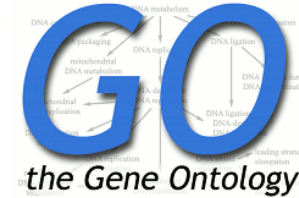
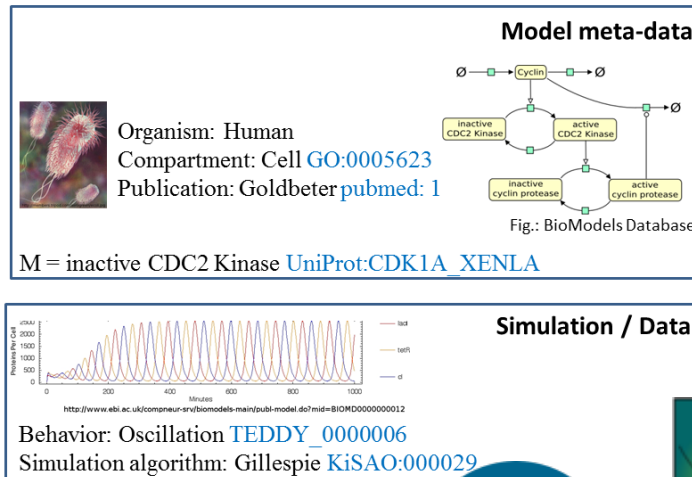
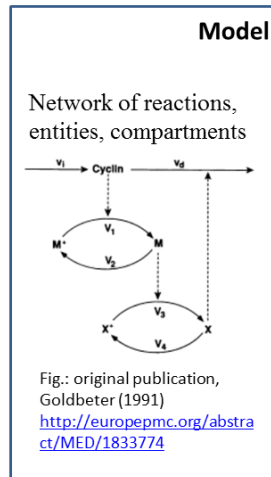


Common annotations for COMBINE efforts

Dagmar Waltemath

Use of meta-data



Search

Curation

Provenance

Graphical representation

Transformation

Model merging

Model comparison

Entity identity

Data links

Virtual experiments

Result reproducibility

MAMO



We have well-specified standards.

MIRIAM guidelines

SBML Core Annotations

SBML L3 Annot
Package Proposal
(2011)

biomodels.net qualifiers

CellML Meta-data v1.0

CellML Meta-data v2.0
(2011)

MIRIAM Annotations

Package Annotations, e.g.
FBA

...

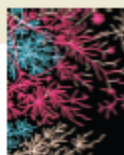
identifiers.org
(2012)

RDF

OWL

vcard

Dublin Core



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

Nicolas Le Novère^{1,15}, Andrew Finney^{2,15}, Michael Hucka³, Upinder S Bhalla⁴, Fabien Campagne⁵, Julio Collado-Vides⁶, Edmund J Crampin⁷, Matt Halstead⁷, Edda Klipp⁸, Pedro Mendes⁹, Poul Nielsen⁷, Herbert Sauro¹⁰, Bruce Shapiro¹¹, Jacky L Snoep¹², Hugh D Spence¹³ & Barry L Wanner¹⁴

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 these 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 application will enable users to (i) have confidence that curated models are an accurate reflection of their associated reference descriptions, (ii) search collections of curated models with precision, (iii) quickly identify the biological phenomena that a given curated model or model constituent represents and (iv) facilitate model reuse and composition into large subcellular models.

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

Box 1 Glossary

Some terms are used in a very specific way throughout the article. We provide here a precise definition of each one.

Quantitative biochemical model. A formal model of a biological system, based on the mathematical description of its molecular and cellular components, and the interactions between those components.

Encoded model. A mathematical model written in a formal machine-readable language, such that it can be systematically parsed and employed by simulation and analysis software without further human translation.

MIRIAM-compliant model. A model that passes all the tests and fulfills all the conditions listed in MIRIAM.

Reference description. A unique document that describes, or references the description of the model, the structure of the model, the numerical values necessary to instantiate a simulation from the model, or to perform a mathematical analysis of the model, and the results one expects from such a simulation or analysis.

Curation process. The process by which the compliance of an encoded model with MIRIAM is achieved and/or verified. The curation process may encompass some or all of the following tasks: encoding of the model, verification of the reference correspondence and annotation of the model.

Reference correspondence. The fact that the structure of a model and the results of a simulation or an analysis match the information present in the reference description.

+



META

¹European Bioinformatics Institute, Hinxton, CB10 1SD, UK. ²Physiomics PLC, Magdalen Centre, Oxford Science Park, Oxford, OX4 4GA, UK. ³Control and Dynamical Systems, California Institute of Technology, Pasadena, California 91125, USA. ⁴National Centre for Biological Sciences, TIFR, UAS-GVK Campus, Bangalore 560005, India. ⁵Institute for Computational Biomedicine, Weill Medical College of Cornell University, New York, New York 10021, USA. ⁶Center for Genomic Sciences, Universidad Nacional Autónoma de México, Av. Universidad 485, Cuernavaca, Morelia, 62100, Mexico. ⁷Bioengineering Institute and Department of Engineering Science, The University of Auckland, Private Bag 92019, Auckland, New Zealand. ⁸Max-Planck Institute for Molecular Genetics, Berlin Center for Genome-based Bioinformatics (BCB), Innestr. 73, 14195 Berlin, Germany. ⁹Virginia Bioinformatics Institute, Virginia Tech, Washington St., Blacksburg, Virginia 24061-0477, USA. ¹⁰Rack Graduate Institute, 535 Watson Drive, Claremont, California 91711, USA. ¹¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA. ¹²Triple-J Group for Molecular Cell Physiology, Department of Biochemistry, Stellenbosch University Private Bag X1, Matieland 7602, South Africa. ¹³Department of Scientific Computing & Mathematical Modeling, GlaxoSmithKline Research & Development Limited, Medicines Research Centre, Gunnels Wood Road, Stevenage, Herts, SG2 2NR, UK. ¹⁴Purdue University, Department of Biological Sciences, Lilly Hall of Life Sciences, 915 W. State Street, West Lafayette, Indiana 47907-2054, USA. ¹⁵These authors have contributed equally to the work. Correspondence should be addressed to N.L.N. (e-mail: lenov@ebi.ac.uk).



Guideline - standard - recommendation
for the annotation of
models - simulations - all kinds of data
within the COMBINetwork

Questions

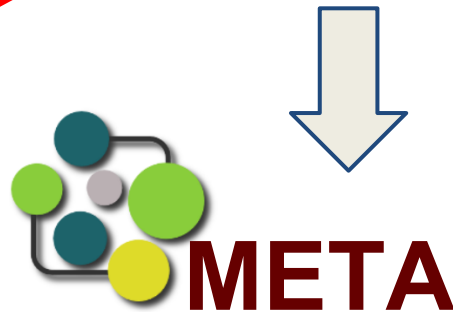
1. Can we drop the SBML Annot proposal in favor of a COMBINE Meta proposal?

[Version 0](#) - package meeting, May 2010

[Version 1](#) – Proposal document, January 2011

SBML Level 3 Package Proposal: Annot

Dagmar Waltemath^{1,*}, Neil Swainston^{2,*}, Allison Lister^{3,*}, Frank Bergmann⁴, Ron Henkel¹, Stefan Hoops⁵, Michael Hucka⁶, Nick Juty⁷, Sarah Keating⁷, Christian Knuepfer⁸, Falko Krause⁹, Camille Laibe⁷, Wolfram Liebermeister⁹, Catherine Lloyd¹⁰, Goksel Misirli³, Marvin Schulz⁹, Morgan Taschuk³, Nicolas Le Novère⁷



META

Questions

2. Should we update the bio- and model-qualifier provided by biomodels.net?

SUBJECT has PREDICATE whose value is OBJECT

e.g. bqmodel:is vs bqmodel:identity

Questions

3. Do we want RDF, or do we not want RDF?

```
< SBML_ELEMENT +++ metaid="SBML_META_ID" +++ >
+++
<annotation>
+++
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:vcard="http://www.w3.org/2001/vcard-rdf/3.0#"
  xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
  xmlns:bqmodel="http://biomodels.net/model-qualifiers/" >
  <rdf:Description rdf:about="#SBML_META_ID">
    [HISTORY]
    <RELATION_ELEMENT>
      <rdf:Bag>
        <rdf:li rdf:resource="URI" />
        ...
      </rdf:Bag>
    </RELATION_ELEMENT>
    ...
  </rdf:Description>
  ...
</rdf:RDF>
+++
</annotation>
+++
</ SBML_ELEMENT >
```

 `<reaction id="E1" reversible="false" sboTerm="SB0:0000182" >`
`<listOfReactants>`
`<speciesReference species="S" sboTerm="SB0:0000015" />`
`</listOfReactants>`
`<listOfProducts>`
`<speciesReference species="P" sboTerm="SB0:0000011" />`
`</listOfProducts>`
`<listOfModifiers>`
`<modifierSpeciesReference species="E" sboTerm="SB0:0000013" />`
`</listOfModifiers>`
`<kineticLaw sboTerm="SB0:0000029">`
`$[...]$`
`<listOfParameters>`
`<parameter id="K" value="1" sboTerm="SB0:0000027" />`
`</listOfParameters>`
`</kineticLaw>`
`</reaction>`

Conversion


Substrate

Product

Catalyst

Henri-Michaelis-Menten rate law

Michaelis constant

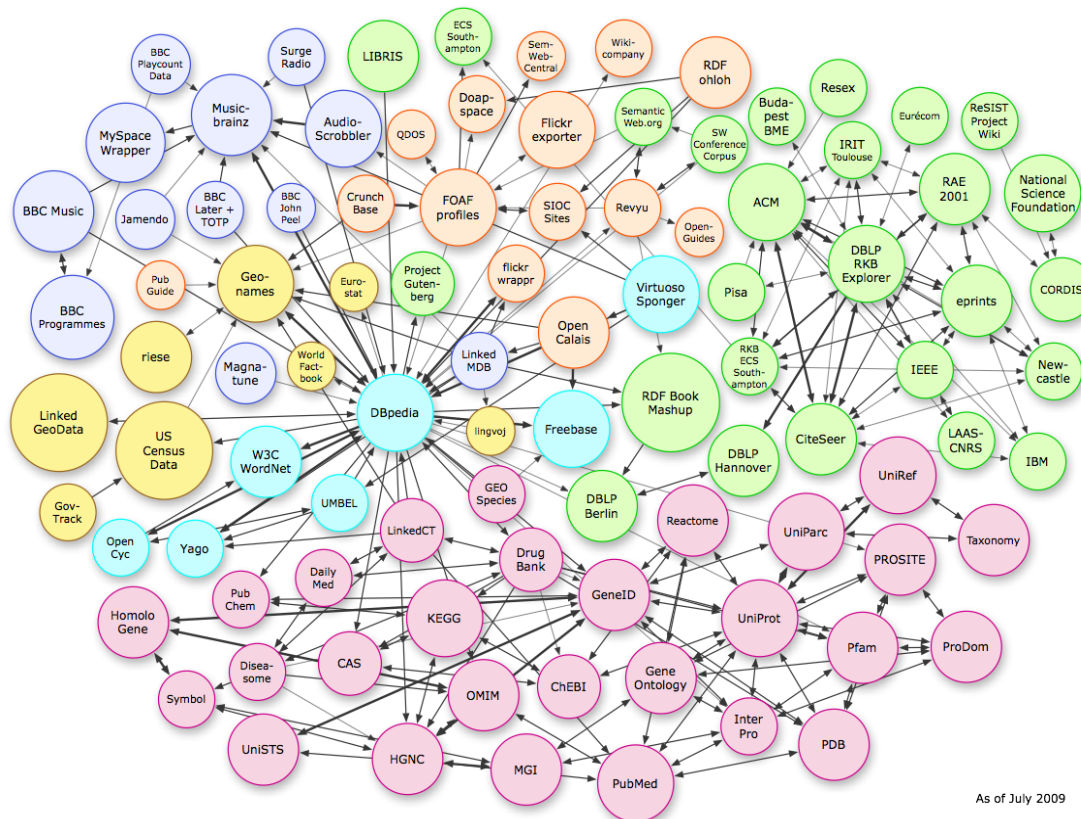
 `<listOfSimulations>`
`<uniformTimeCourse id="sim1"`
`initialTime="0"`
`outputStartTime="0"`
`outputEndTime="5"`
`numberOfPoints="51">`
`<algorithm kisaoID="KISA0:0000019" />`
`</uniformTimeCourse>`

CVODE

<http://www.nature.com/msb/journal/v7/n1/images/msb201177-f2.jpg>

Questions

3. Do we want RDF, or do we not want RDF?



As of July 2009

http://segnalazionit.files.wordpress.com/2009/11/post_d.png

Common annotations for COMBINE efforts | COMBINE, Paris 2013 | D. Waltemath

Questions

4. Would other efforts [beside SBML] contribute and then support such a recommendation?

CellML Metadata 2.0 Draft Specification

This is a proposal for the CellML metadata 2.0 specification written by Mike Cooling, intended for testing and discussion.

- CellML Metadata Framework Core Specification 2.0
- CellML Basic Model Information Metadata Specification 2.0
- CellML Licensing Metadata Specification 2.0
- CellML Citation Metadata Specification 2.0
- CellML Biological Annotation Metadata Specification 2.0

[Metadata Version 1, 2009](#)

[Metadata Version 2, 2011](#)

Last question

5. Would you like to see a working group taking care of these issues? - Would you like to join?

