lenov</td>
<td>5</td>
</tr>
<tr>
<td>2610942</td>
<td>synthetic lethality</td>
<td>Open</td>
<td>2009-06-23</td>
<td>nobody</td>
<td>lenov</td>
<td>5</td>
</tr>
<tr>
<td>2799371</td>
<td>error in term 324 and 325</td>
<td>Closed</td>
<td>2009-06-01</td>
<td>nobody</td>
<td>luen</td>
<td>5</td>
</tr>
<tr>
<td>2790105</td>
<td>activator</td>
<td>Open</td>
<td>2009-05-11</td>
<td>nobody</td>
<td>wliemhofer</td>
<td>5</td>
</tr>
<tr>
<td>2790100</td>
<td>standard chemical potential</td>
<td>Open</td>
<td>2009-05-11</td>
<td>nobody</td>
<td>wliemhofer</td>
<td>5</td>
</tr>
<tr>
<td>2790038</td>
<td>products or geometric mean of kinetic constants</td>
<td>Open</td>
<td>2009-05-11</td>
<td>nobody</td>
<td>wliemhofer</td>
<td>5</td>
</tr>
<tr>
<td>2714205</td>
<td>Implicit Compartment</td>
<td>Closed</td>
<td>2009-03-26</td>
<td>nobody</td>
<td>Forgmann</td>
<td>5</td>
</tr>
</tbody>
</table>
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
<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>
<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

Functional compartment
Simple chemical
Simple chemical
Enzyme
Catalysis
Substrate
Product
Catalyst
Briggs-Haldane equation
Km
Kcat

SBML and SBO

MLSB09 Ljubljana, Slovenia 5-6th September, 2009
<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:000031">
      <listOfParameters>
        <parameter id="U" sboTerm="SBO:0000008"/>
        <parameter id="V" sboTerm="SBO:0000025"/>
      </listOfParameters>
    </kineticLaw>
  </reaction>
</listOfReactions>

http://www.sbgn.org/
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
  - ...

MLSB09 Ljubljana, Slovenia 5-6th September, 2009
Minimum information requested in the annotation of biochemical models (MIRIAM)

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

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

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. 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
- 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
Model characteristics

Spatial resolution

Deterministic / stochastic

Numerical characteristics

Adaptive / fixed timesteps

Continuous / discrete variables
• MIASE guidelines implemented in SED-ML

<?xml version="1.0" encoding="utf­8"?>
<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 target="'/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8">
      </listOfChanges>
    </model>
  </listOfModels>
</sedML>
Any model description in XML such as SBML, CellML, VCML etc.
<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 target="/sbml/model/listOfParameters/parameter[@id='V_dT']/@value"
          newValue="4.8"/>
      </listOfChanges>
    </model>
  </listOfModels>
</sedML>
<?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 target="/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8"/>
      </listOfChanges>
    </model>
  </listOfModels>
</sedML>

http://biomodels.net/database
<?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>
</sedML>
<?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 target="/sbml/model/listOfParameters/parameter[@id='V_dT']/@value" newValue="4.8"/>
      </listOfChanges>
    </model>
  </listOfModels>
</sedML>
Simulation approach

<listOfSimulations>
  <uniformTimeCourse id="simulation1"
    algorithm="KiSAO:0000071"
    initialTime="0"
    outputStartTime="50"
    outputEndTime="1000"
    numberOfPoints="1000" />
</listOfSimulations>
<listOfSimulations>
  <uniformTimeCourse id="simulation1"
    algorithm="KiSAO:0000071"
    initialTime="0"
    outputStartTime="50"
    outputEndTime="1000"
    numberOfPoints="1000"/>
</listOfSimulations>
<listOfTasks>
  <task id="task1"
    name="Baseline"
    modelReference="model1"
    simulationReference="simulation1"/>
  <task id="task2"
    name="Modified parameters"
    modelReference="model2"
    simulationReference="simulation1"/>
  </task>
</listOfTasks>
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.