Presentation by Michael Hucka
at the
Cell Behavior Ontology and Standards for Multicellular Modeling Workshop,
16 Oct. 2010,
Edinburgh, UK.

SBML fly-by

Michael Hucka

(On behalf of many people)

Senior Research Fellow
California Institute of Technology
Pasadena, California, USA

SBML = Systems Biology Markup Language

- Machine-readable format for representing computational models
  - Central concept: **process**
    - Literally called **reaction** in SBML, but can be anything similar
  - Can include
    - Compartments (i.e., where pools of entities are located)
    - Mathematical "extras" (e.g., additional assignments, ODEs, etc.)
    - Discontinuous events

$$n_aA + n_bB \xrightarrow{f([A],[B],[P],\ldots)} n_pP$$

$$n_cC + n_dD + n_eE \xrightarrow{f(\ldots)} n_qQ + n_rR$$

$$\vdots$$

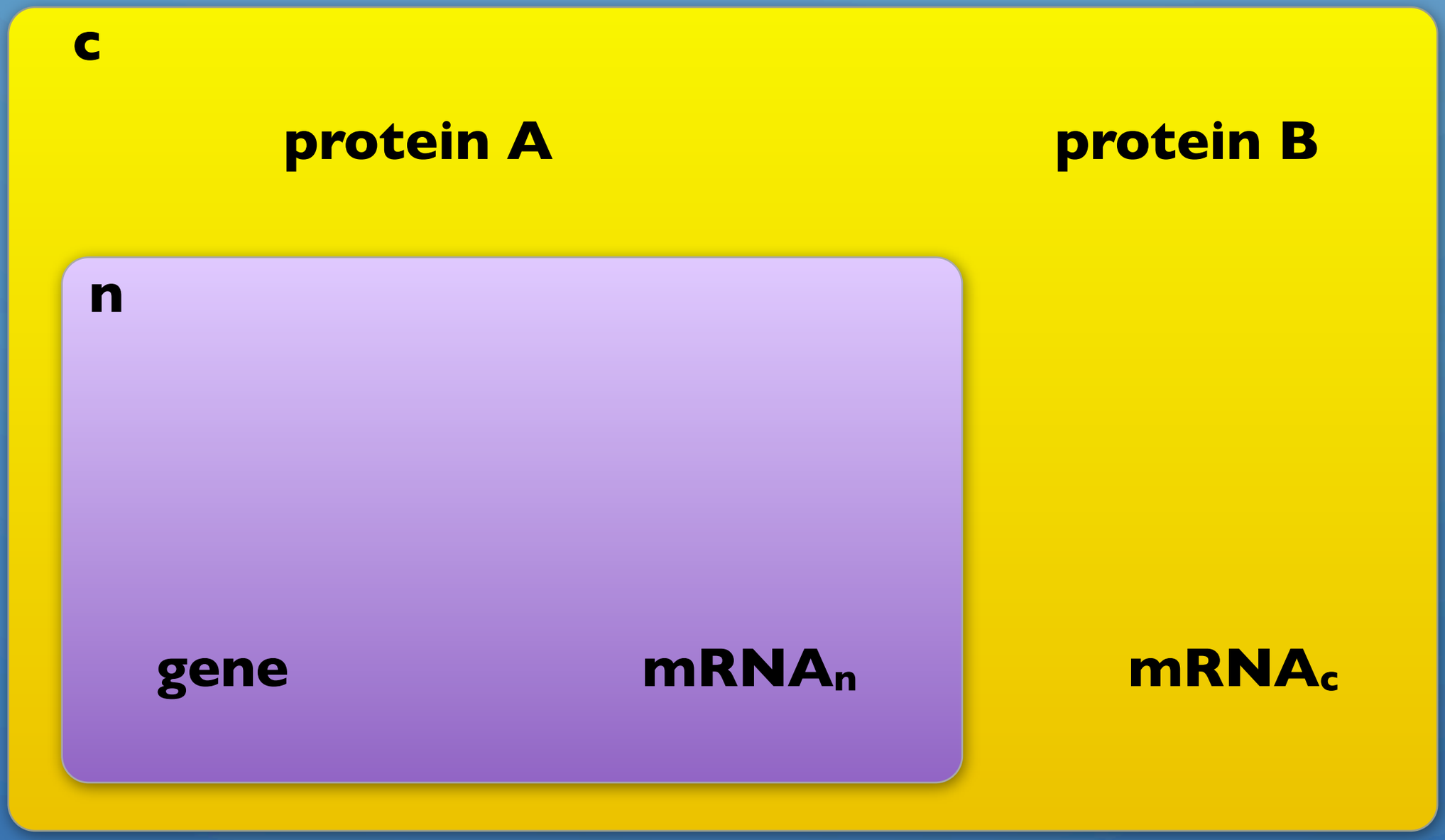
- A declarative description, not procedural
  - Not a script for a simulation
- Not specific to a particular formalism (ODE, stochastic, ...)

Some basics of SBML model encoding

- Well-stirred compartments

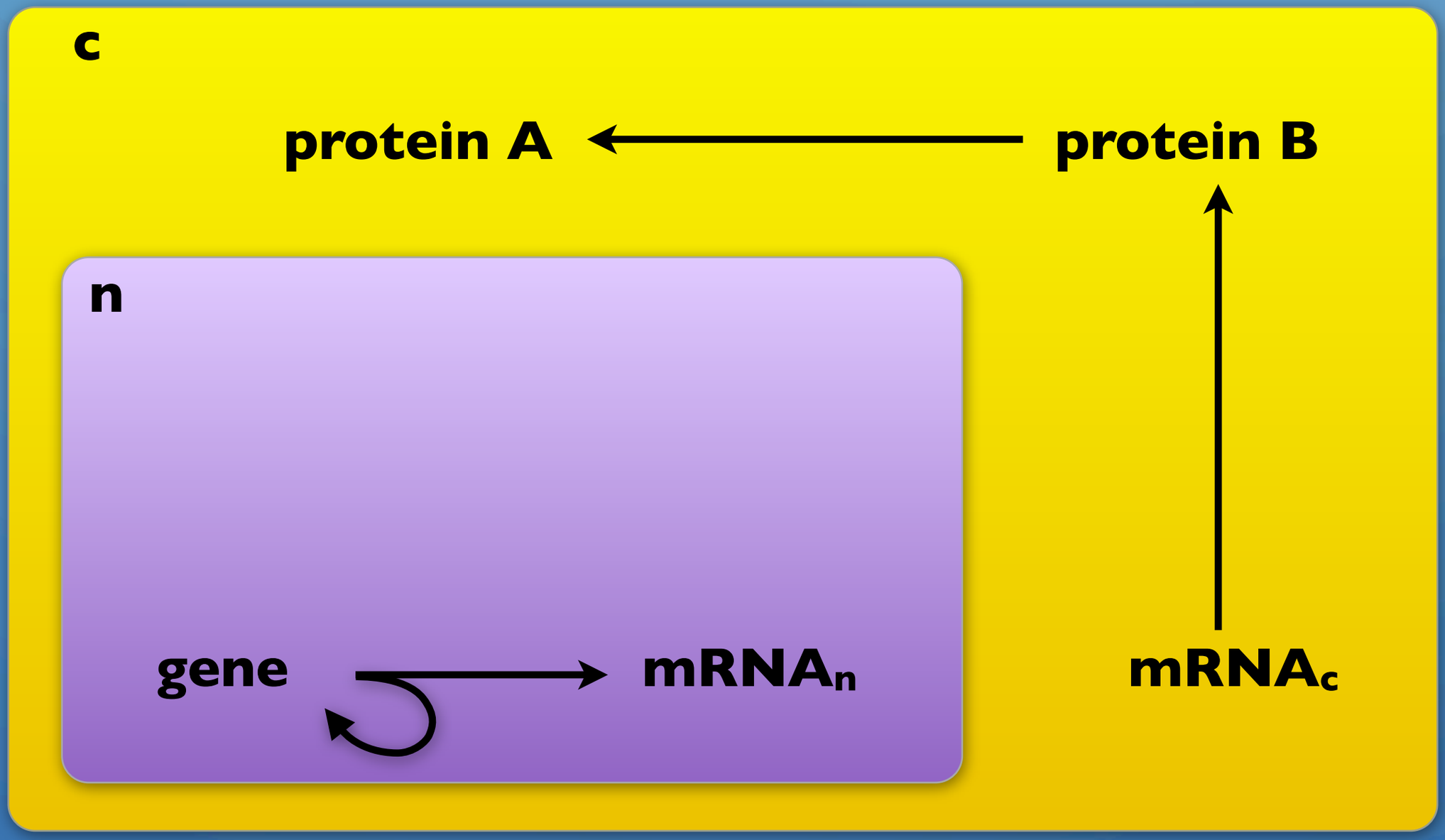
Some basics of SBML model encoding

- Species pools are located in compartments



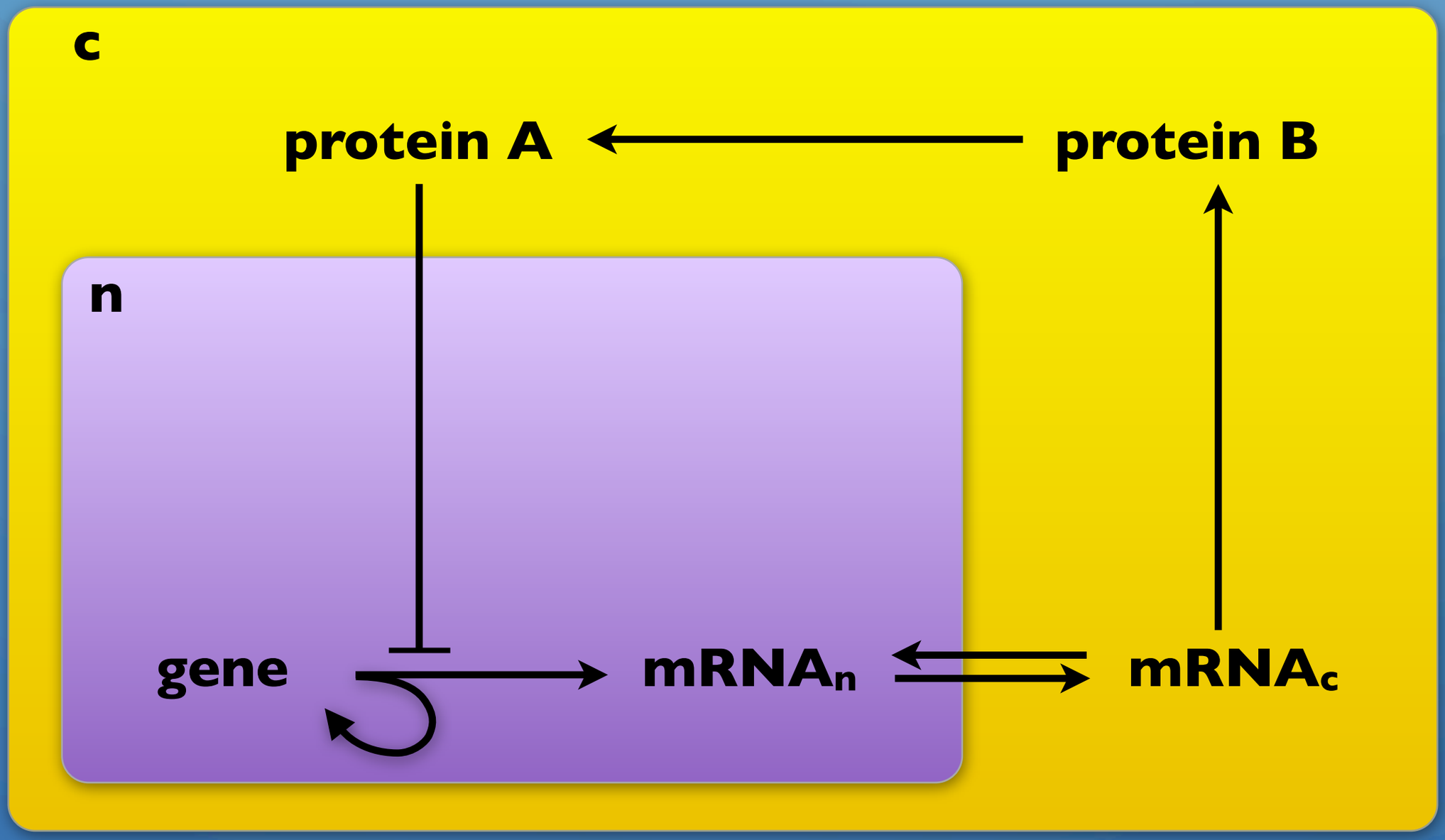
Some basics of SBML model encoding

- Reactions can involve any species anywhere



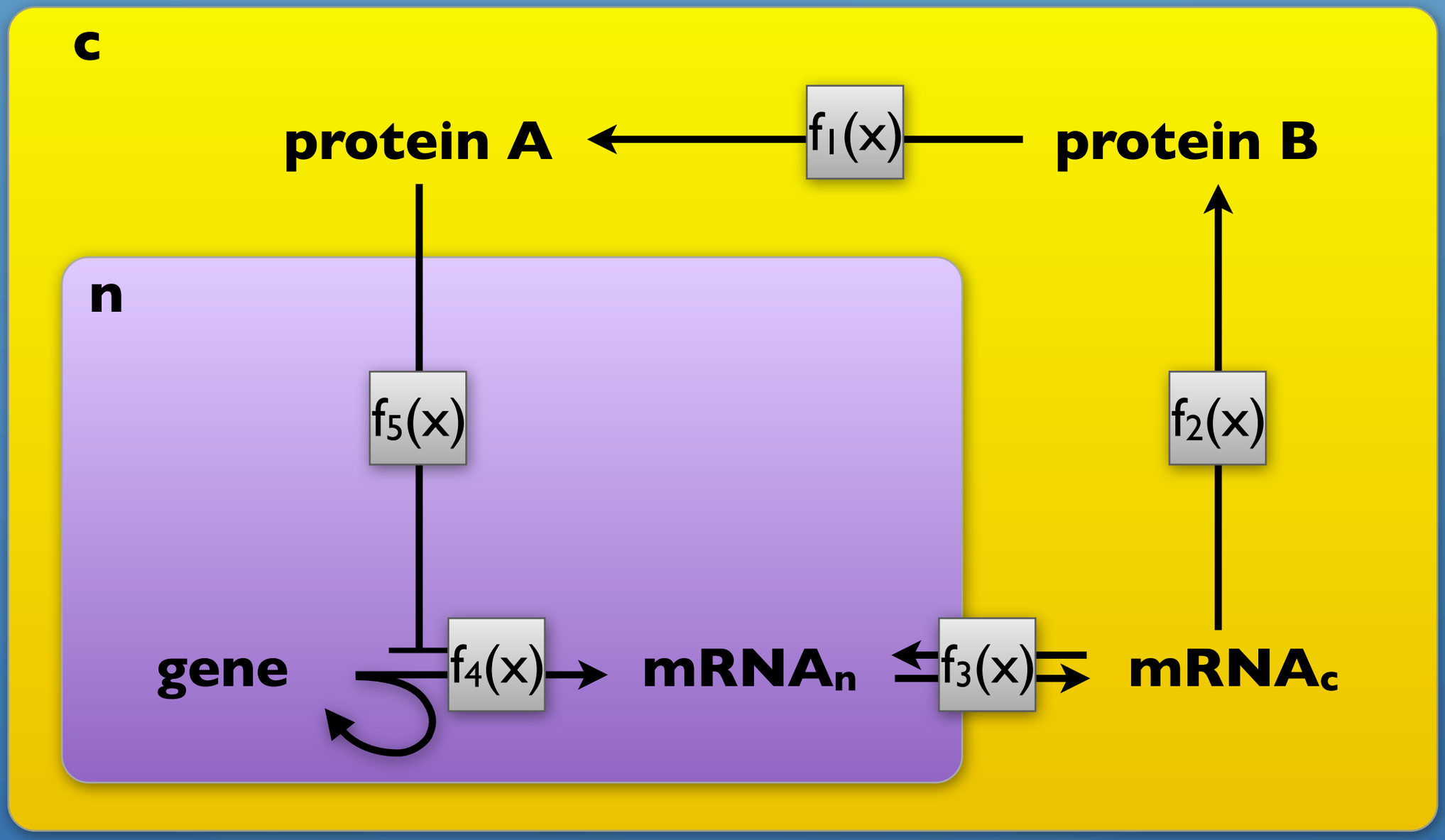
Some basics of SBML model encoding

- Reactions can cross compartment boundaries



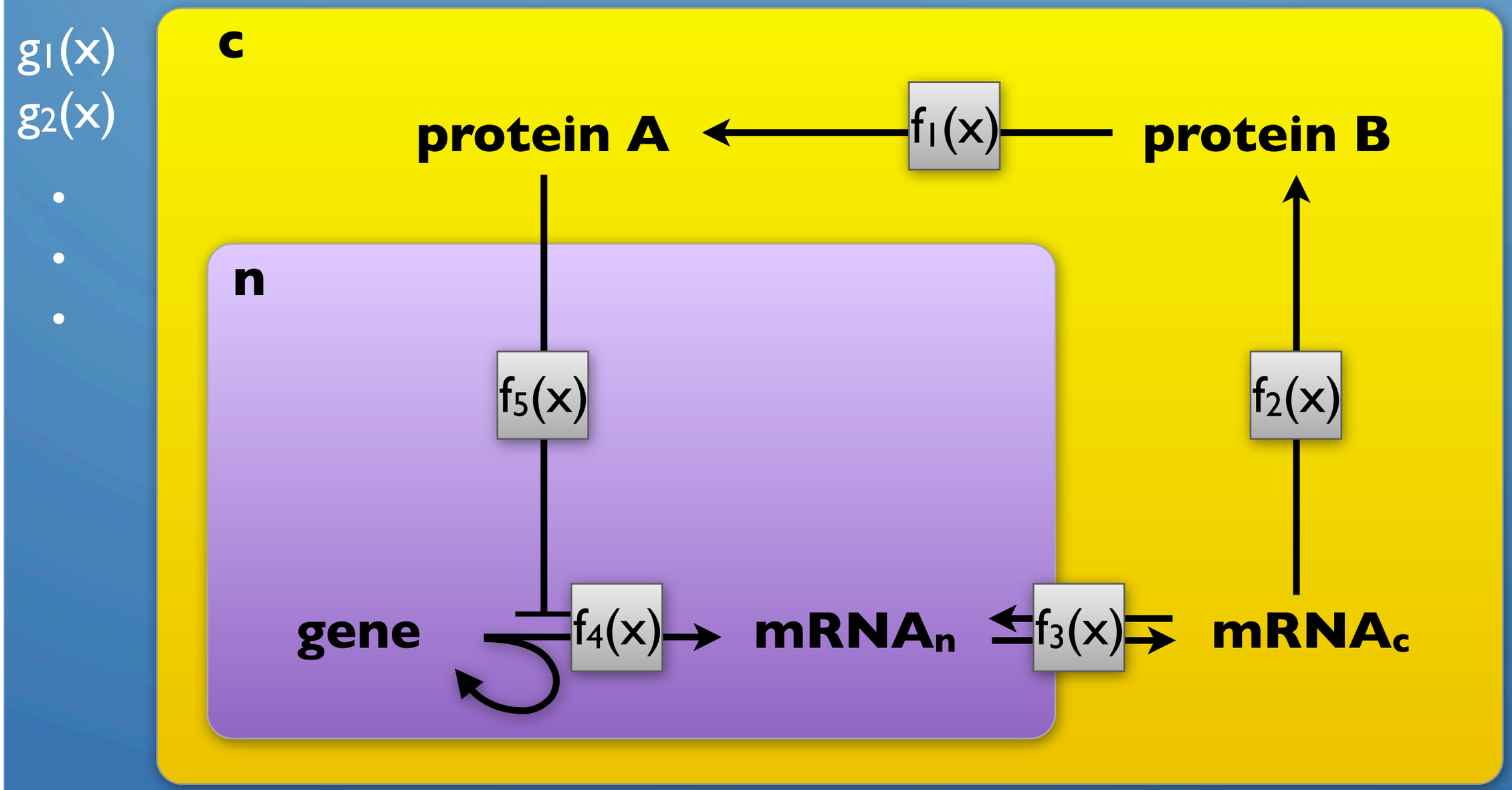
Some basics of SBML model encoding

- Reaction/process rates can be (almost) arbitrary formulas



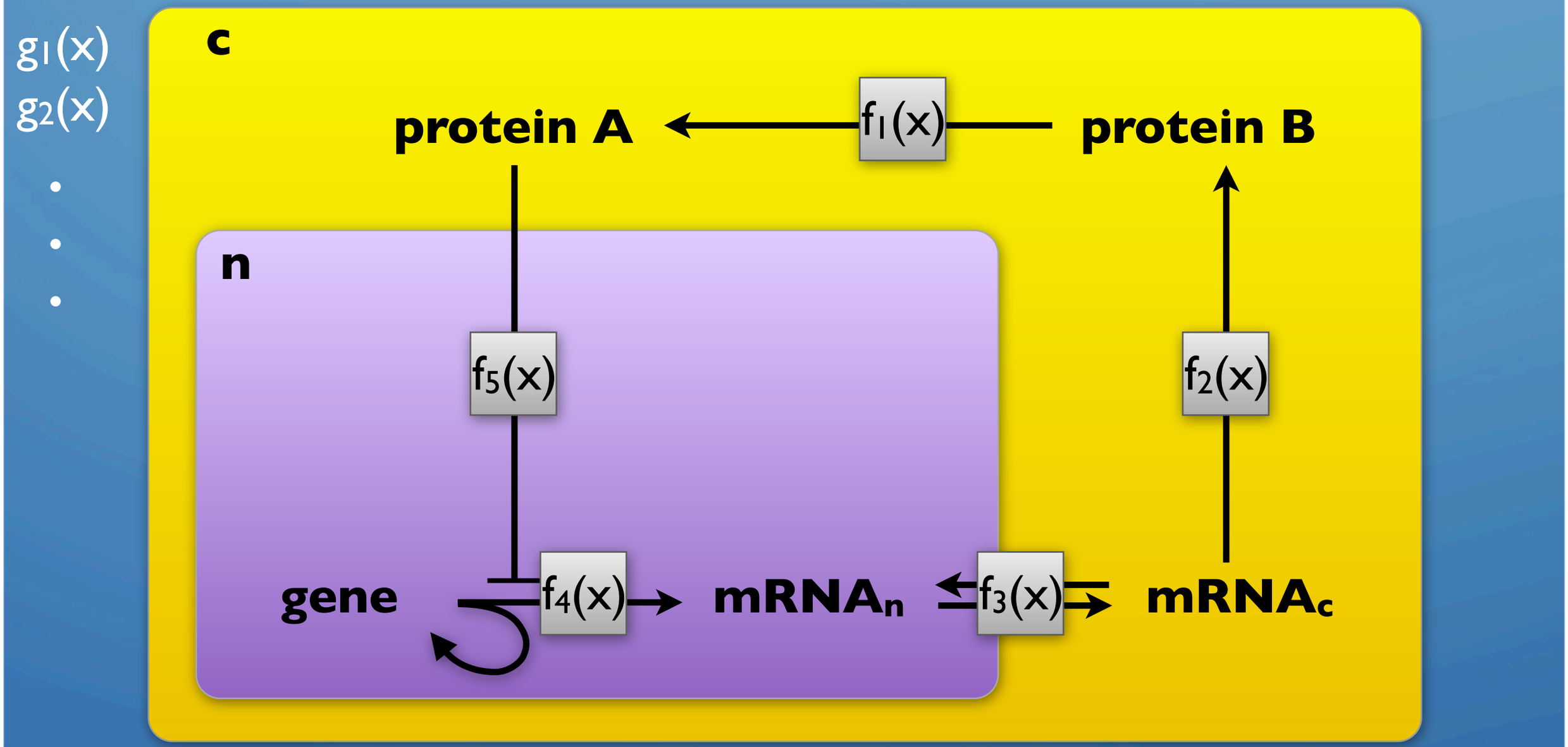
Some basics of SBML model encoding

- "Rules": equations expressing relationships in addition to reaction sys.



Some basics of SBML model encoding

- "Events": discontinuous actions triggered by system conditions

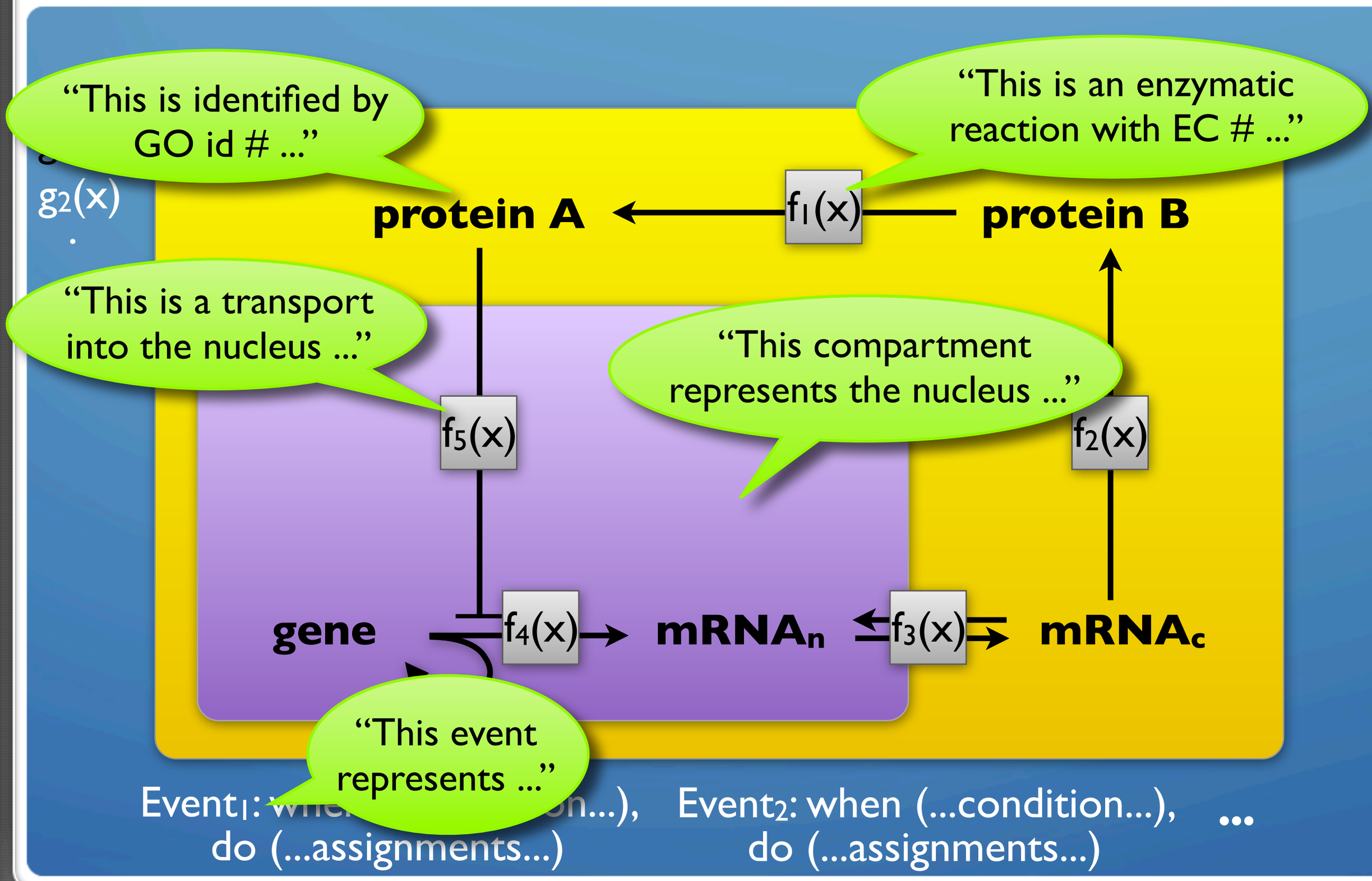


Event$_1$: when (...condition...), do (...assignments...)

Event$_2$: when (...condition...), do (...assignments...)

...

Some basics of SBML model encoding



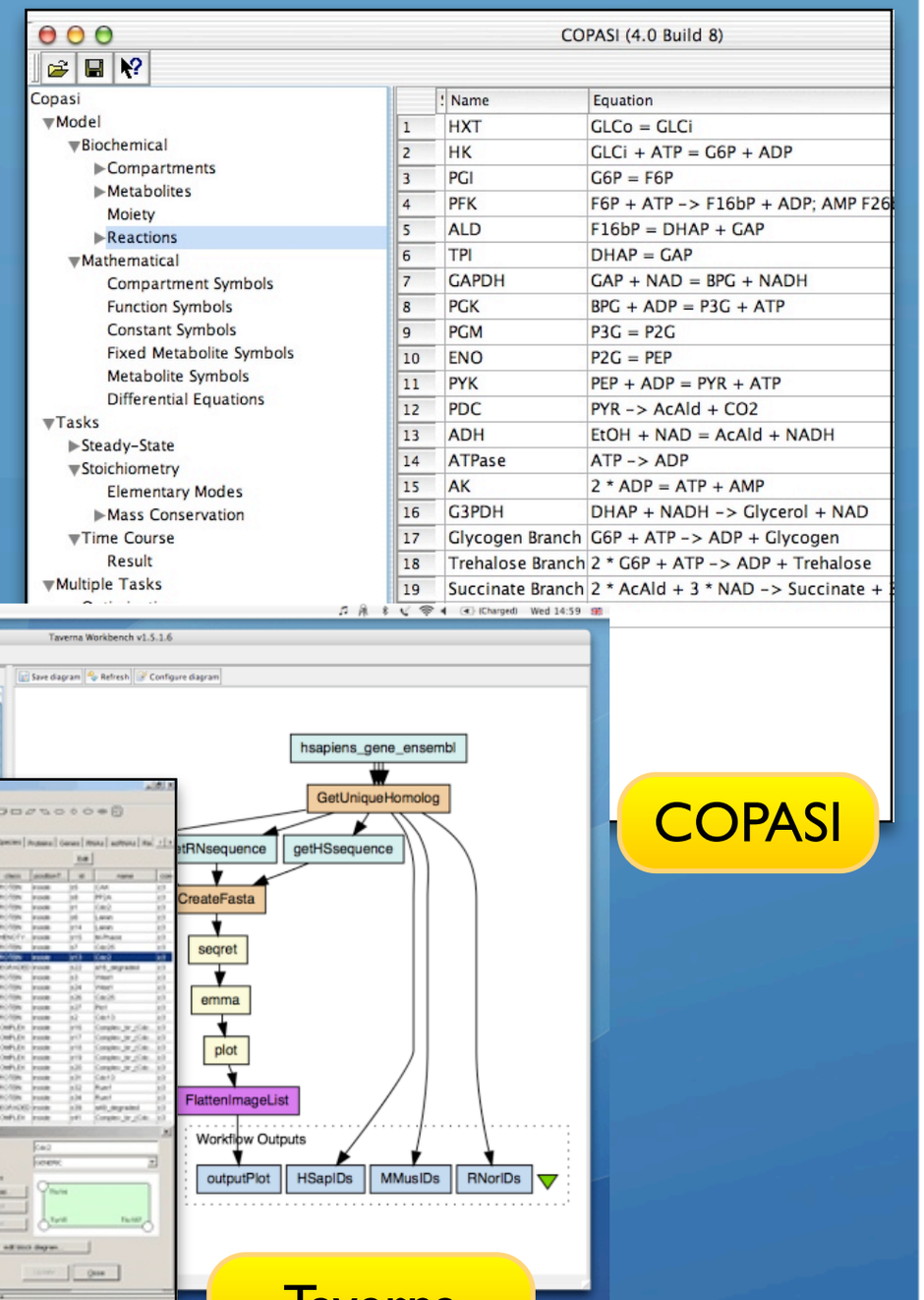
SBML is fundamentally an exchange format

- XML-based
- Lingua franca for software exchange of models
  - Not for humans to edit directly
  - Not necessarily a software system's *internal* format
  - Not suited for experimental or numerical results
  - Not a union of all possible features
- Open & free

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    <species compartment="cytosol" id="P"  initi
    <species compartment="cytosol" id="S"  initi
    <species compartment="cytosol" id="E"  initi
</listOfSpecies>
<listOfReactions>
    <reaction id="veq">
        <listOfReactants>
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        <listOfProducts>
            <speciesReference species="ES"/>
        </listOfProducts>
        <kineticLaw>
            <math xmlns="http://www.w3.org/1998/
                <apply>
                    <times/>
                    <ci>cytosol</ci>
                    <apply>
                        <minus/>
                        <apply>
                            <times/>
                            <ci>kon</ci>
                            <ci>E</ci>
                            <ci>S</ci>
                        </apply>
                        <apply>
                            <times/>
                            <ci>koff</ci>
                            <ci>ES</ci>
                        </apply>
```

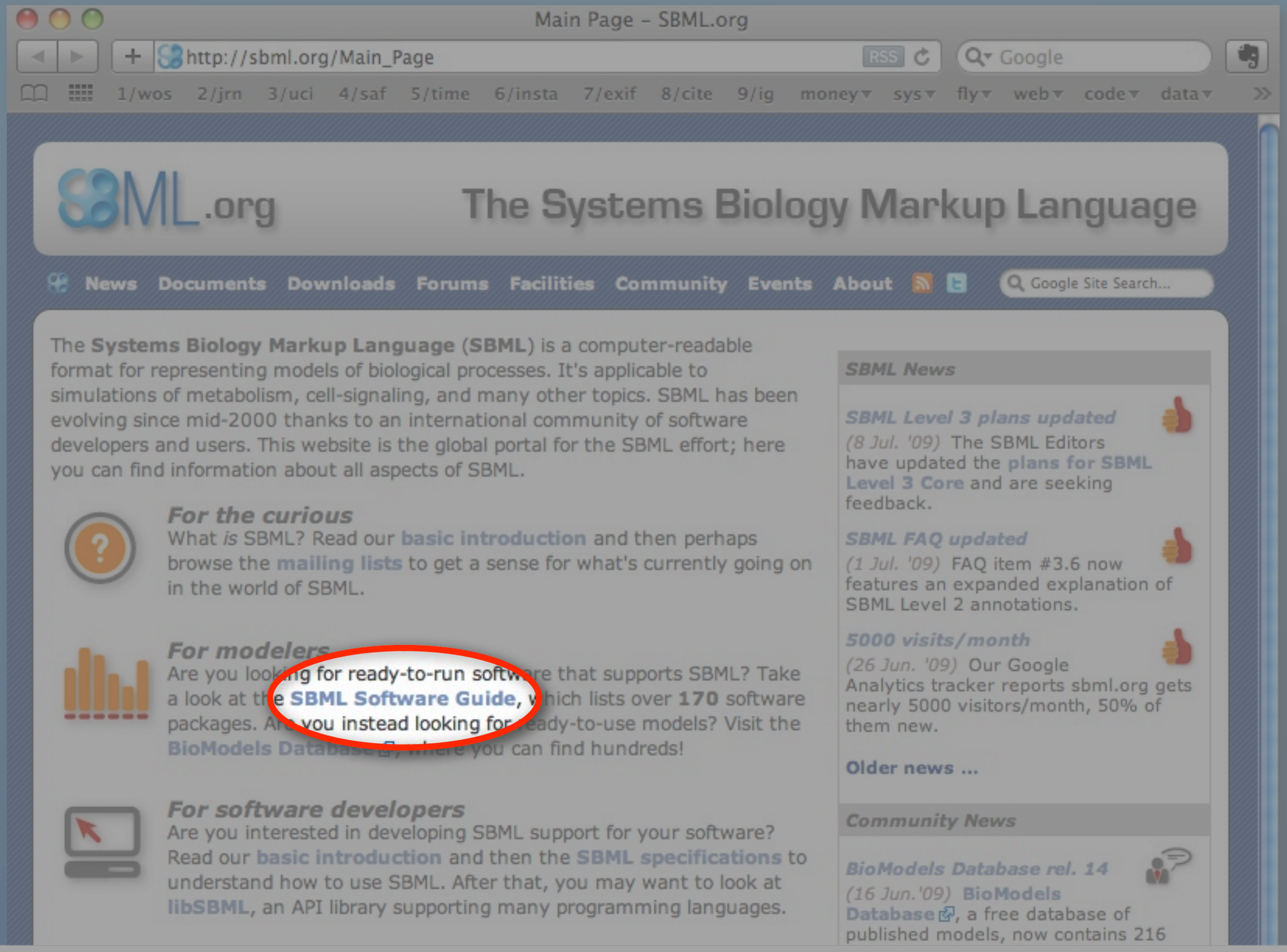
Broad support of SBML

- Supported by 200+ software systems
  - Open-source & commercial
    - General-purpose environments
      - Mathematica, MATLAB, etc.
    - Special-purpose software
      - model editing
      - simulation
      - analysis
      - visualization

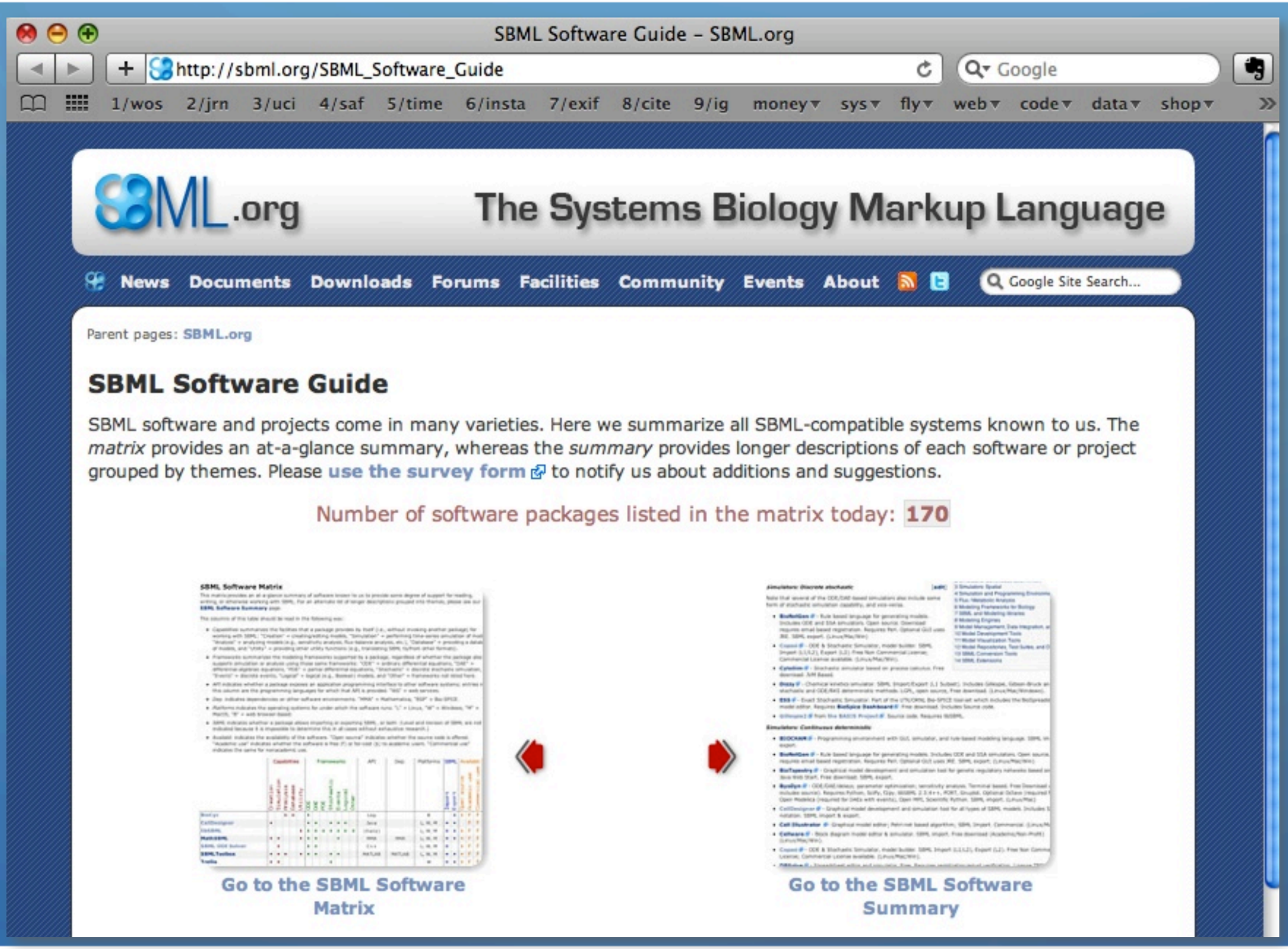


SBML Software Guide

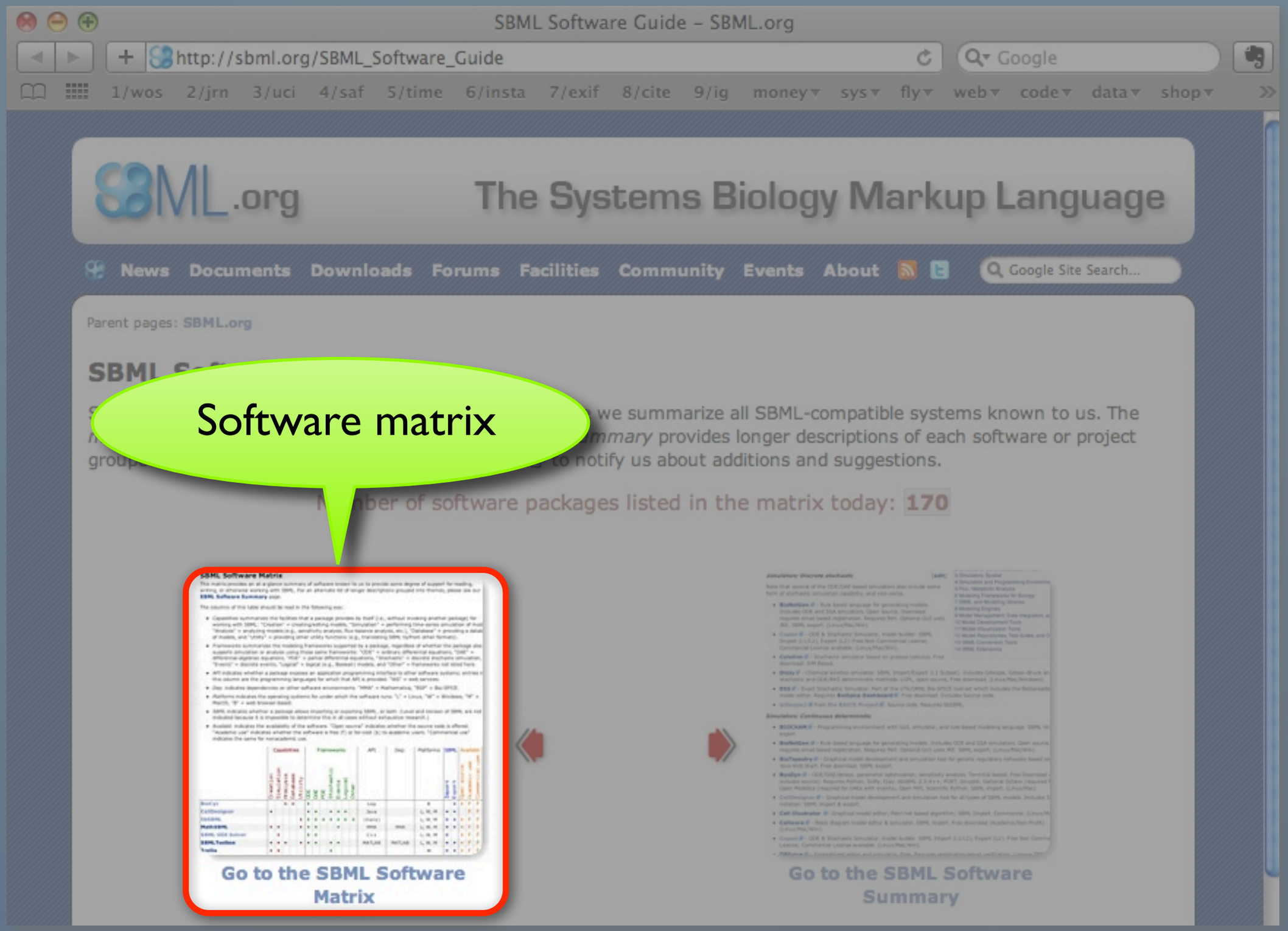
SBML Software Guide



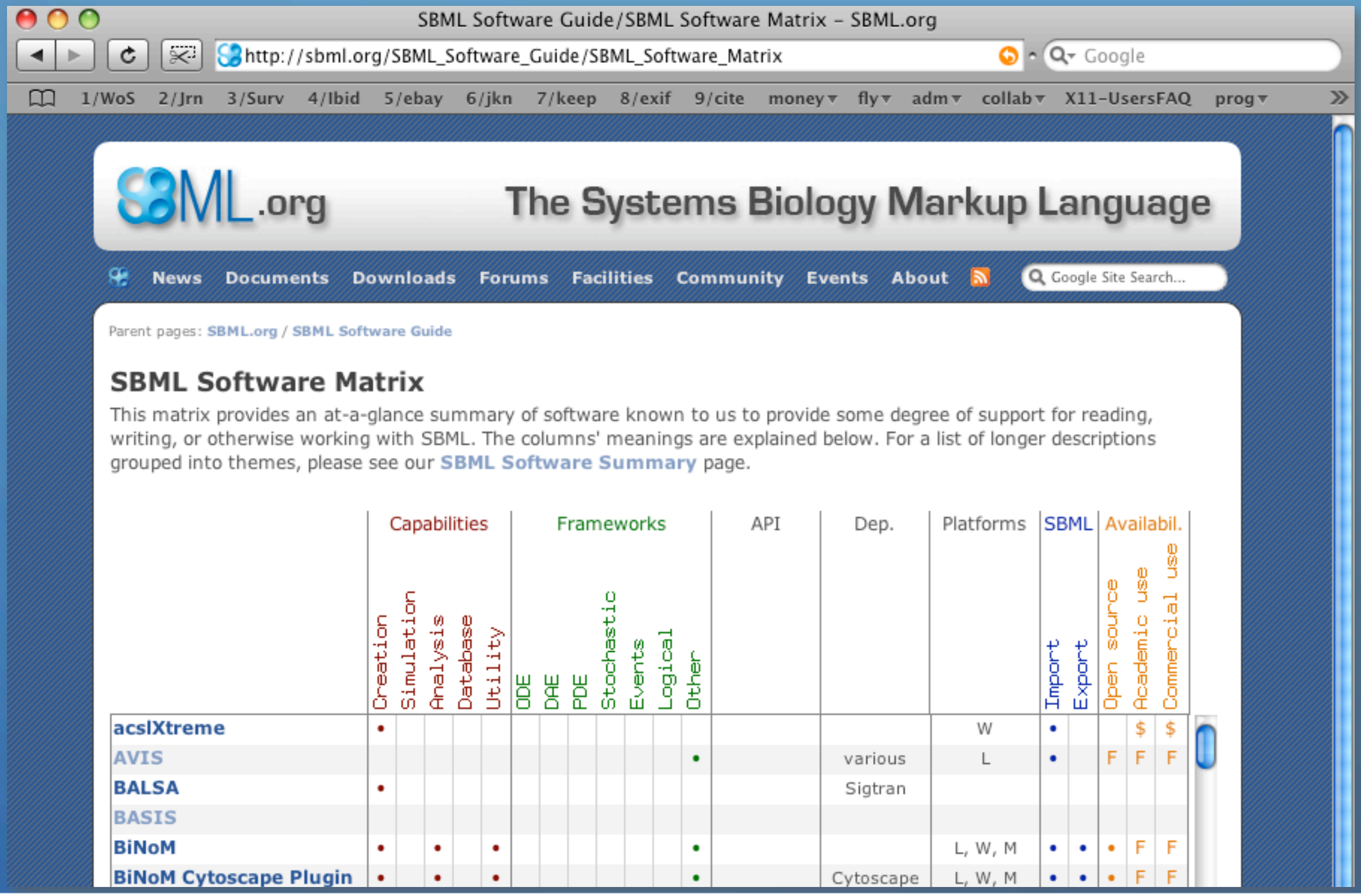
Guide has both matrix and summary views



Guide has both matrix and summary views



SBML Software Matrix

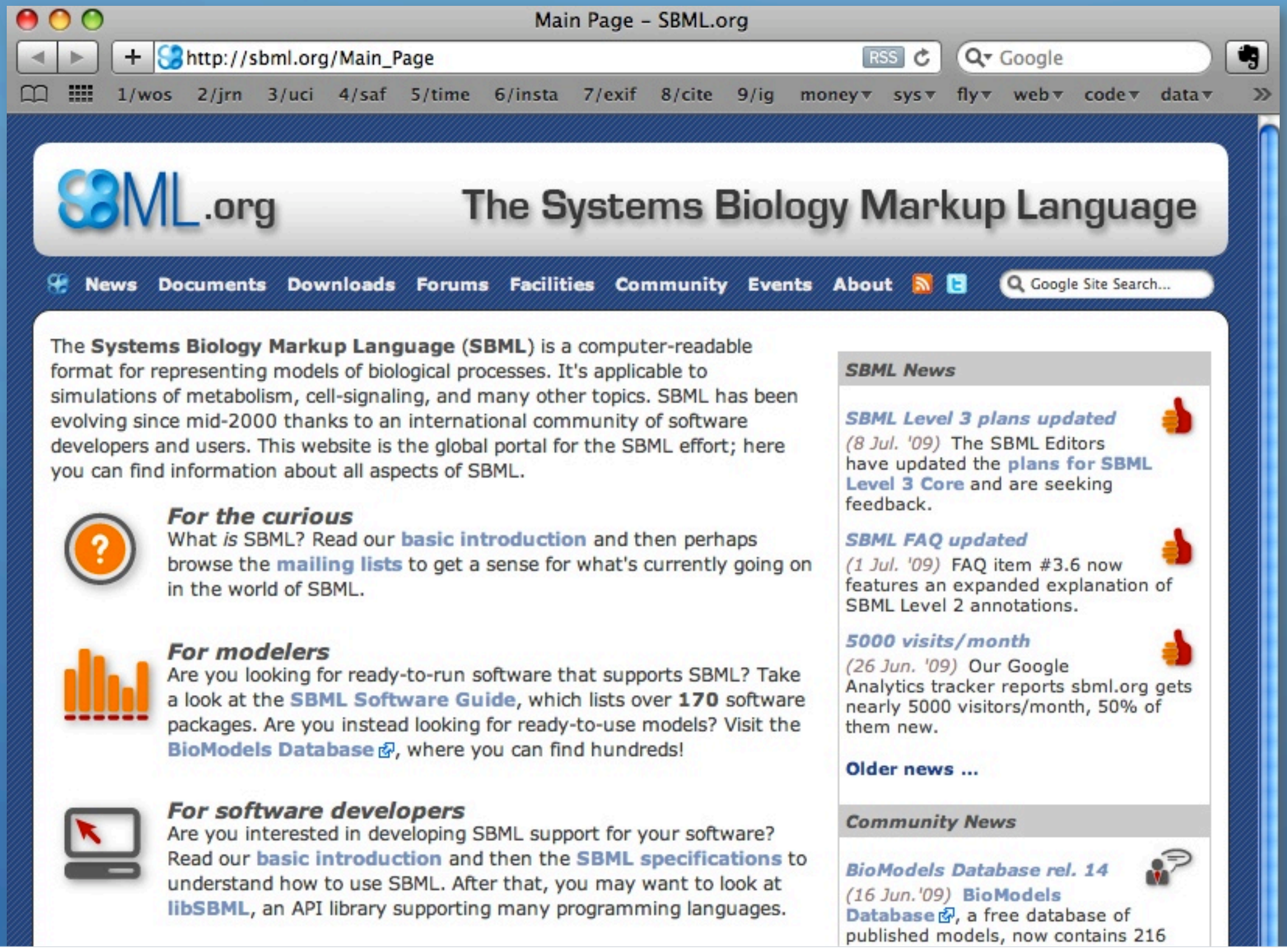


	Capabilities					Frameworks							API	Dep.	Platforms	SBML		Availabil.		
	Creation	Simulation	Analysis	Database	Utility	ODE	DAE	PDE	Stochastic	Events	Logical	Other				Import	Export	Open source	Academic use	Commercial use
acslXtreme	•														W	•			$	$
AVIS												•		various	L	•		F	F	F
BALSA	•													Sigtran						
BASIS																				
BiNoM	•		•		•							•			L, W, M	•	•	•	F	F
BiNoM Cytoscape Plugin	•		•		•							•		Cytoscape	L, W, M	•	•	•	F	F

Some particularly full-featured tools for simulation

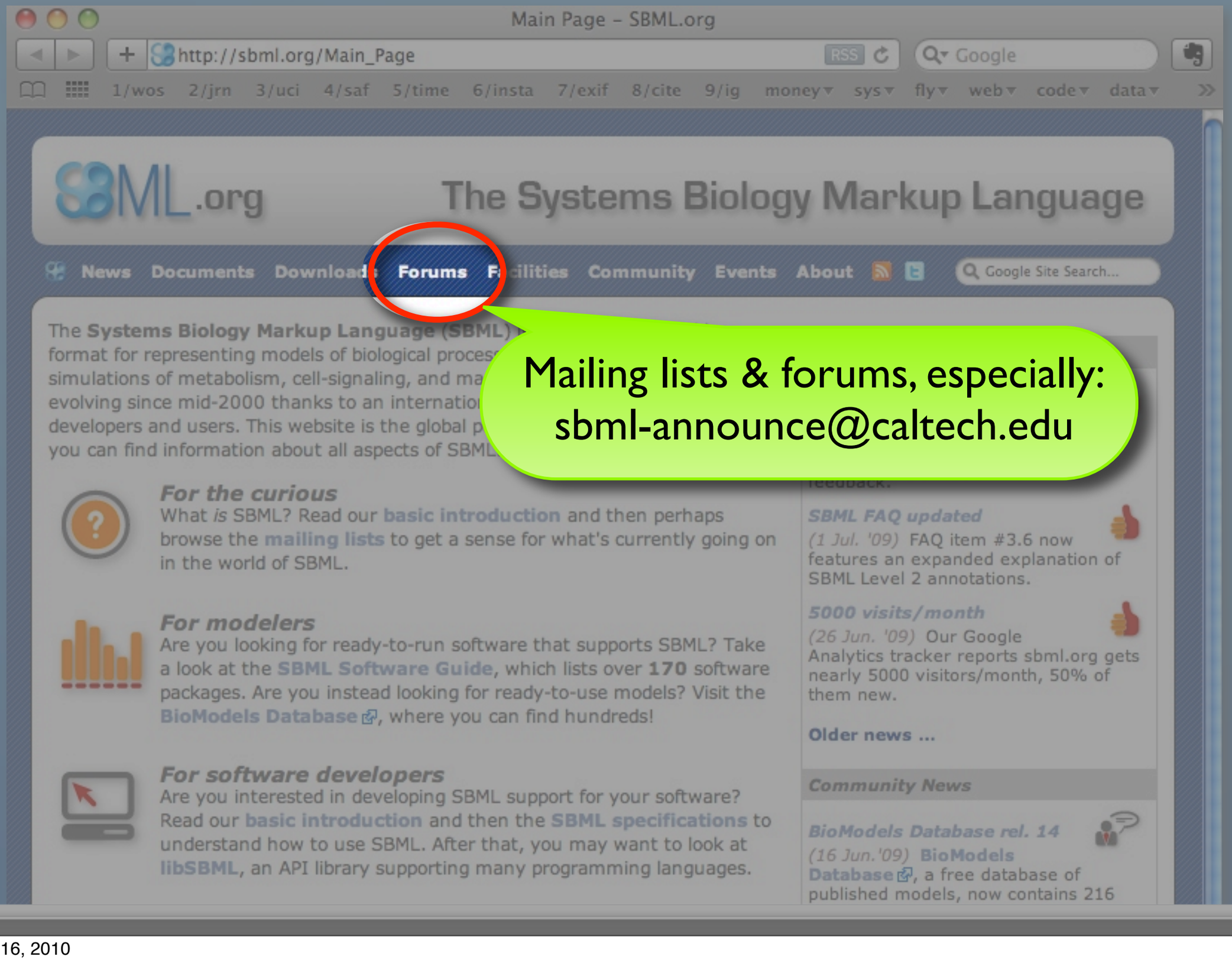
- **COPASI**: ODE & stochastic simulation, parameter scanning, powerful plotting capabilities, much more
- **iBioSim**: special features for genetic circuit models for synthetic bio
- **CellDesigner**: graphical editing interface, SBGN support, SABIO-RK integration, much more
  - Can use COPASI as computational engine
- **Virtual Cell**: web-based interface
- **SBtoolbox2**: MATLAB package
- **SBW**: collection of interoperating programs (many Windows-only)

http://sbml.org

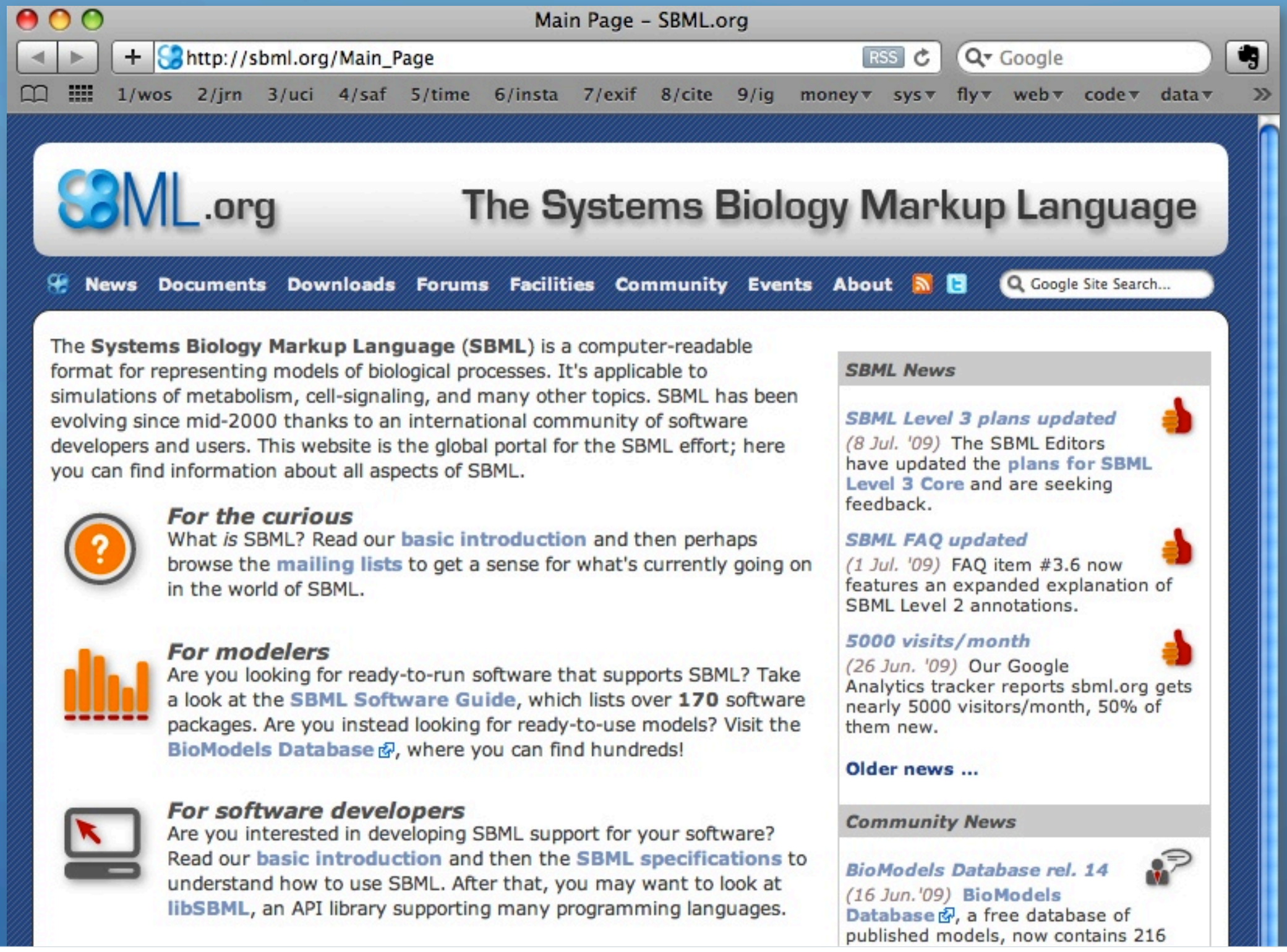


http://sbml.org

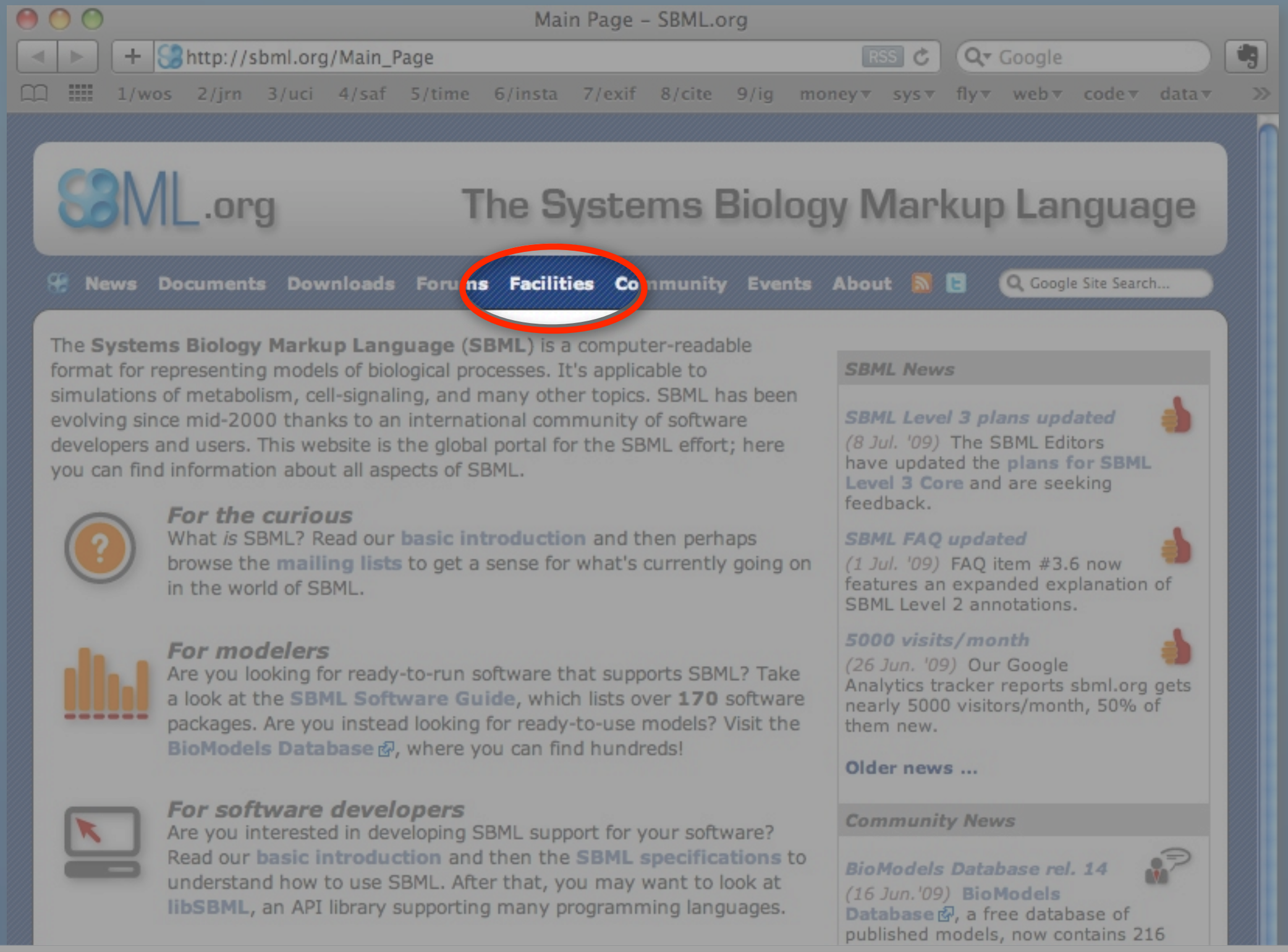
http://sbml.org



Online SBML Validator

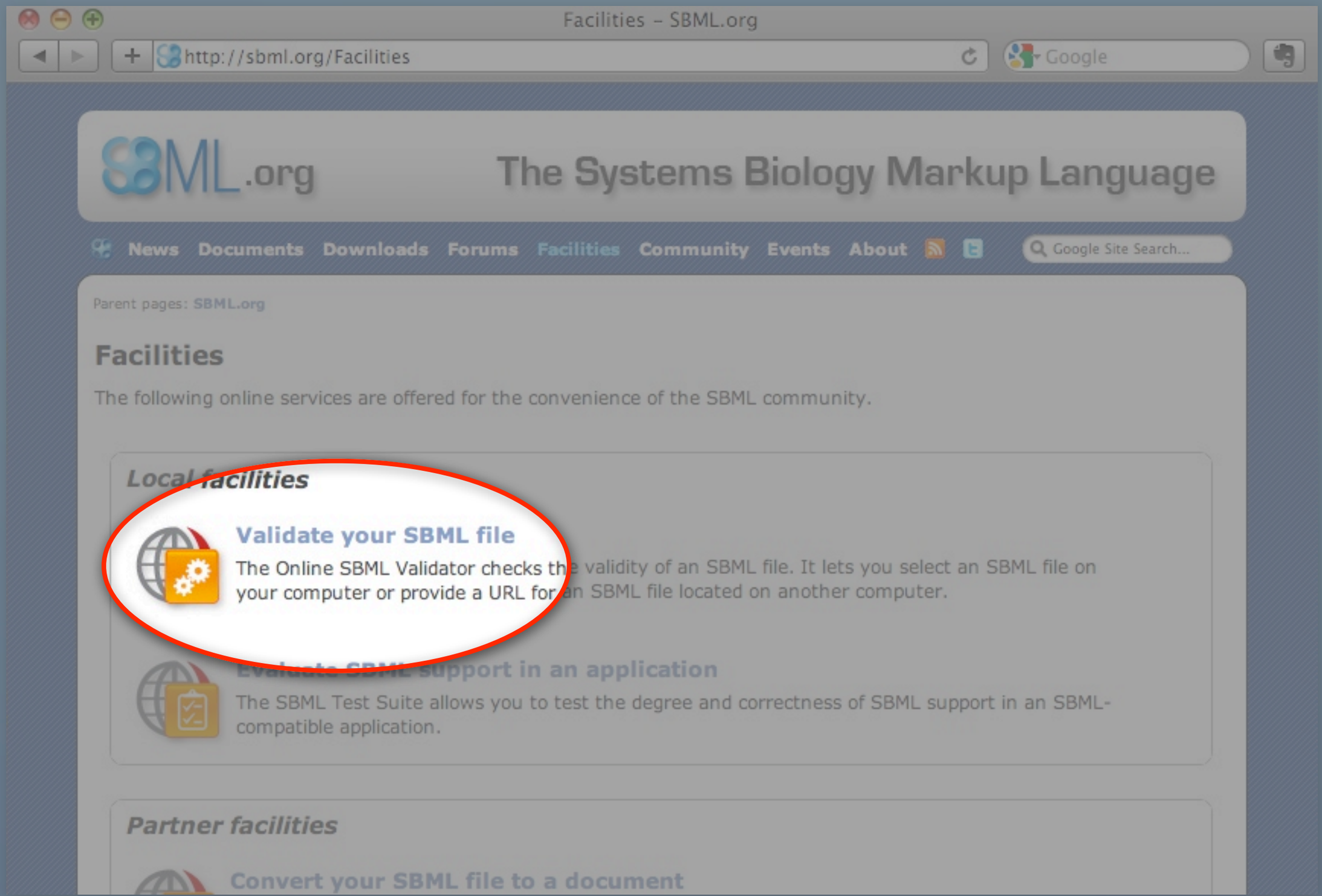


Online SBML Validator



Online SBML Validator

Online SBML Validator



libSBML

http://sbml.org/Software/libSBML

LibSBML, for programming SBML support

- Reads, writes, validates SBML
  - Hundreds of rules for helping to ensure correct SBML
- Unit checking & conversion
- Well-tested
- Written in portable C++
  - Linux, Mac, Windows, FreeBSD
- APIs for C, C++, C#, Java, Octave, Perl, Python, Ruby, MATLAB (some via SWIG)
- Can use Expat, libxml2, or Xerces
- Open-source under LGPL
- 110k+ lines of code + 65k lines of XML tests +18k lines of language interface code

Latest stable version: 4.2.0

http://sbml.org/Software/libSBML

- Developed by Sarah Keating, Ben Bornstein, Akiya Jouraku, & Mike Hucka, with substantial contributions from many other people

LibSBML 5 for SBML Level 3 packages

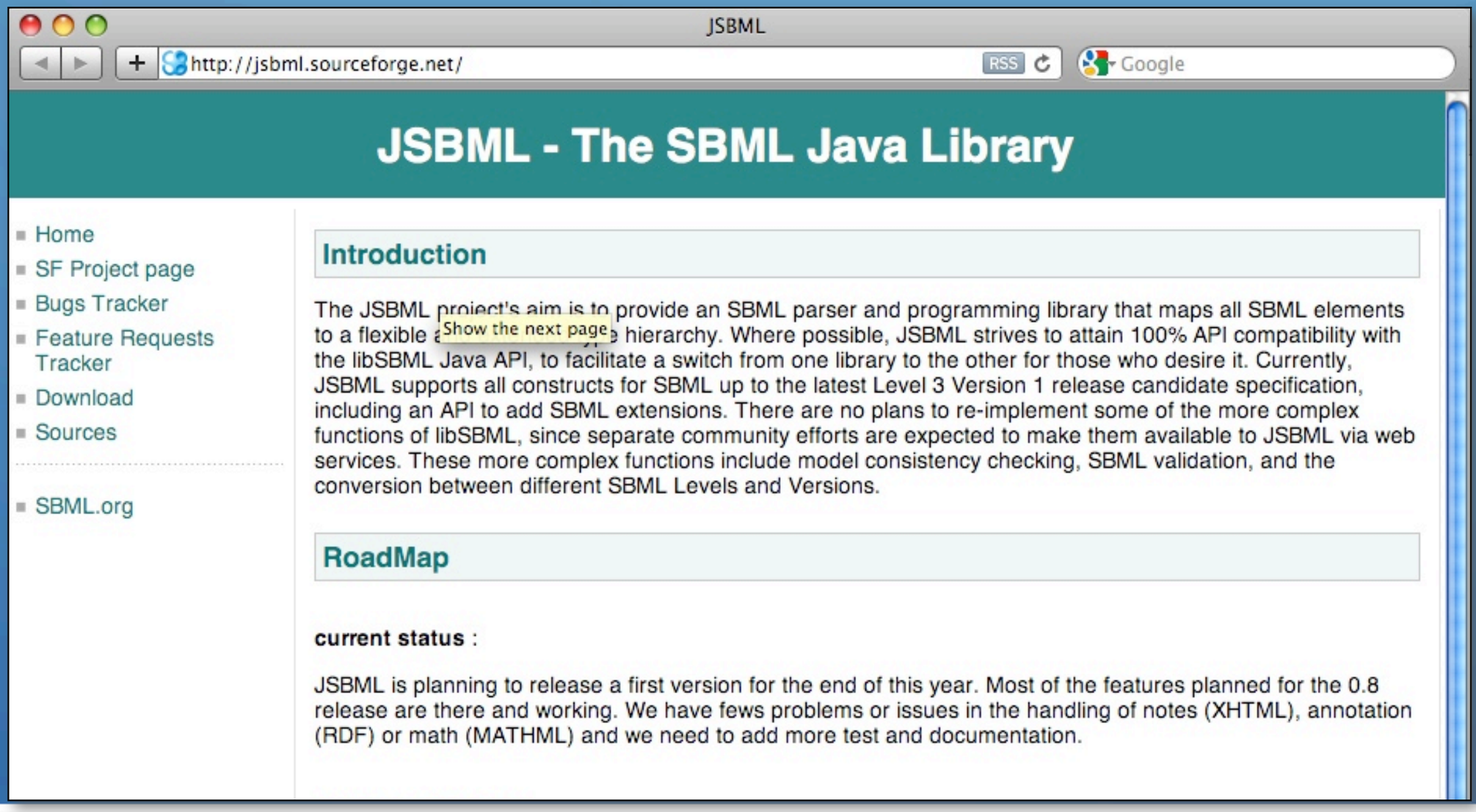
- Relationships between releases:

		libSBML		
		4.0.x	4.2.x	5.x
SBML	Level 2	✔	✔	✔
	Level 3 Core		✔	✔
	Level 3 packages			✔

- libSBML 5 has same API as libSBML 4.2 and adds plug-in mechanism
  - Each plug-in implements one SBML Level 3 package
  - Plug-ins register themselves with libSBML & provide XML namespace
- libSBML 5 is available now as an alpha release

On the horizon: JSBML

- Pure Java work-alike implementation of portions of libSBML
- Still in development: **http://jsbml.sourceforge.net**



BioModels Database

http://biomodels.net/biomodels

http://biomodels.net/biomodels

BioModels Home | Models | Submit | Support | About BioModels | Contact us

BioModels Database - A Database of Annotated Published Models

BioModels Database is a data resource that allows biologists to store, search and retrieve published mathematical models of biological interests. Models present in BioModels Database are annotated and linked to relevant data resources, such as publications, databases of compounds and controlled vocabularies.
BioModels Database also allows users to generate sub-models, provides access to online simulation tools and features programmatic access via Web Services.

Search | Go to the model | Advanced search

Browse models

- Curated models (269)
- Browse models using GO
- Non-curated models (361)

Simulate in JWS Online

Submit a model

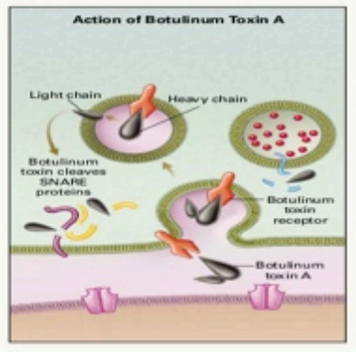
Main instance at European Bioinformatics Institute http://www.ebi.ac.uk/biomodels/

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

Display a menu

Model of the month

August, 2010



Botulinum neurotoxin serotype A (BoNT/A) causes flaccid paralysis by a multi-step mechanisim. Two mathematical models that has been developed, to estimate upper limits of the time during which antitoxin and other impermeable inhibitors of BoNT/A can exert an effect, is described here.
Read more...

News

30 September 2010 **Eighteenth Release!**
Download All Models Under SBML Format

BioModels Database

- Stores & serves quantitative models of biological interest
  - Free, public resource
  - Models must be described in peer-reviewed publication(s)
- All models are **curated by hand** to reproduce published results
- Imports & exports models in several formats
  - SBML, CellML, SciLab, XPP, BioPAX
- Developed by Nicolas Le Novère's group (EBI), funded by EBI & NIH
- Today: 600+ models

Curated, annotated models drawn from the literature

BioModels Database

http://biomodels.caltech.edu/BIOMD0000000234

BioModels Home | Browse models | Support | About BioModels | Search

BIOMD0000000234 - Tham2008_PDmodel_TumourShrinkage

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

Model | Overview | Math | Physical entities | Parameters | Curation

Reference Publication

Publication ID: 18594002

Clin Cancer Res 2008 Jul;14(13):4213-8.
A pharmacodynamic model for the time course of tum
patients.
Tham LS, Wang L, Soo RA, Lee SC, Lee HS, Yong W
Department of Hematology-Oncology, National Unive

Model

Original Model: *BIOMD0000000234.xml.origin*

Submitter: Nick Holford

Submission ID: MODEL0911120001

Submission Date: 16 Nov 2009 02:46:45 UTC

Last Modification Date: 23 Jan 2010 00:08:36 UTC

Creation Date: 16 Nov 2009 20:37:34 UTC

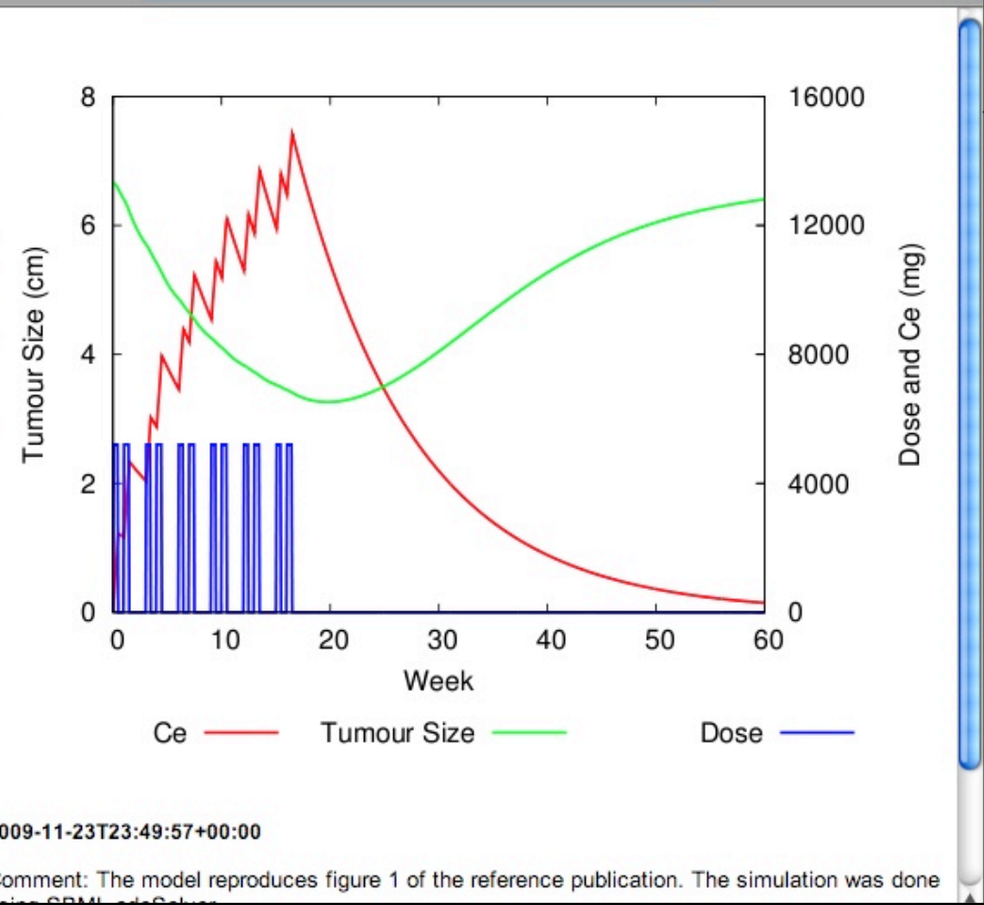
Encoders: Vijayalakshmi Chelliah
Lai-San Tham

set #1	bqbiol:occursIn	Taxonomy Homo sapiens
	bqbiol:isVersionOf	ICD C34 OMIM 211980

Notes

Curation Result

http://biomodels.caltech.edu/simulation-result.



2009-11-23T23:49:57+00:00

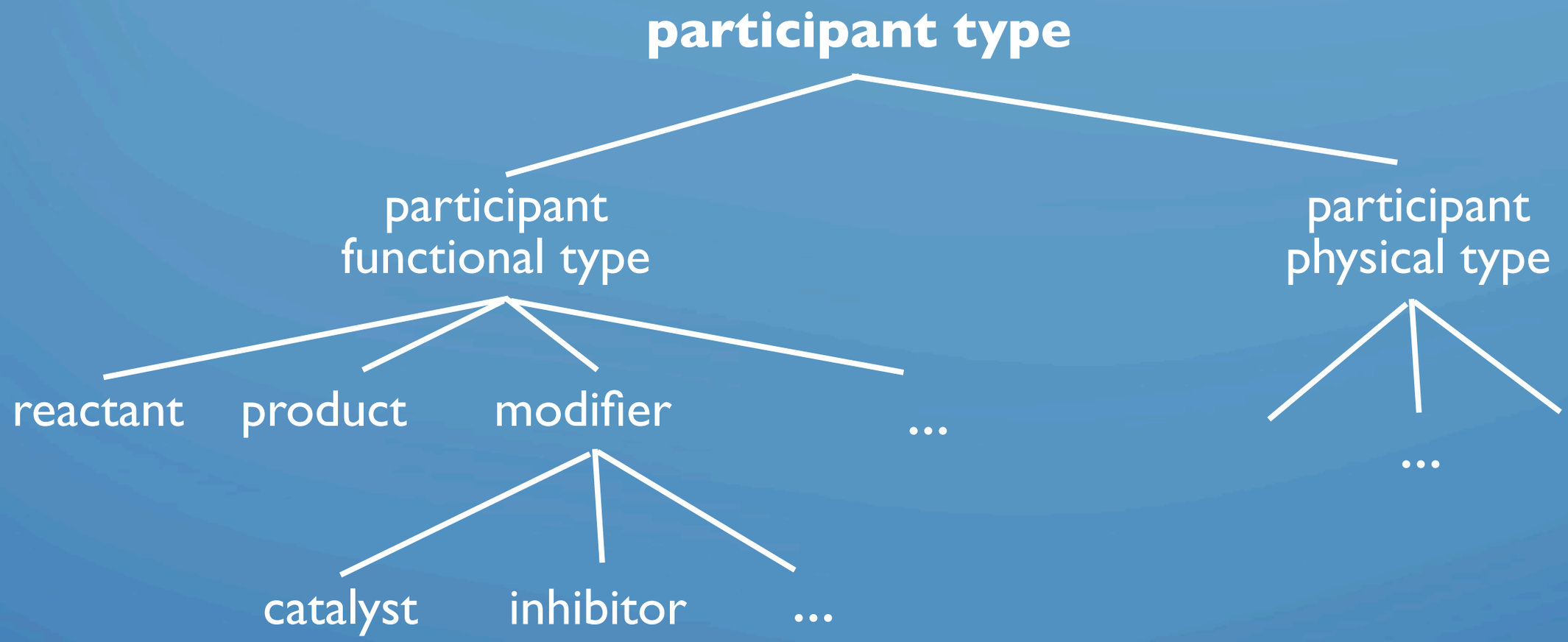
Comment: The model reproduces figure 1 of the reference publication. The simulation was done

Ease of model exchange is driving new developments

- More people are exchanging models and using more software tools
  - Encountering challenges not faced before
    - BioModels Database in particular has uncovered new needs
- Result: development of additional standardization efforts
  - SBO
  - MIRIAM
  - MIASE
  - Also (but not presented here):
    - SBRML
    - SBGN

SBO

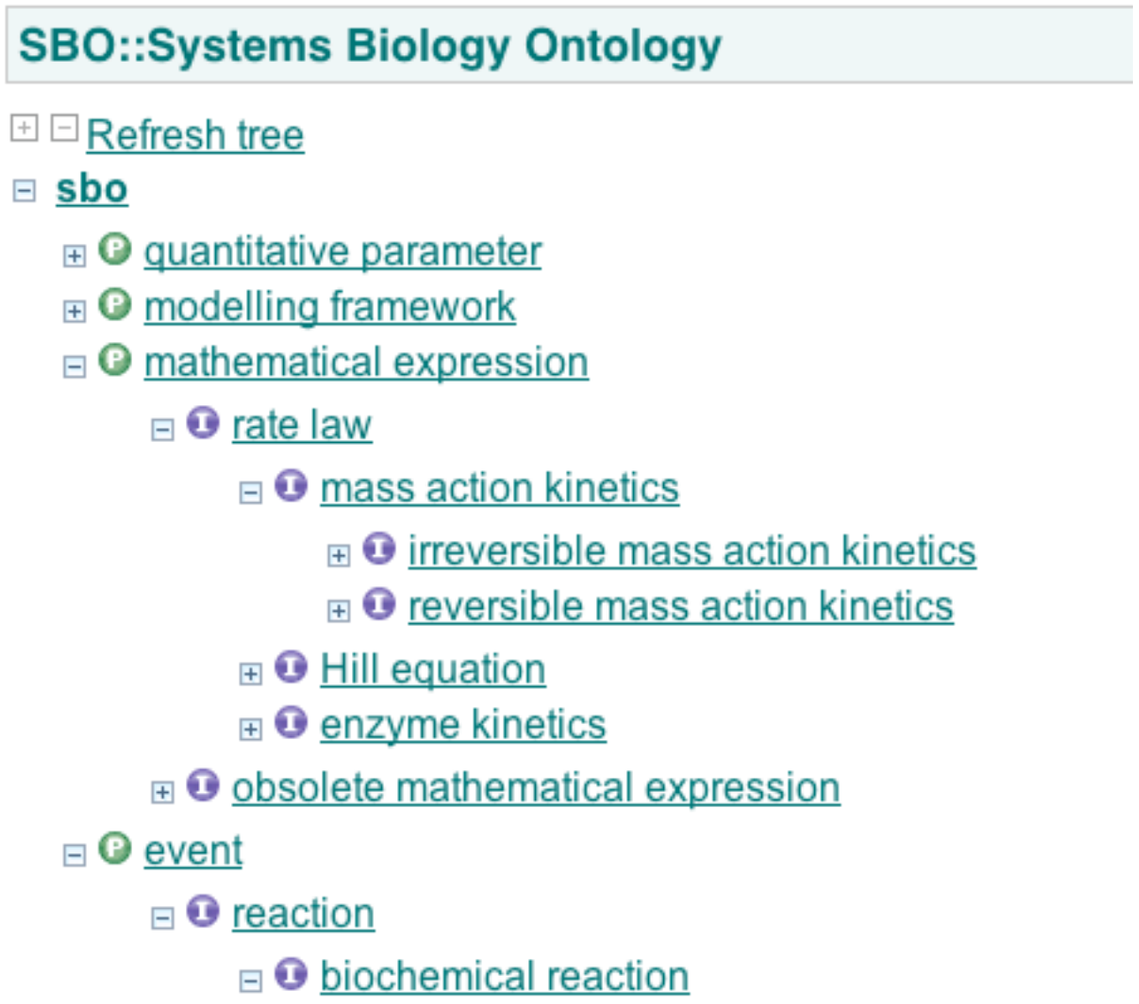
Set of controlled vocabularies



- Not for describing the biology, but rather the **math** in a model
  - E.g.: the different pieces of a rate law, the constants, the mathematical framework assumed (e.g., continuous versus discrete)

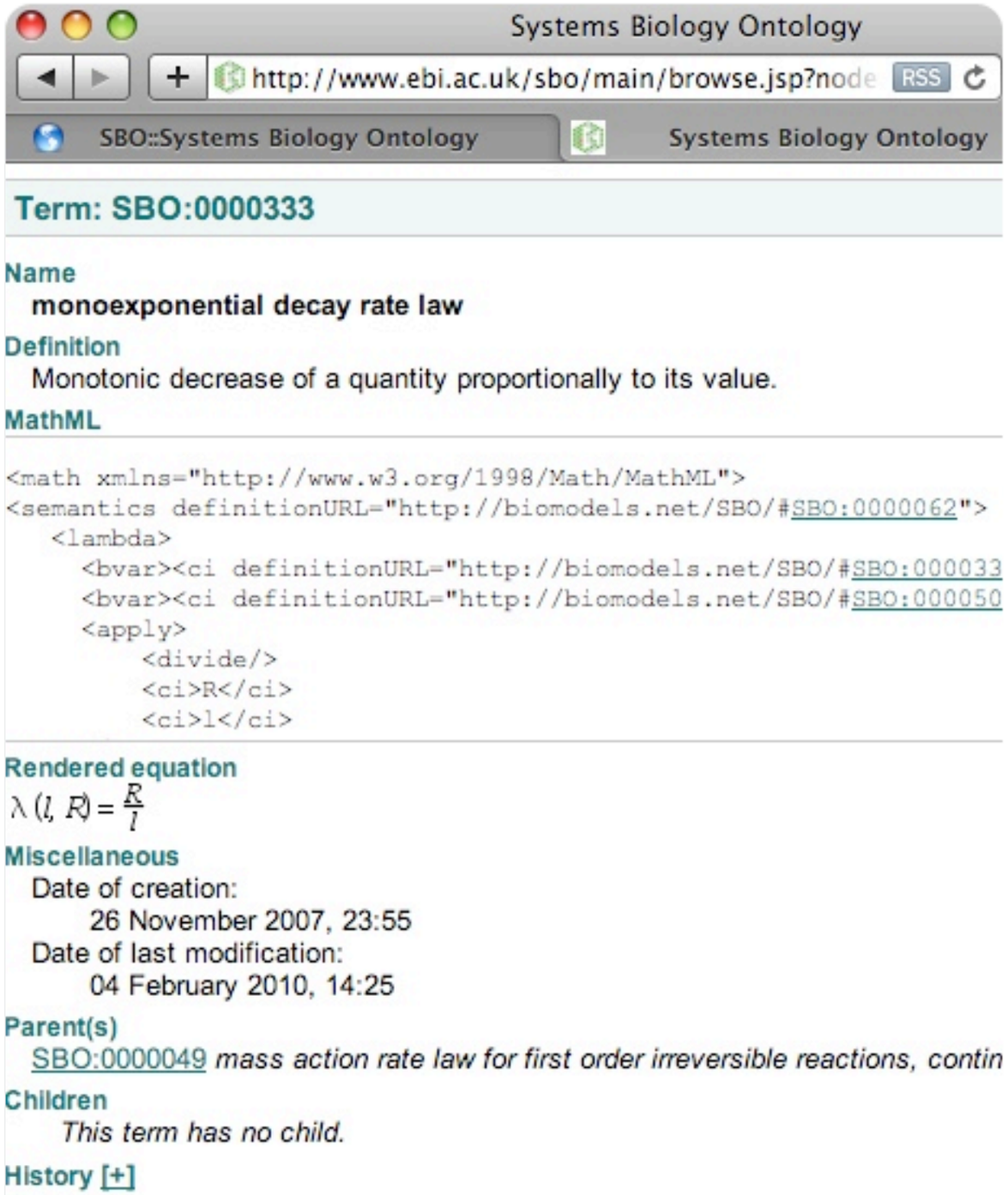
SBO browser (for humans) & web services (for s/w)

- **http://www.ebi.ac.uk/sbo**
- Hierarchical organization
- Machine-readable
- **Community input wanted!**
  - Please comment on ontology structure, suggest additions, etc.
- Software developed by Melanie Courtot and Camile Laibe in the Le Novère group @ EBML-EBI



Entries contain retrievable data

- Entries for mathematical expressions contain MathML
- All entries contain cross links



Direct SBO support in SBML L2V3

- **`sboTerm`** field on all SBML constructs (part of **SBase**)
  - Usage: **`sboTerm="SBO:0123456"`**
- Terms must be drawn from appropriate SBO subtrees

Component	SBO vocabulary	Parent SBO id
Parameter	Quantitative parameter	SBO:0000002
KineticLaw	Mathematical expression	SBO:0000064
SpeciesReference	Functional Participant	SBO:0000003
…	…	…

Example

```
<listOfCompartments>
  <compartment id="cell" size="1e-15" />
</listOfCompartments>
<listOfSpecies>
  <species compartment="cell" id="S1" initialAmount="1000" />
  <species compartment="cell" id="S2" initialAmount="0" />
<listOfSpecies>
<listOfParameters>
  <parameter id="k" value="0.005" sboTerm="SBO:0000339" />
<listOfParameters>
<listOfReactions>
  <reaction id="r1" reversible="false">
    <listOfReactants>
      <speciesReference species="S1" stoichiometry="2" sboTerm="SBO:0000010" />
    </listOfReactants>
    <listOfProducts>
      <speciesReference species="S1" stoichiometry="2" sboTerm="SBO:0000011" />
    </listOfProducts>
    <kineticLaw sboTerm="SBO:0000052">
      <math>
       ...
      <math>
```

Example

```
<listOfCompartments>
  <compartment id="cell" size="1e-15" />
</listOfCompartments>
<listOfSpecies>
  <species compartment="cell" id="S1" initialAmount="1000" />
  <species compartment="cell" id="S2" initialAmount="0" />
<listOfSpecies>
<listOfParameters>
  <parameter id="k" value="0.005" sboTerm="SBO:0000339" />
<listOfParameters>
```

SBO::Systems Biology Ontology

- SBO:0000000 - sbo
  - SBO:0000002 - quantitative parameter
    - SBO:0000226 - number density
    - SBO:0000227 - mass density
    - SBO:0000255 - physical characteristic
    - SBO:0000256 - biochemical parameter
      - SBO:0000009 - kinetic constant
        - SBO:0000016 - unimolecular rate constant
        - SBO:0000017 - bimolecular rate constant
          - SBO:0000023 - forward bimolecular rate constant
            - SBO:0000036 - forward bimolecular rate constant, continuous case
              - SBO:0000302 - catalytic efficiency
              - SBO:0000339 - bimolecular association rate constant
            - SBO:0000067 - forward bimolecular rate constant, discrete case

```
2" sboTerm="SBO:0000010" />

2" sboTerm="SBO:0000011" />
```

Example

```
<listOfCompartments>
  <compartment id="cell" size="1e-15" />
</listOfCompartments>
<listOfSpecies>
  <species compartment="cell" id="S1" initialAmount="1000" />
  <species compartment="cell" id="S2" initialAmount="0" />
<listOfSpecies>
<listOfParameters>
  <parameter id="k" value="0.005" sboTerm="SBO:0000339" />
<listOfParameters>
```

SBO::Systems Biology Ontology

- SBO:0000000 - sbo
  - SBO:0000002 - quantitative parameter
    - SBO:0000226 - number density
    - SBO:0000227 - mass density
    - SBO:0000255 - physical characteristic
    - SBO:0000256 - biochemical parameter
      - SBO:0000009 - kinetic constant
        - SBO:0000016 - unimolecular rate constant
        - SBO:0000017 - bimolecular rate constant
          - SBO:0000023 - forward bimolecular rate constant
            - SBO:0000036 - [illegible]
              - SBO:0000[illegible]
              - SBO:0000339 - bimolecular association rate constant
            - SBO:0000067 - forward bimolecular rate constant, discrete case

```
2" sboTerm="SBO:0000010" />

2" sboTerm="SBO:0000011" />
```

“forward bimolecular rate constant, continuous case”

MIRIAM

http://www.biomodels.net/miriam

SBML provides syntax

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    <species compartment="cytosol" id="P"  initi
    <species compartment="cytosol" id="S"  initi
    <species compartment="cytosol" id="E"  initi
</listOfSpecies>
<listOfReactions>
    <reaction id="veq">
        <listOfReactants>
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        <listOfProducts>
            <speciesReference species="ES"/>
        </listOfProducts>
        <kineticLaw>
            <math xmlns="http://www.w3.org/1998/
                <apply>
                    <times/>
                    <ci>cytosol</ci>
                    <apply>
                        <minus/>
                        <apply>
                            <times/>
                            <ci>kon</ci>
                            <ci>E</ci>
                            <ci>S</ci>
                        </apply>
                        <apply>
                            <times/>
                            <ci>koff</ci>
                            <ci>ES</ci>
```

SBML provides syntax

- Raw models alone are insufficient
  - SBML only provides syntax

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    <species compartment="cytosol" id="P"  initi
    <species compartment="cytosol" id="S"  initi
    <species compartment="cytosol" id="E"  initi
</listOfSpecies>
<listOfReactions>
    <reaction id="veq">
        <listOfReactants>
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        <listOfProducts>
            <speciesReference species="ES"/>
        </listOfProducts>
        <kineticLaw>
            <math xmlns="http://www.w3.org/1998/
                <apply>
                    <times/>
                    <ci>cytosol</ci>
                    <apply>
                        <minus/>
                        <apply>
                            <times/>
                            <ci>kon</ci>
                            <ci>E</ci>
                            <ci>S</ci>
                        </apply>
                        <apply>
                            <times/>
                            <ci>koff</ci>
                            <ci>ES</ci>
```

SBML provides syntax

- Raw models alone are insufficient
  - SBML only provides syntax

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    <species compartment="cytosol" id="P"  initi
    <species compartment="cytosol" id="S"  initi
    <species compartment="cytosol" id="E"  initi
</listOfSpecies>
<listOfReactions>
    <reaction id="veq">
        <listOfReactants>
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        <listOfProducts>
            <speciesReference species="ES"/>
        </listOfProducts>
        <kineticLaw>
            <math xmlns="http://www.w3.org/1998/
                <apply>
                    <times/>
                    <ci>cytosol</ci>
                    <apply>
                        <minus/>
                        <apply>
                            <times/>
                            <ci>kon</ci>
                            <ci>E</ci>
                            <ci>S</ci>
                        </apply>
                        <apply>
                            <times/>
                            <ci>koff</ci>
                            <ci>ES</ci>
```

SBML provides syntax

- Raw models alone are insufficient
  - SBML only provides syntax

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    <species compartment="cytosol" id="P"  initi
    <species compartment="cytosol" id="S"  initi
    <species compartment="cytosol" id="E"  initi
</listOfSpecies>
<listOfReactions>
    <reaction id="kon">
        <listOfReactants>
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        <listOfProducts>
            <speciesReference species="ES"/>
        </listOfProducts>
        <kineticLaw>
            <math xmlns="http://www.w3.org/1998/
                <apply>
                    <times/>
                    <ci>cytosol</ci>
                    <apply>
                        <minus/>
                        <apply>
                            <times/>
                            <ci>kon</ci>
                            <ci>E</ci>
                            <ci>S</ci>
                        </apply>
                        <apply>
                            <times/>
                            <ci>koff</ci>
                            <ci>ES</ci>
```

Low info content

Unregulated

SBML provides syntax

- Raw models alone are insufficient
  - SBML only provides syntax
- Need standard schemes for machine-readable annotations
  - For authorship, publication info
  - For links to other data resources
  - For semantics of mathematics

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    <species compartment="cytosol" id="P"  initi
    <species compartment="cytosol" id="S"  initi
    <species compart          id="E"  initi
</listOfSpecies
<listOfReactio
    <reaction
        <listOfReac
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        <listOfProducts>
            <speciesReference species="ES"/>
        </listOfProducts>
        <kineticLaw>
<kineticLaw>
    <math xmlns="http://www.w3.org/1998/
        <apply>
            <times/>
            <ci>cytosol</ci>
            <apply>
                <minus/>
                <apply>
                    <times/>
                    <ci>kon</ci>
                    <ci>E</ci>
                    <ci>S</ci>
                ly>
                pply>
                    <times/>
                    <ci>koff</ci>
                    <ci>ES</ci>
```

Low info content

Unregulated

SBML provides syntax

- Raw models alone are insufficient
  - SBML only provides syntax
- Need standard schemes for machine-readable annotations
  - For authorship, publication info
  - For links to other data resources
  - For semantics of mathematics
- Need common guidelines for minimal model quality and content

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    <species compartment="cytosol" id="P"  initi
    <species compartment="cytosol" id="S"  initi
    <species compart       "cytosol" id="E"  initi
</listOfSpecies
<listOfReactio
    <reaction i
        <listOfReac
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        <listOfProducts>
            <speciesReference species="ES"/>
        </listOfProducts>
        <kineticLaw>
<kineticLaw>
    <math xmlns="http://www.w3.org/1998/
        <apply>
            <times/>
            <ci>cytosol</ci>
            <apply>
                <minus/>
                <apply>
                    <times/>
                    <ci>kon</ci>
                    <ci>E</ci>
                    <ci>S</ci>
                ly>
                pply>
                    <times/>
                    <ci>koff</ci>
                    <ci>ES</ci>
```

Low info content

Unregulated

SBML provides syntax

- Raw models alone are insufficie
  - SBML only provides syntax
- Need standard schemes for machine-readable annotations
  - For authorship, publication info
  - For links to other data resources
  - For semantics of mathematics
- Need common guidelines for minimal model quality and content

Defined by SBML

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    ies compartment="cytosol" id="P"  initi
    es compartment="cytosol" id="S"  initi
    es compartment="cytosol" id="E"  initi
    pecies>
    ofReactions>
    <reaction id="veq">
        <listOfReactants>
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        <listOfProducts>
            <speciesReference species="ES"/>
        </listOfProducts>
        <kineticLaw>
            <math xmlns="http://www.w3.org/1998/
                <apply>
                    <times/>
                    <ci>cytosol</ci>
                    <apply>
                        <minus/>
                        <apply>
                            <times/>
                            <ci>kon</ci>
                            <ci>E</ci>
                            <ci>S</ci>
                        </apply>
                        <apply>
                            <times/>
                            <ci>koff</ci>
                            <ci>ES</ci>
                        </apply>
```

SBML provides syntax

- Raw models alone are insufficie
  - SBML only provides syntax
- Need standard schemes for machine-readable annotations
  - For authorship, publication info
  - For links to other data resources
  - For semantics of mathematics
- Need common guidelines for minimal model quality and content

Defined by SBML

Defined by MIRIAM

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    ies compartment="cytosol" id="P"  initi
    es compartment="cytosol" id="S"  initi
    es compartment="cytosol" id="E"  initi
    pecies>
    ofReactions>
    <reaction id="veq">
        <listOfReactants>
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        <listOfProducts>
            <speciesReference species="ES"/>
        </listOfProducts>
        <kineticLaw>
            <math xmlns="http://www.w3.org/1998/
                <apply>
                    <times/>
                    <ci>cytosol</ci>
                    <apply>
                        <minus/>
                        <apply>
                            <times/>
                            <ci>kon</ci>
                            <ci>E</ci>
                            <ci>S</ci>
                        </apply>
                        <apply>
                            <times/>
                            <ci>koff</ci>
                            <ci>ES</ci>
```

SBML provides syntax

- Raw models alone are insufficie
  - SBML only provides syntax
- Need standard schemes for machine-readable annotations
  - For authorship, publication info
  - For links to other data resource
  - For semantics of mathematics
- Need common guidelines for minimal model quality and content

Defined by SBML

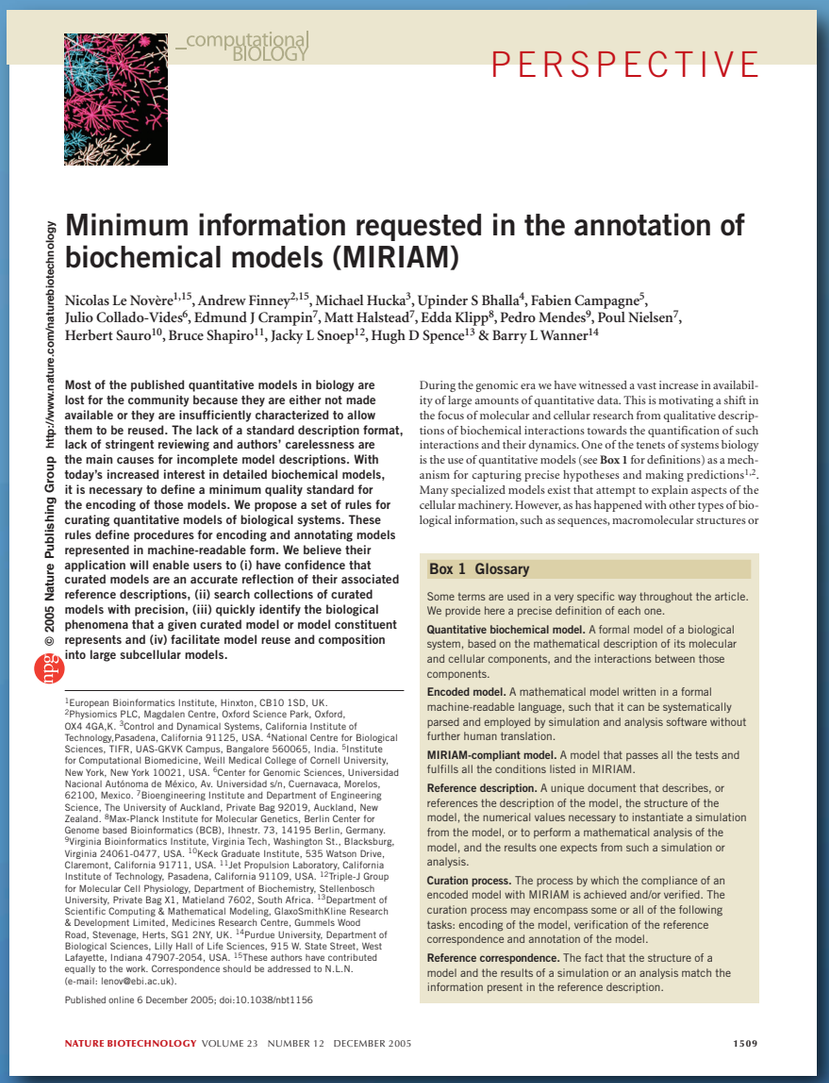
Defined by SBO

Defined by MIRIAM

```
<listOfSpecies>
    <species compartment="cytosol" id="ES" initi
    ies compartment="cytosol" id="P"  initi
    s compartment="cytosol" id="S"  initi
    s compartment="cytosol" id="E"  initi
    pecies>
    ofReactions>
    <reaction id="veq">
        <listOfReactants>
            <speciesReference species="E"/>
            <speciesReference species="S"/>
        </listOfReactants>
        listOfProducts>
        speciesReference species="ES"/>
        fProducts>
        cLaw>
        math xmlns="http://www.w3.org/1998/
            <apply>
                <times/>
                <ci>cytosol</ci>
                <apply>
                    <minus/>
                    <apply>
                        <times/>
                        <ci>kon</ci>
                        <ci>E</ci>
                        <ci>S</ci>
                    </apply>
                    <apply>
                        <times/>
                        <ci>koff</ci>
                        <ci>ES</ci>
```

Need standardize annotations added to a model

- "**M**inimum **i**nformation **r**equested **i**n the **a**nnotation of biochemical **m**odels"
- Proposed guidelines for annotation and curation of **quantitative** models
  - Specifically about encoding & annotation
  - Limited to models that can be simulated
- **Not** specific to SBML—applicable to any structured model format
- Community effort, originally lead by Andrew Finney and Nicolas Le Novère



_computational BIOLOGY

PERSPECTIVE

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 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.

[1]European Bioinformatics Institute, Hinxton, CB10 1SD, UK. [2]Physiomics PLC, Magdalen Centre, Oxford Science Park, Oxford, OX4 4GA,K. [3]Control and Dynamical Systems, California Institute of Technology,Pasadena, California 91125, USA. [4]National Centre for Biological Sciences, TIFR, UAS-GKVK Campus, Bangalore 560065, India. [5]Institute for Computational Biomedicine, Weill Medical College of Cornell University, New York, New York 10021, USA. [6]Center for Genomic Sciences, Universidad Nacional Autónoma de México, Av. Universidad s/n, Cuernavaca, Morelos, 62100, Mexico. [7]Bioengineering Institute and Department of Engineering Science, The University of Auckland, Private Bag 92019, Auckland, New Zealand. [8]Max-Planck Institute for Molecular Genetics, Berlin Center for Genome based Bioinformatics (BCB), Ihnestr. 73, 14195 Berlin, Germany. [9]Virginia Bioinformatics Institute, Virginia Tech, Washington St., Blacksburg, Virginia 24061-0477, USA. [10]Keck Graduate Institute, 535 Watson Drive, Claremont, California 91711, USA. [11]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA. [12]Triple-J Group for Molecular Cell Physiology, Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa. [13]Department of Scientific Computing & Mathematical Modeling, GlaxoSmithKline Research & Development Limited, Medicines Research Centre, Gummels Wood Road, Stevenage, Herts, SG1 2NY, UK. [14]Purdue University, Department of Biological Sciences, Lilly Hall of Life Sciences, 915 W. State Street, West Lafayette, Indiana 47907-2054, USA. [15]These authors have contributed equally to the work. Correspondence should be addressed to N.L.N. (e-mail: lenov@ebi.ac.uk).

Published online 6 December 2005; doi:10.1038/nbt1156

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.

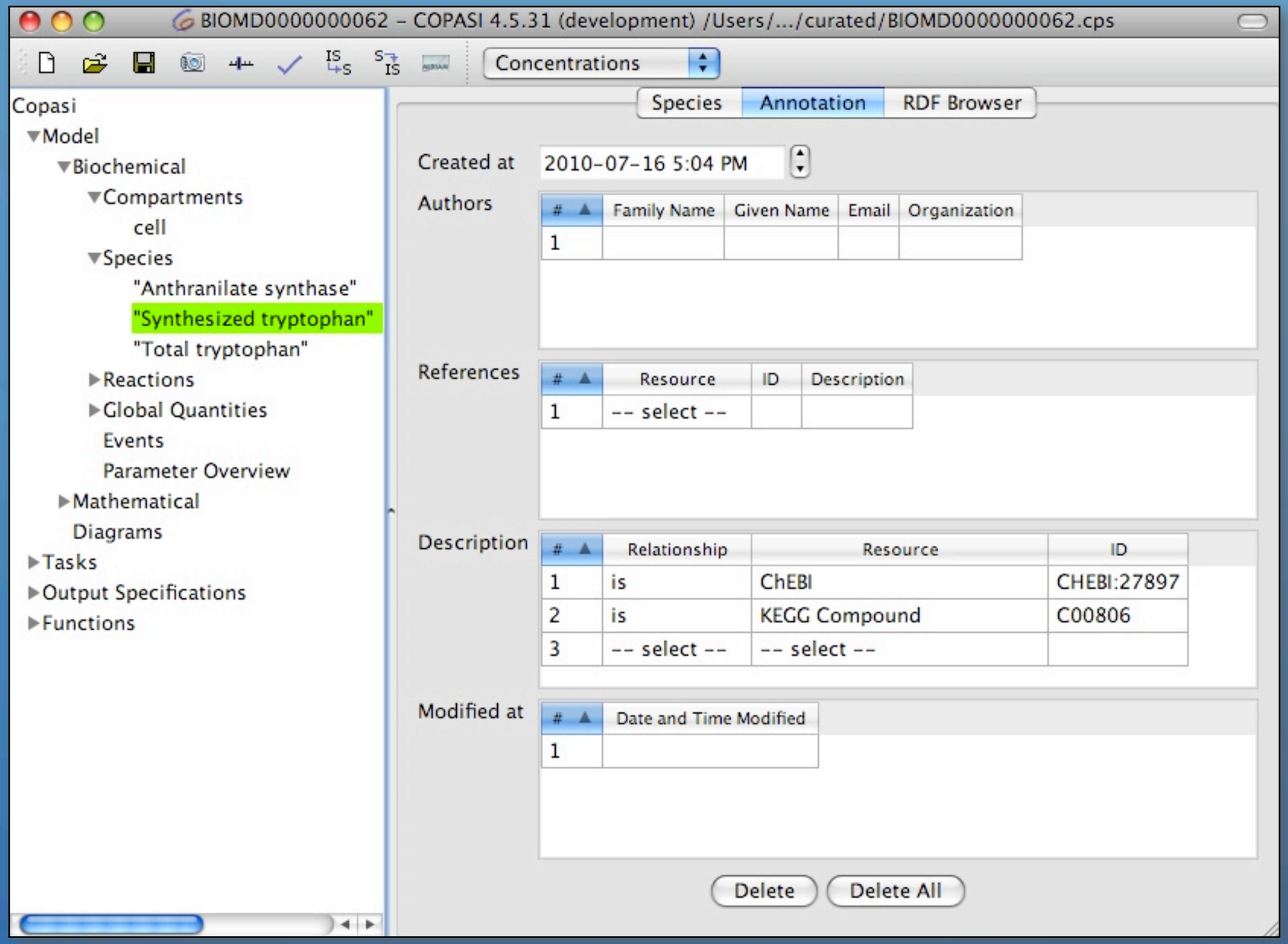
Reference description. A unique document that describes, or references the description of the model, the structure of the model, the numerical values necessary to instantiate a simulation from the model, or to perform a mathematical analysis of the model, and the results one expects from such a simulation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

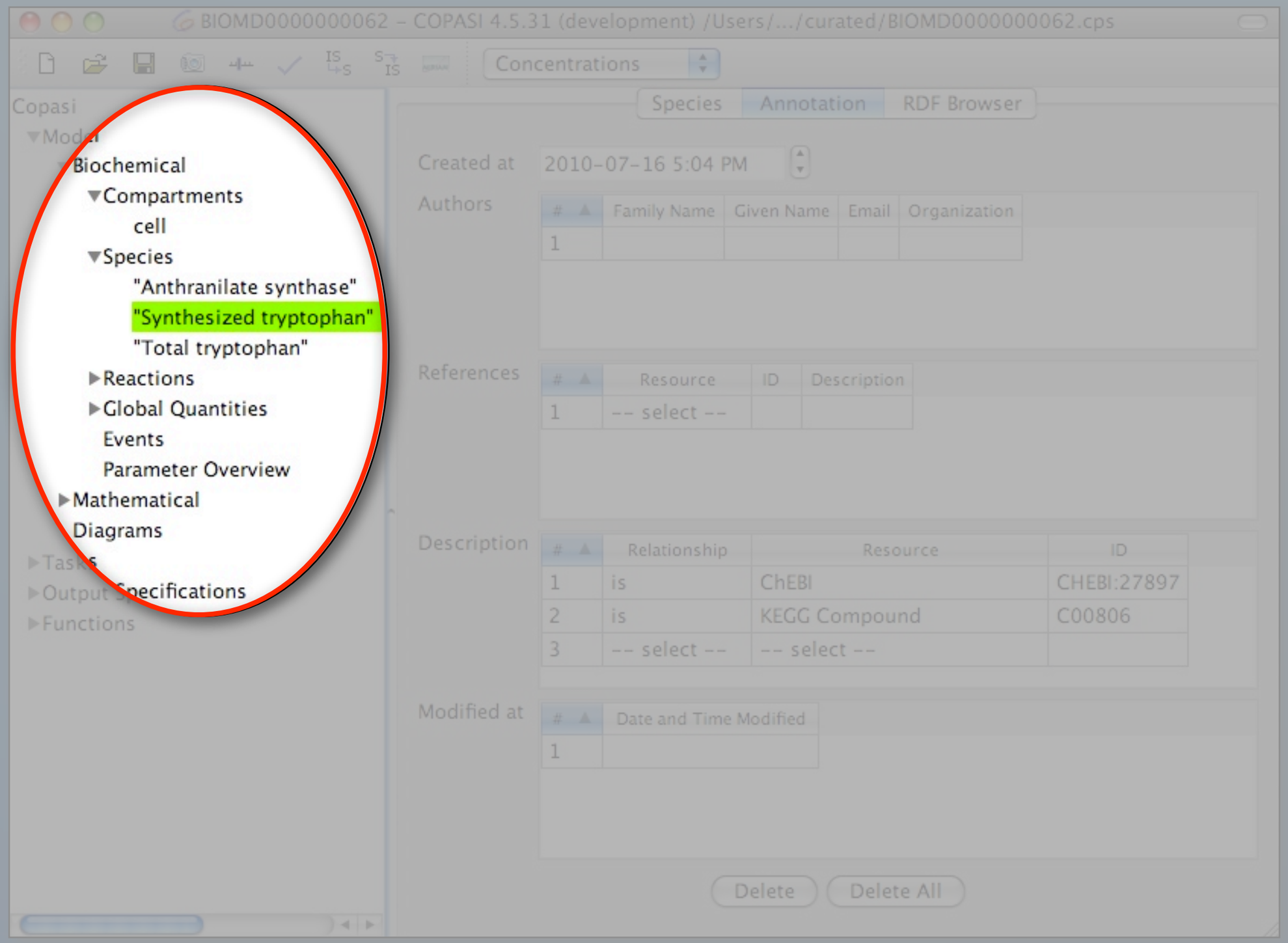
Reference correspondence. The fact that the structure of a model and the results of a simulation or an analysis match the information present in the reference description.

Le Novère et al., *Nat. Biotech.* 23 '05

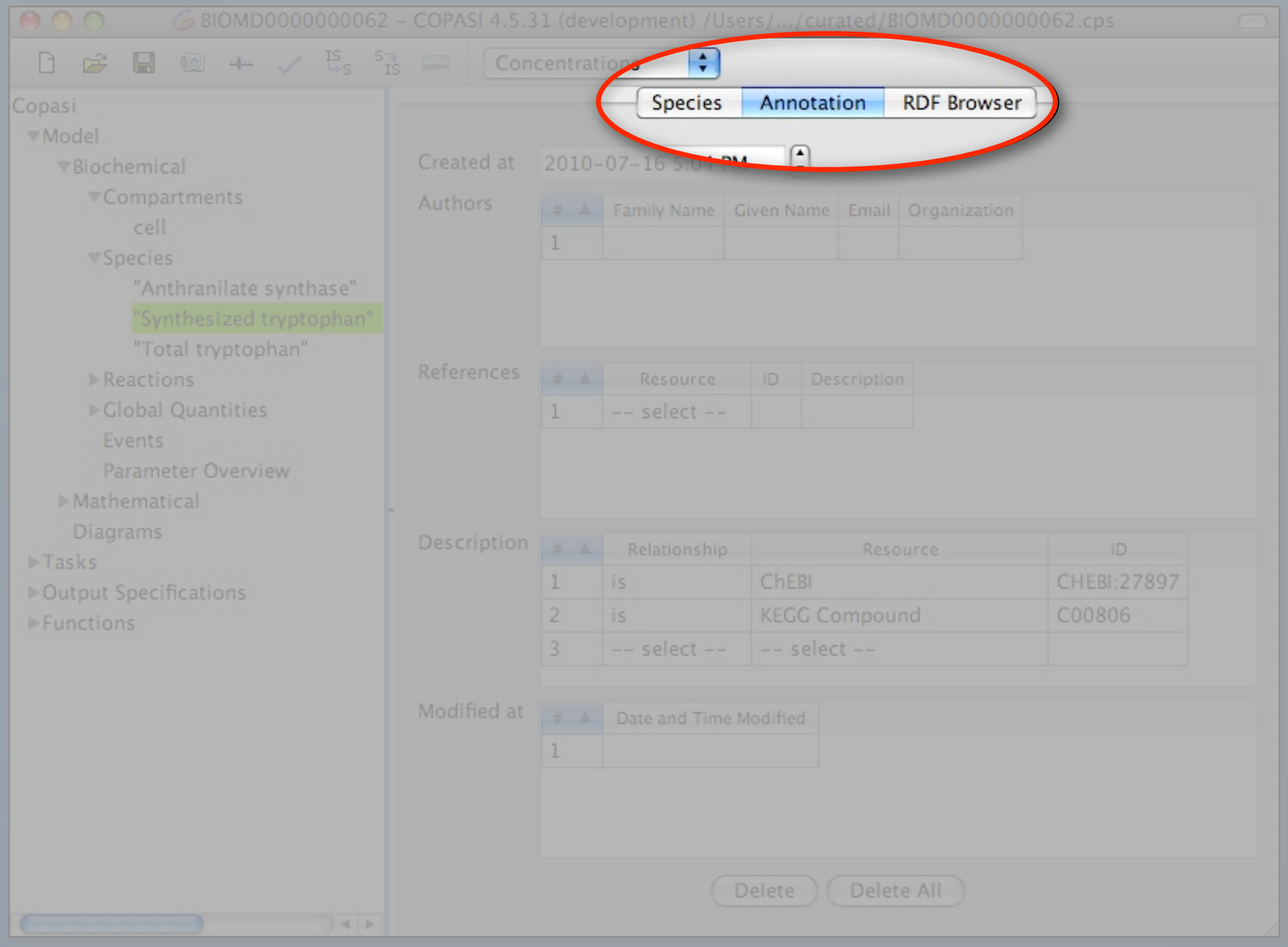
Example interface in software: COPASI



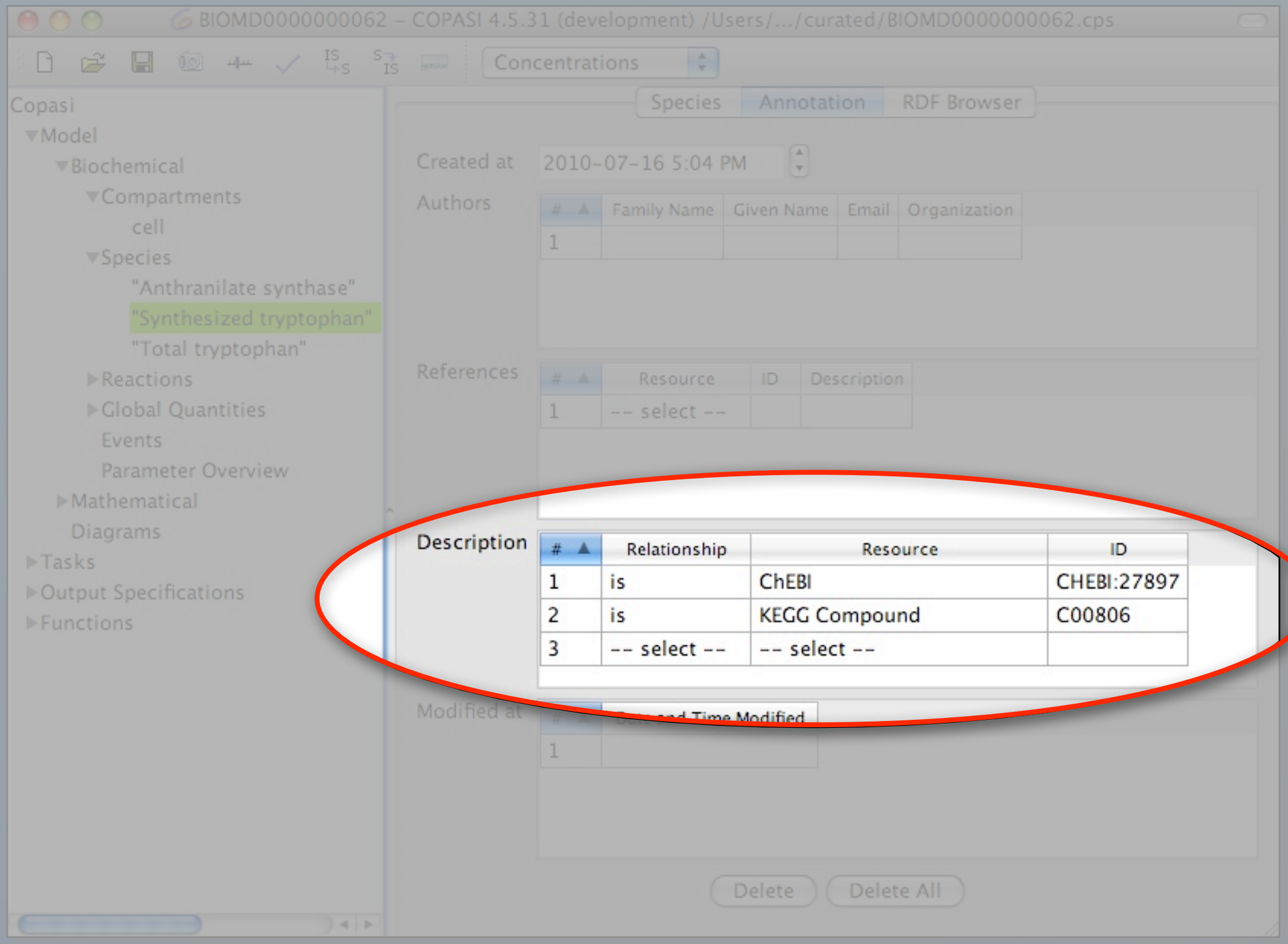
Example interface in software: COPASI



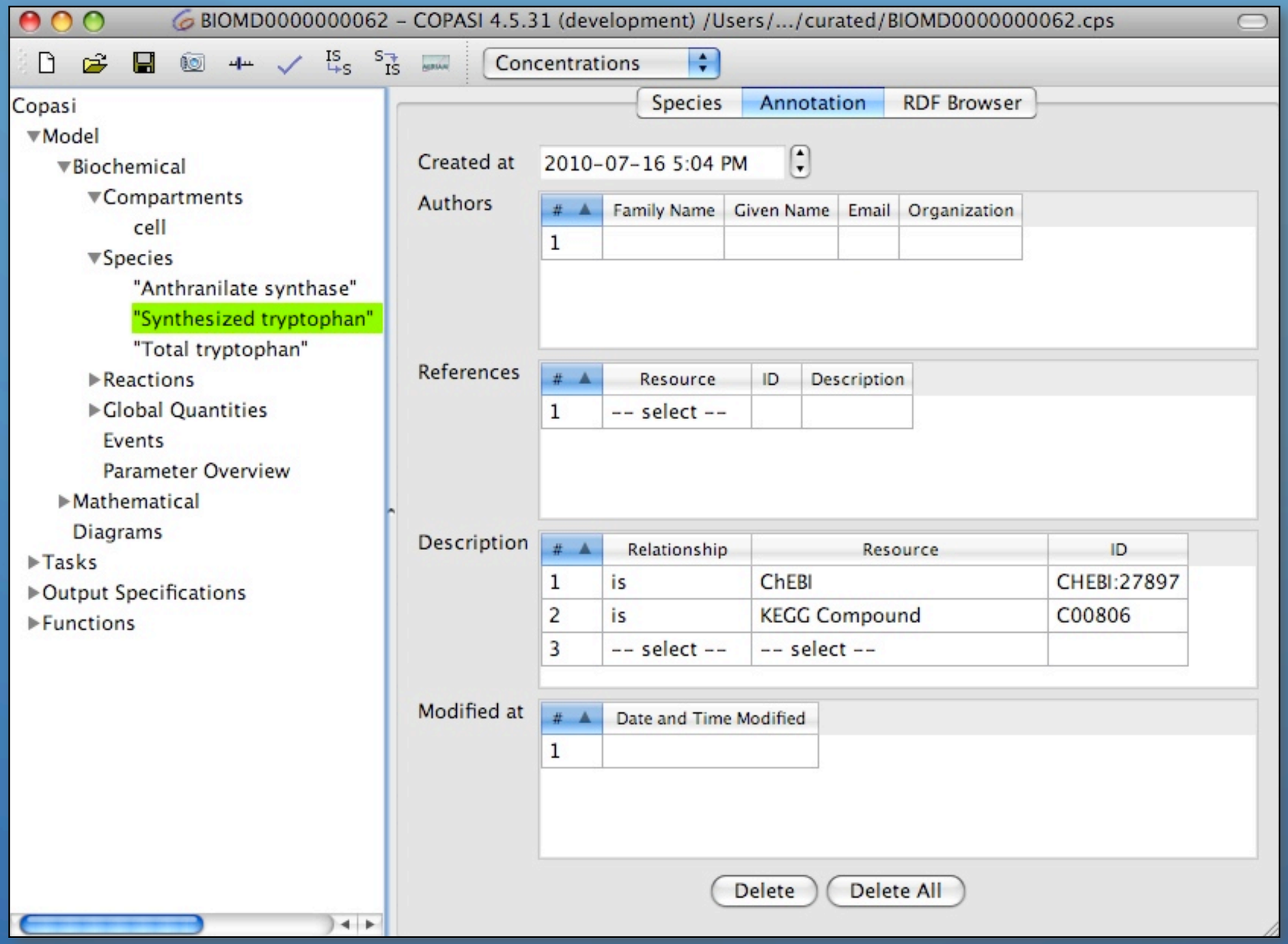
Example interface in software: COPASI



Example interface in software: COPASI



Example interface in software: COPASI



Example of semantic annotation in SBML

```
<species metaid="metaid_0000009" id="species_3" compartment="c_1">
  <annotation>
    <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
             xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
             xmlns:bqmodel="http://biomodels.net/model-qualifiers/">
      <rdf:Description rdf:about="#metaid_0000009">
        <bqbiol:is>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:obo.chebi:CHEBI%3A15996"/>
            <rdf:li rdf:resource="urn:miriam:kegg.compound:C00044"/>
          </rdf:Bag>
        </bqbiol:is>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

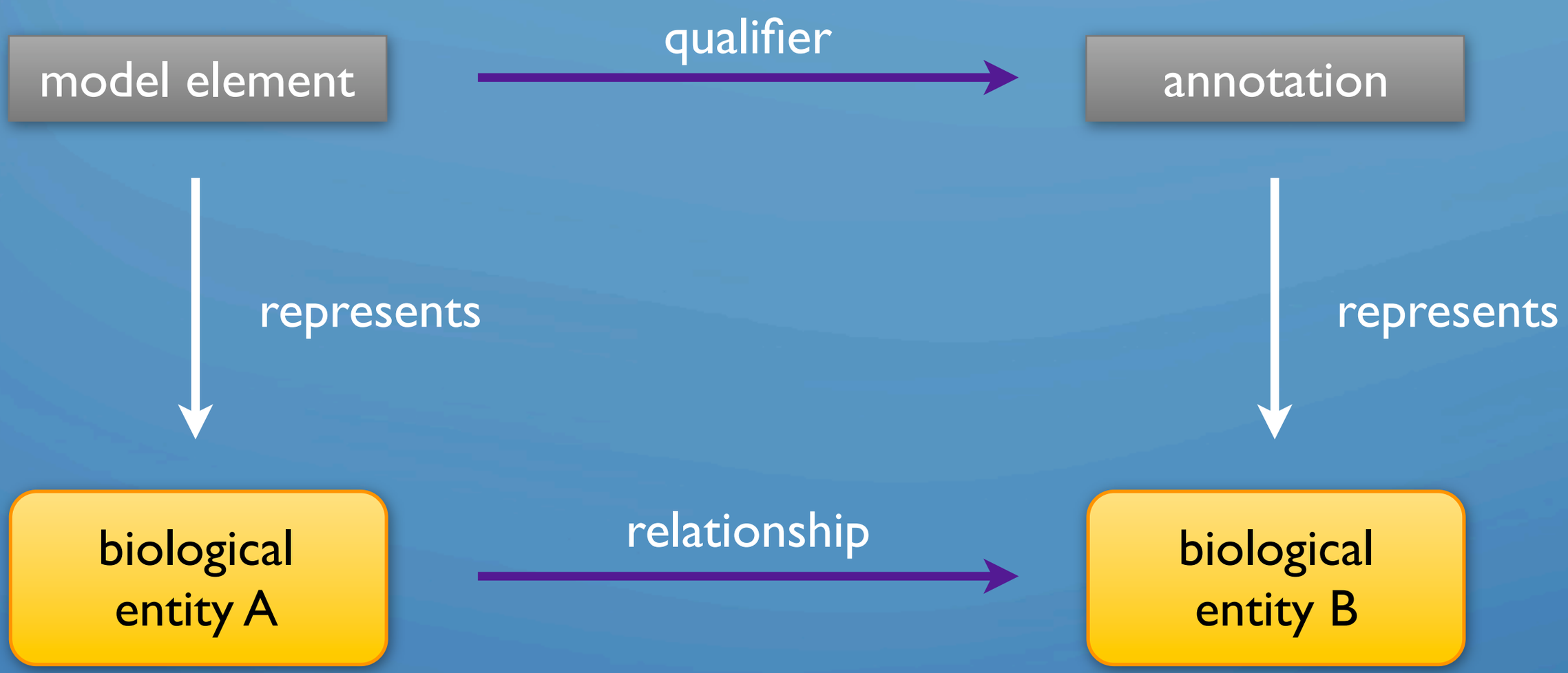
Example of semantic annotation in SBML

```
<species metaid="metaid_0000009" id="species_3" compartment="c_1">
  <annotation>
    <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
             xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
             xmlns:bqmodel="http://biomodels.net/model-qualifiers/">
      <rdf:Description rdf:about="#metaid_0000009">
        <bqbiol:is>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:obo.chebi:CHEBI%3A15996"/>
            <rdf:li rdf:resource="urn:miriam:kegg.compound:C00044"/>
          </rdf:Bag>
        </bqbiol:is>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

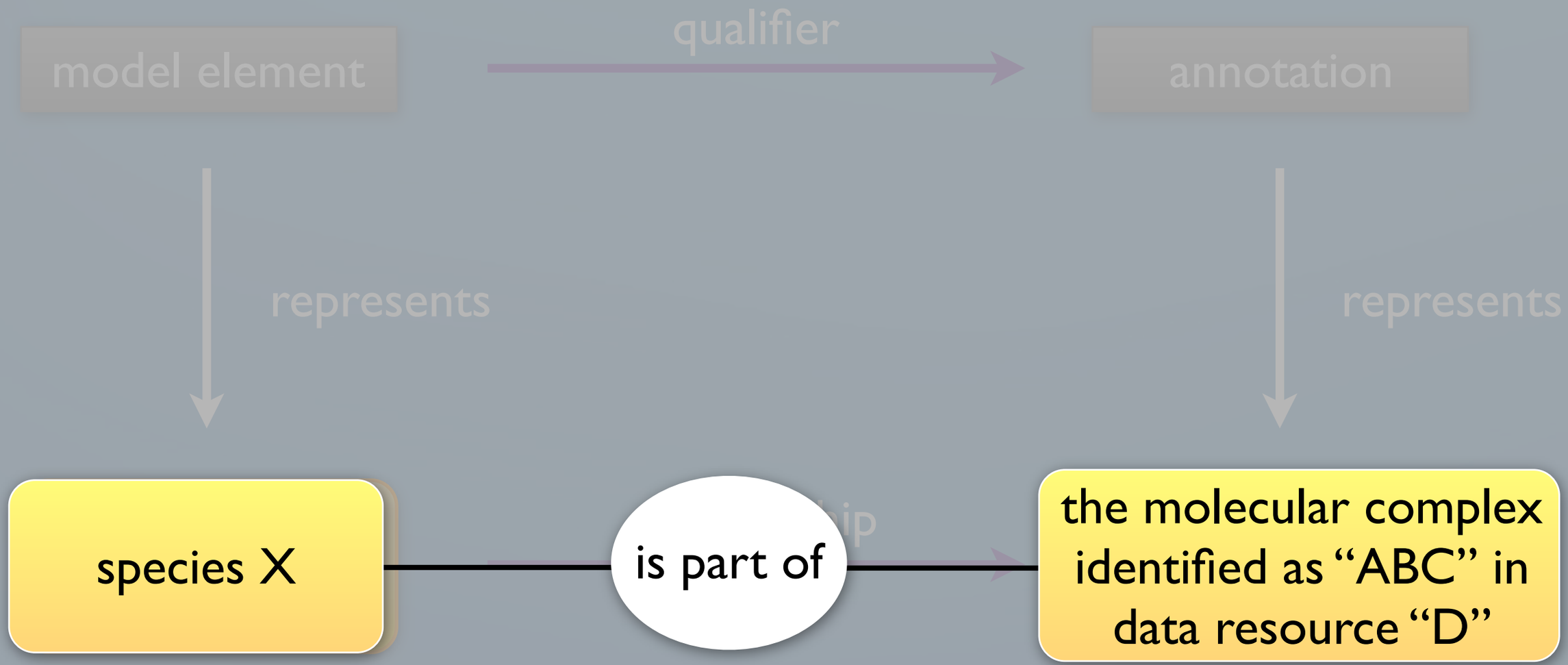
Example of semantic annotation in SBML

```
<species metaid="metaid_0000009" id="species_3" compartment="c_1">
  <annotation>
    <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
             xmlns:bqbiol="http://biomodels.net/biology-qualifiers/"
             xmlns:bqmodel="http://biomodels.net/model-qualifiers/">
      <rdf:Description rdf:about="#metaid_0000009">
        <bqbiol:is>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:obo.chebi:CHEBI%3A15996"/>
            <rdf:li rdf:resource="urn:miriam:kegg.compound:C00044"/>
          </rdf:Bag>
        </bqbiol:is>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

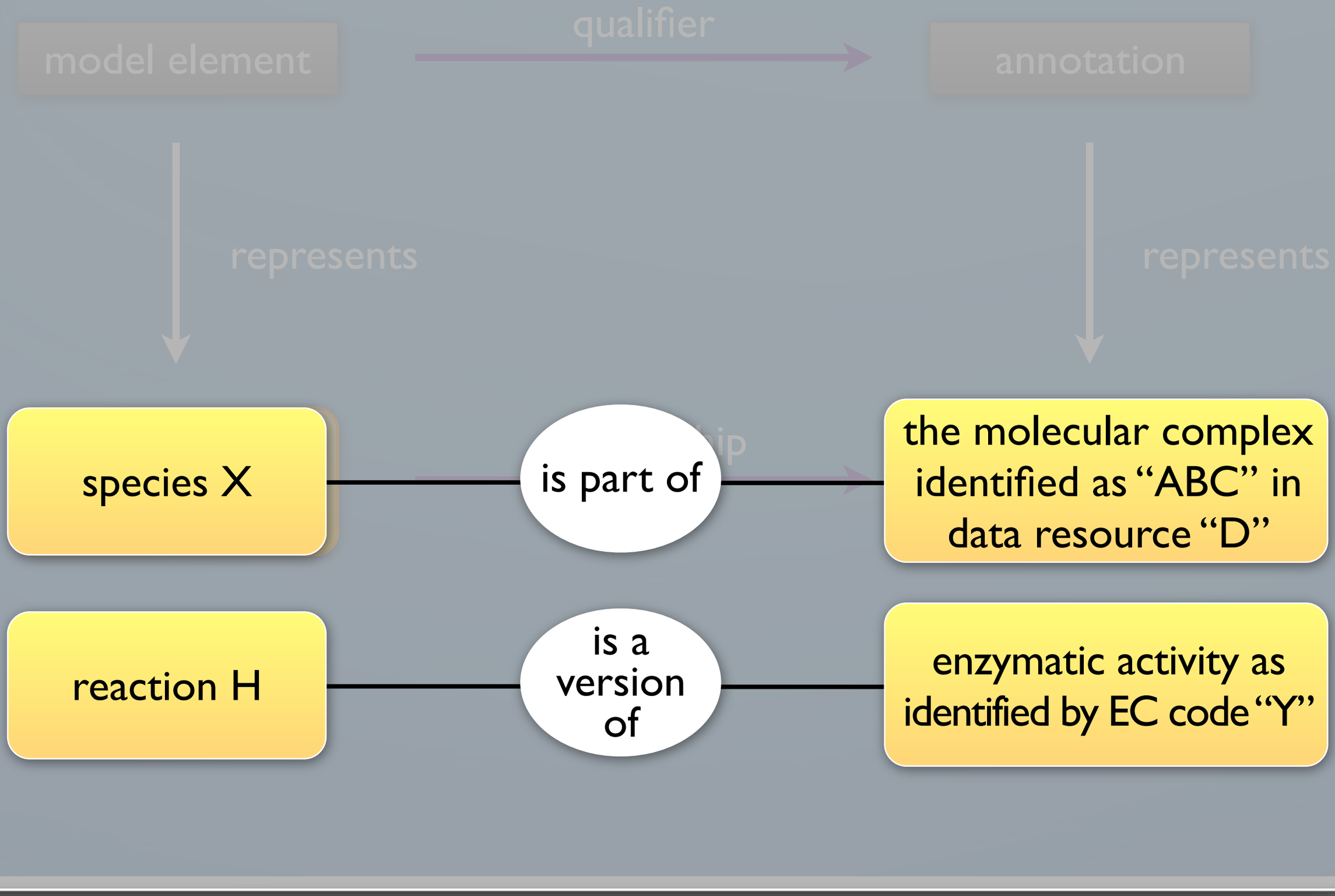
What are the relationships?



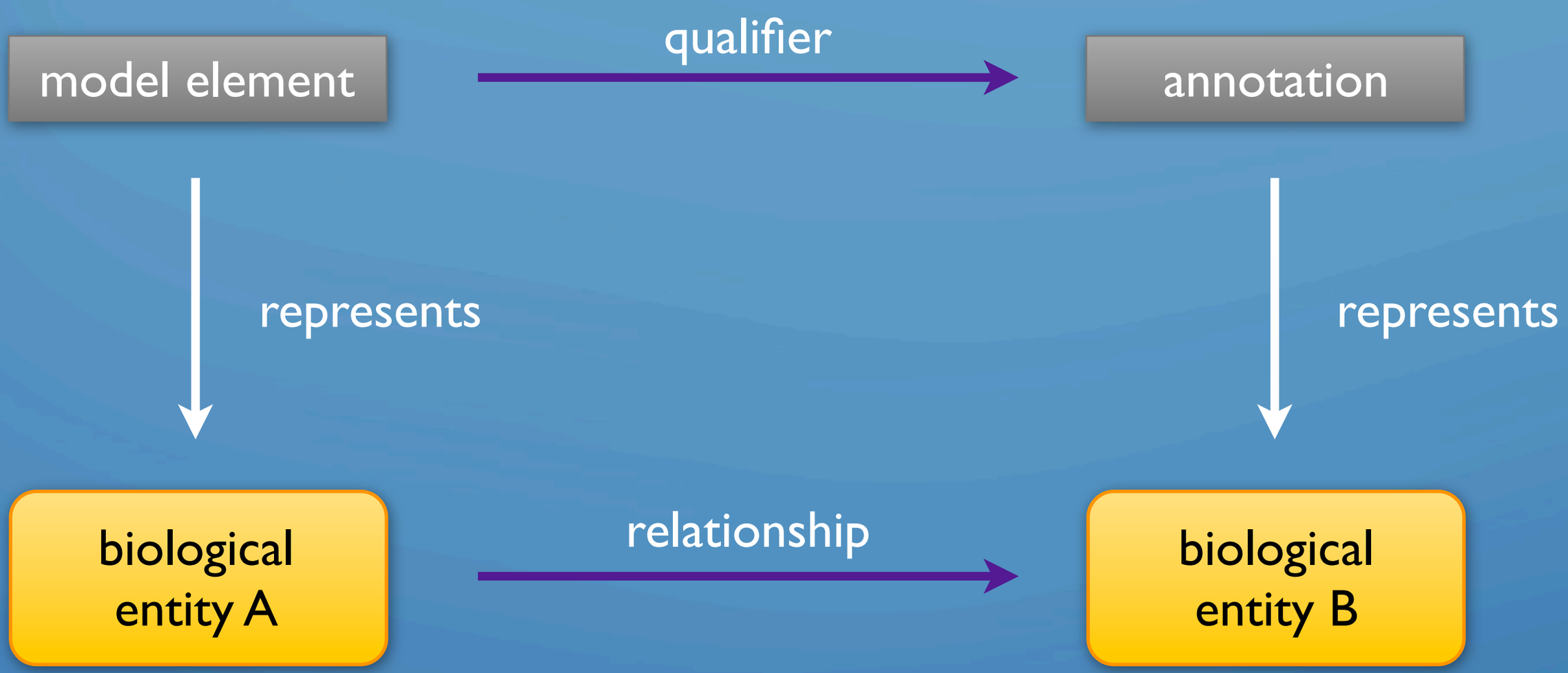
What are the relationships?



What are the relationships?



What are the relationships?

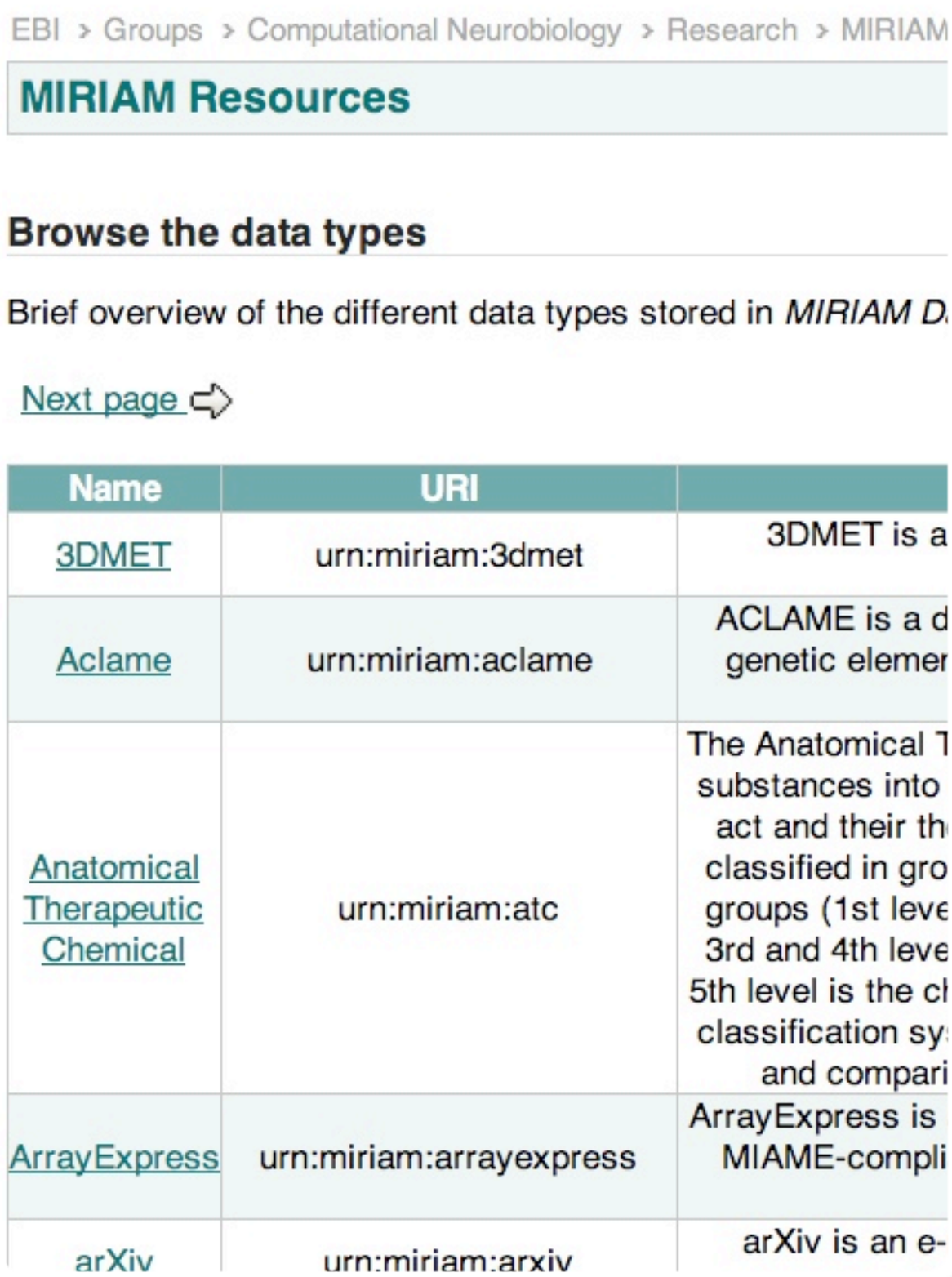


What are the relationships?

The relationship is rarely “is”, hence the need for a dictionary of qualifiers

MIRIAM Resources

- Provides 3 key resources:
  - Central catalog of agreed-upon standard URNs for data types
  - Facility for the community to add and update entries
    - People can "own" an entry
  - Resolution services for software via standard protocols (web services)
- Developed by Camille Laibe in Nicolas Le Novère's group @ EBI
- **biomodels.net/miriam**

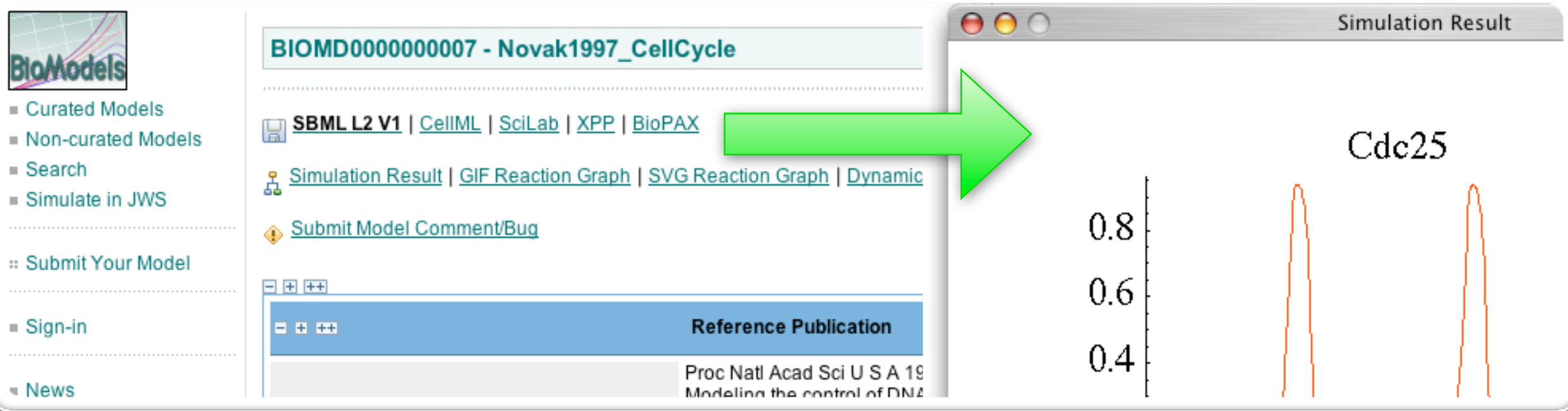


MIASE

http://biomodels.net/miase/

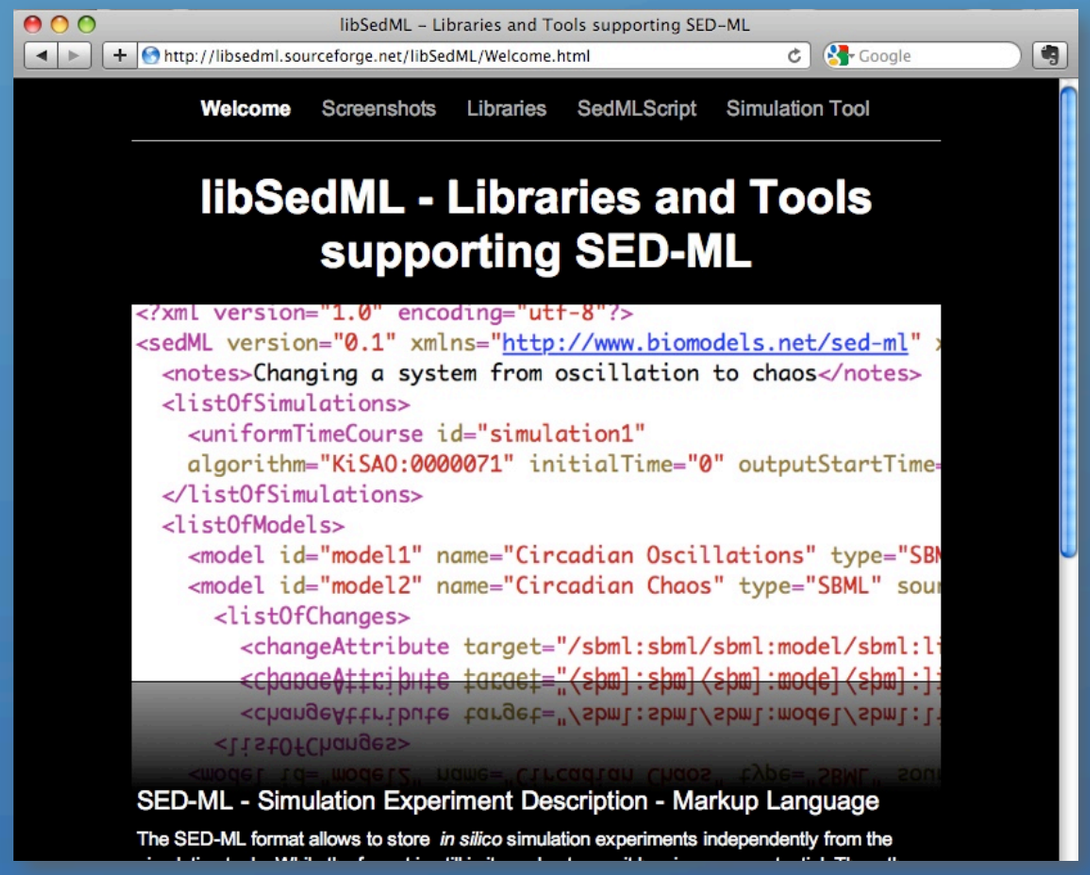
How was a given simulation output obtained?

- Procedures are not captured in SBML
  - It's a **model** representation formats, **not a simulation language**
- Want to record essential steps to go from model to output data
  - Essentially explaining how a figure in a paper (or database) was obtained
  - Format must be independent of any particular simulation tool
  - Format must be machine-readable to allow automation



MIASE & SED-ML

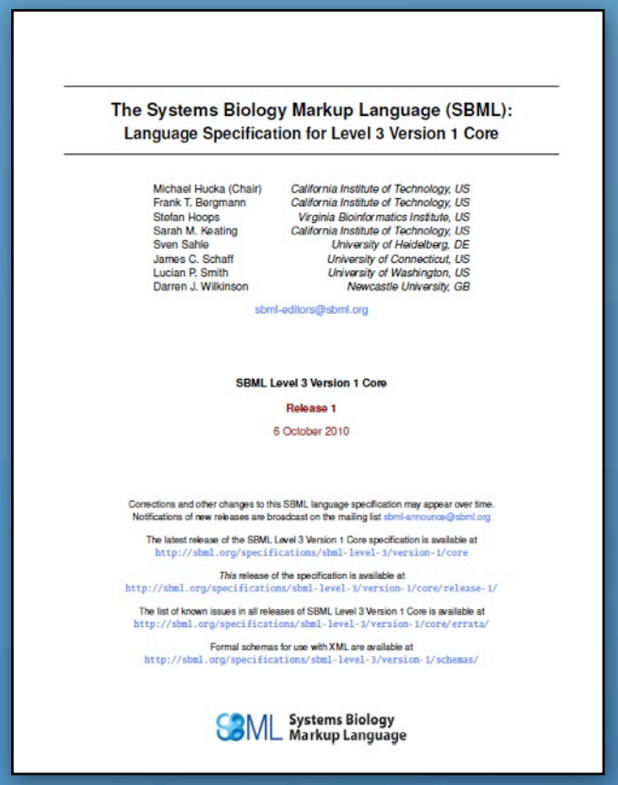
- MIASE = "**Mi**nimum **I**nformation **a**bout a **S**imulation **E**xperiment"
  - Allow person or software to reproduce the results of a simulation
    - E.g.: software used, procedures used, transformations applied, etc.
- SED-ML = "**S**imulation **E**xperiment **D**escription **M**arkup **L**anguage"
  - Formal representation satisfying the MIASE guidelines
  - The libSedML project:
    - **http://libsedml.sf.net/**



SBML Level 3

What are SBML "Levels"?

- Specification document available from **http://sbml.org/Documents**
- Newest: **Level 3 Version 1 Core**
  - Oct. 2010



The Systems Biology Markup Language (SBML):
Language Specification for Level 3 Version 1 Core

Michael Hucka (Chair) — *California Institute of Technology, US*
Frank T. Bergmann — *California Institute of Technology, US*
Stefan Hoops — *Virginia Bioinformatics Institute, US*
Sarah M. Keating — *California Institute of Technology, US*
Sven Sahle — *University of Heidelberg, DE*
James C. Schaff — *University of Connecticut, US*
Lucian P. Smith — *University of Washington, US*
Darren J. Wilkinson — *Newcastle University, GB*

sbml-editors@sbml.org

SBML Level 3 Version 1 Core

Release 1

6 October 2010

Corrections and other changes to this SBML language specification may appear over time. Notifications of new releases are broadcast on the mailing list sbml-announce@sbml.org

The latest release of the SBML Level 3 Version 1 Core specification is available at http://sbml.org/specifications/sbml-level-3/version-1/core

This release of the specification is available at http://sbml.org/specifications/sbml-level-3/version-1/core/release-1/

The list of known issues in all releases of SBML Level 3 Version 1 Core is available at http://sbml.org/specifications/sbml-level-3/version-1/core/errata/

Formal schemas for use with XML are available at http://sbml.org/specifications/sbml-level-3/version-1/schemas/

SBML Systems Biology Markup Language

- About SBML "Levels":
  - Levels help manage significant restructuring of SBML architecture
  - Levels coexist
    - E.g., Level 2 models will remain valid and exist for a long time
  - A Level is *not* solely a vertical change (i.e., more features)—there is horizontal change too (i.e., changes to existing elements)

Modularizing SBML

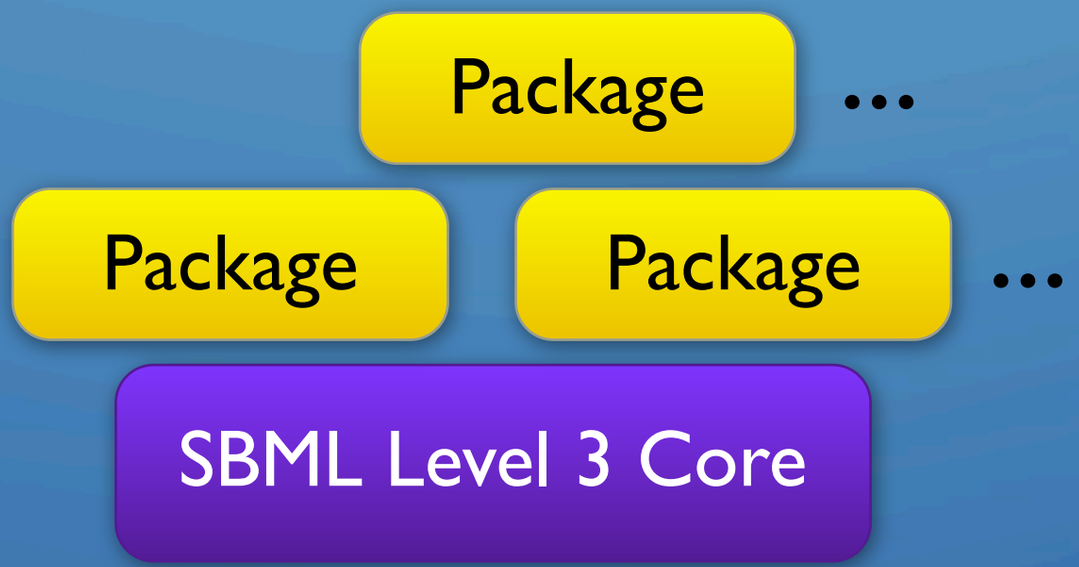
- The future of SBML
  - Major extensions have been requested and are sorely needed
  - But SBML is complicated enough
    - Only the largest-scale efforts have implemented most of Level 2
- Idea (a natural one): modularize
  - Layer additional language extensions on top of a core
    - Similar to XML: XML is core, then there's MathML, etc.
  - Additional benefits:
    - Can decouple development of individual extensions
    - Software can recognize & work with extensions it needs

Evolution of features took time & practical experience

Level 1	Level 2	Level 3
predefined math functions	user-defined functions	user-defined functions
text-string math notation	MathML subset	MathML subset
reserved namespaces for annotations	no reserved namespaces for annotations	no reserved namespaces for annotations
no controlled annotation scheme	RDF-based controlled annotation scheme	RDF-based controlled annotation scheme
no discrete events	discrete events	discrete events
default values defined	default values defined	no default values
monolithic	monolithic	modular

SBML Level 3

- *SBML Level 3* is modular:
  - "Core" defines common aspects
  - "Packages" add optional features
    - Models declare which packages they use
    - Tools can tell their users which packages they support



SBML Level 3 packages in development

Package	Specification status
Diagram layout	Finalized & updated for L3; needs review
Diagram rendering	Finalized & updated for L3; needs review
Multi* species	Finalized & updated for L3; needs review
Hierarchical composition	New L3 specification now under discussion
Qualitative models	Draft specification proposed
Groups	Draft specification proposed
Steady-state models	Draft specification proposed
Spatial geometry	Draft specification proposed
Annotations	Specification in development
Distribution & ranges	Past spec. proposed; needs update for final L3
Arrays & sets	Past spec. proposed; needs significant update
Spatial diffusion	No specification yet
Dynamic structures	No specification yet

Hierarchical Model Composition

Composition as it is currently envisioned

- Goals of model composition extension is to support the following:
  - Models containing/including other models ("hierarchical models")
    - Arbitrary nesting depths
    - Models may be defined internally or externally (in separate files)
  - Allow linking entities (species, compartments, reactions) between models
    - Support links between internal and external models
  - Allow definition of explicit interfaces ("ports")
    - But make use of ports optional

Models & submodels

Current SBML

Model "A"

- Compartments ...
- Species ...
- Parameters ...
- Reactions ...

Models & submodels

Current SBML

Model "A"

- Compartments ...
- Species ...
- Parameters ...
- Reactions ...

With model composition

Model "A"

- Compartments ...
- Species ...
- Parameters ...
- Reactions ...

Model "B"

- Compartments ...
- Species ...
- Parameters ...
- Reactions ...

Model "C"

- Compartments ...
- Species ...
- Parameters ...
- Reactions ...

More scenarios

Model "A"

Compartments ...

Species ...

Parameters ...

Reactions ...

Model "B"

Compartments ...

Species ...

Parameters ...

Reactions ...

Model "B"

Compartments ...

Species ...

Parameters ...

Reactions ...

Model "B"

Compartments ...

Species ...

Parameters ...

Reactions ...

...

More scenarios

Model "A"

Compartments ...

Species ...

Parameters ...

Reactions ...

Model "B"

Compartments ...

Species ...

Parameters ...

Reactions ...

Model "C"

Compartments ...

Species ...

Parameters ...

Reactions ...

Model "D"

Compartments ...

Species ...

Parameters ...

Reactions ...

More scenarios

Separate files/resources

Model "A"

Compartments ...

Species ...

Parameters ...

Reactions ...

Model "B"

Model "C"

Model "D"

Model "B"

Compartments ...

Species ...

Parameters ...

Reactions ...

Model "C"

Compartments ...

Species ...

Parameters ...

Reactions ...

Model "D"

Compartments ...

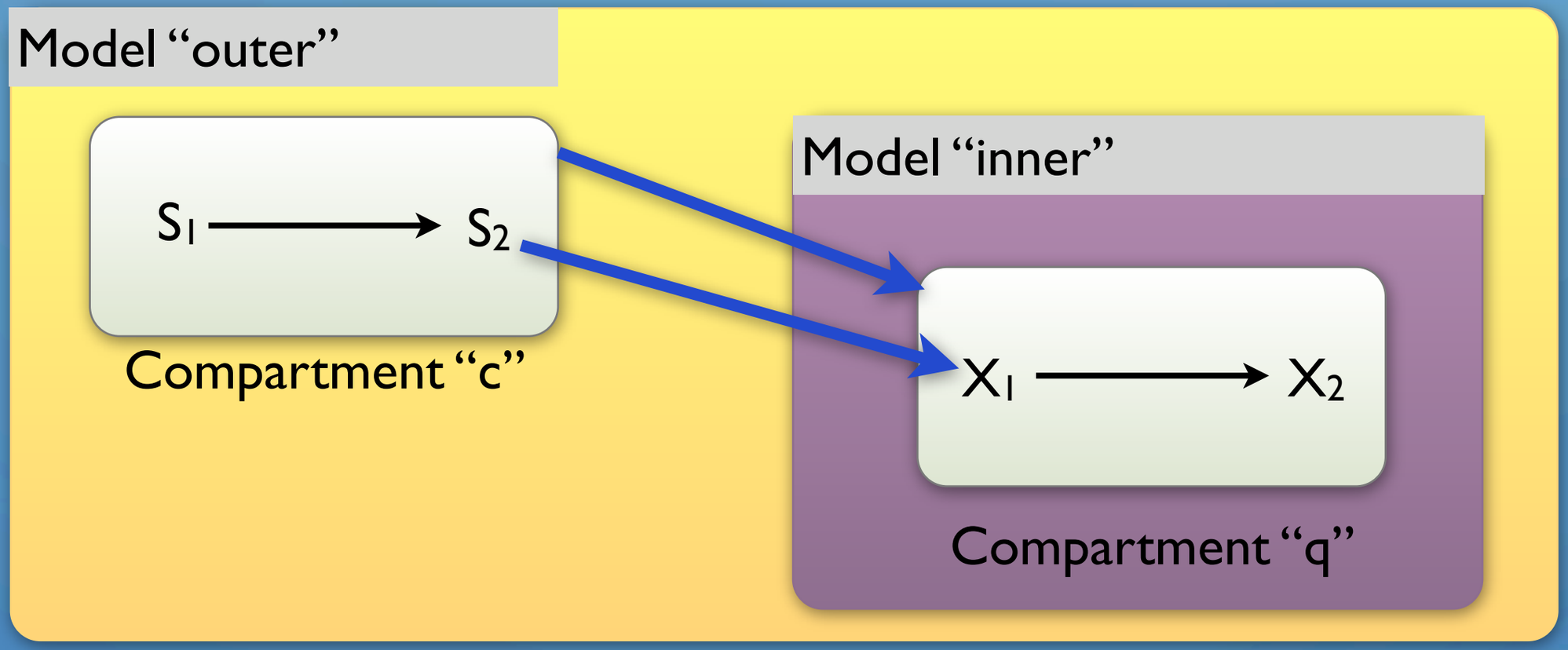
Species ...

Parameters ...

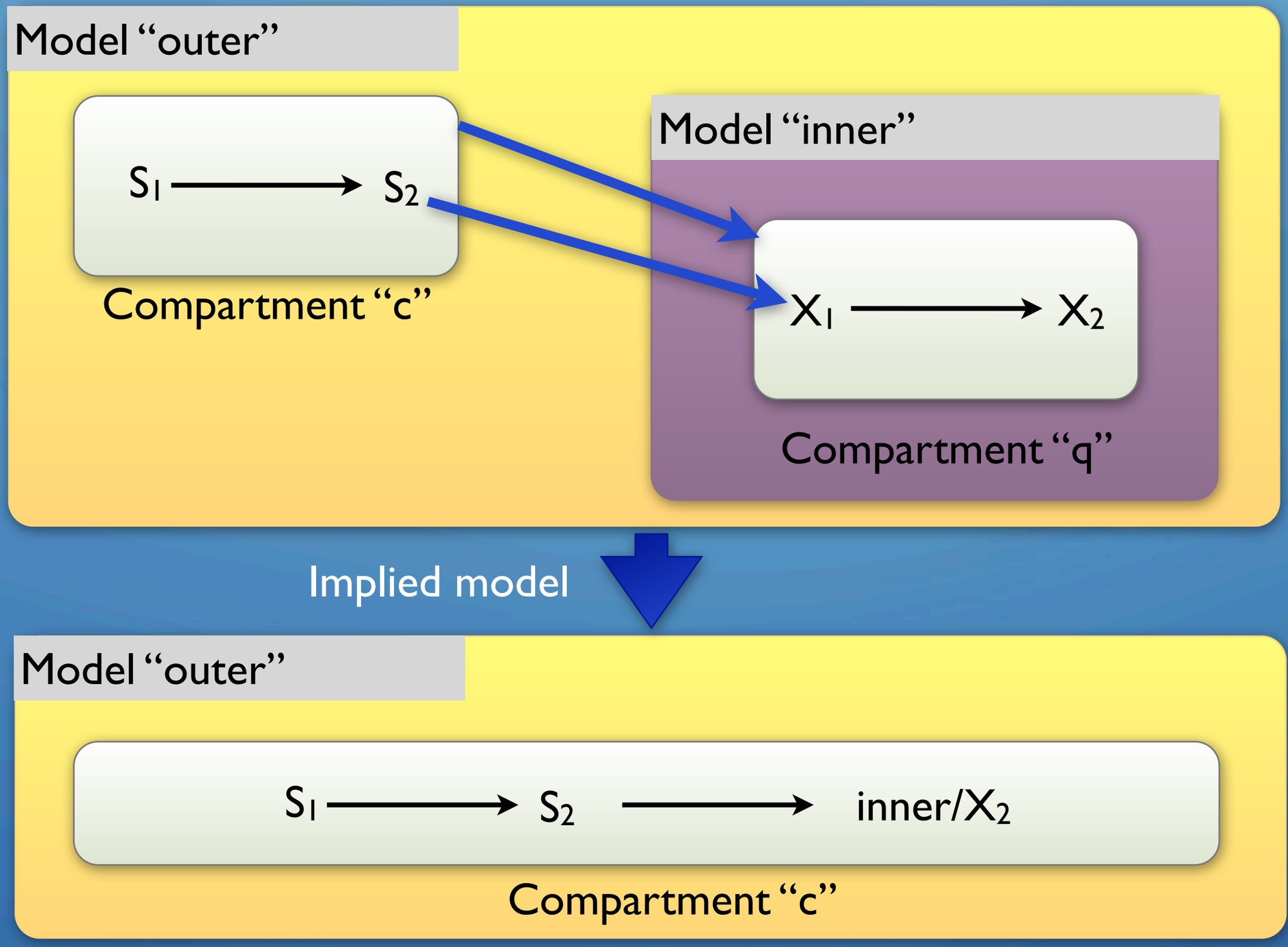
Reactions ...

Think: libraries of well-tested models

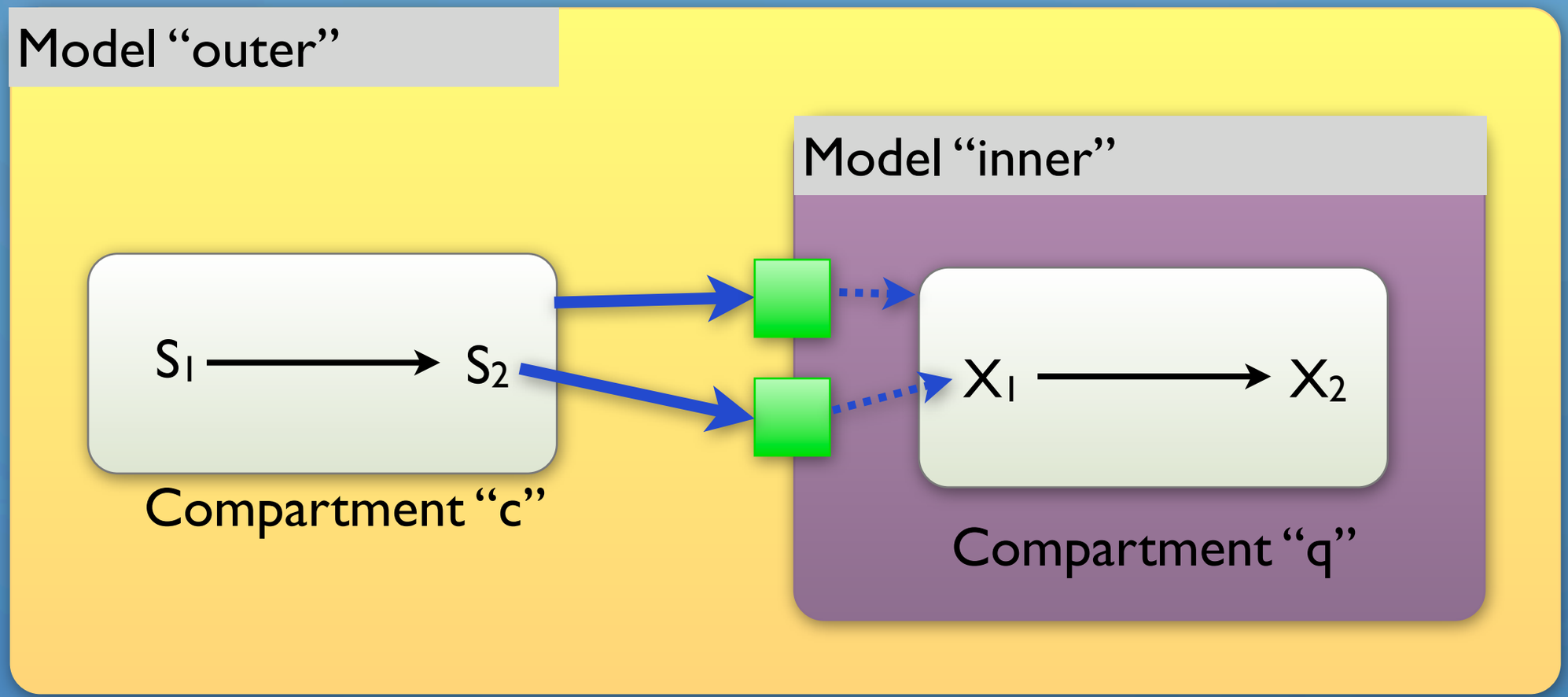
Links/references/replacements



Links/references/replacements



Interfaces/ports



- Defined interfaces act as "contracts"
  - Promote designed encapsulation & reuse of components

Gory bits

- Unit conversions
  - Need to specify how base units of one model are related to the base units of another model
  - For each entity link (species, compartment parameter), need to allow conversion factors
    - E.g.: X amount of glucose-6-phosphate in model M1 should be considered equal to Y amount in model M2

Acknowledgements: SBML

Sarah Keating

Frank Bergmann

Lucian Smith

SBML Team
Michael Hucka
Sarah Keating
Linda Taddeo
Frank Bergmann
Lucian Smith
Ben Bornstein
Andrew Finney
Akira Funahashi
Akiya Joukarou
Ben Kovitz
Bruce Shapiro
Herbert Sauro
Maria Schilstra
Jo Matthews
Hamid Bolouri
John Doyle
Hiroaki Kitano

SBML Editors
Michael Hucka
Frank Bergmann
Sven Sahle
James Schaff
Lucian Smith
Darren Wilkinson

SBML community

Dozens of contributors

Sven Sahle

Darren Wilkinson

Jim Schaff

Acknowledgements: BioModels Database

- EBI
  - Nicolas Le Novère
  - Alexander Broicher
  - Mélanie Courtot
  - Marco Donizelli
  - Arnaud Henry
  - Chen Li
  - Lu Li
  - Camille Laibe
  - Nicolas Rodriguez
- SBML team (Caltech)
  - Michael Hucka
  - Andrew Finney
  - Benjamin Borstein
  - Harish Dharuri
  - Enuo He
  - Sarah Keating
  - Maria Schilstra
  - Bruce Shapiro
- NCBS (Bangalore)
  - Upinder Bhalla
  - Harsha Rani
- University of Washington
  - Herbert Sauro
- Vienna TBI
  - Rainer Machne
- Systems Biology Institute (Tokyo)
  - Hiroaki Kitano
  - Akira Funahashi
- JWS Online (Stellenbosh)
  - Jacky Snoep
- Virtual Cell (UCHC)
  - Ion Moraru
- Journals supporting BioModels Database
  - Molecular Systems Biology
  - All PLoS Journals
  - All BioMedCentral Journals
- Programs used for curation
  - CellDesigner/SBMLodeSolver
  - COPASI
  - Jarnac/JDesigner
  - MathSBML
  - RoadRunner
  - SBMLeditor
  - XPP-Aut

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

(Slide from Nicolas Le Novère)

Acknowledgments: the funding

- **National Institute of General Medical Sciences (USA)**
- **European Molecular Biology Laboratory (EMBL)**
- **Beckman Institute, Caltech (USA)**
- Keio University (Japan)
- JST ERATO Kitano Symbiotic Systems Project (Japan) (to 2003)
- National Science Foundation (USA)
- International Joint Research Program of NEDO (Japan)
- JST ERATO-SORST Program (Japan)
- Japanese Ministry of Agriculture
- Japanese Ministry of Educ., Culture, Sports, Science and Tech.
- BBSRC (UK)
- DARPA IPTO Bio-SPICE Bio-Computation Program (USA)
- Air Force Office of Scientific Research (USA)
- STRI, University of Hertfordshire (UK)
- Molecular Sciences Institute (USA)

NIH/NIGMS

EMBL-EBI

Caltech

Thanks for listening!

- How to keep informed about SBML and learn more:
  - Join **sbml-announce@caltech.edu** (low-volume, broadcast-only)
  - See **http://sbml.org/Forums**
- SBML: http://sbml.org
- MIRIAM: http://biomodels.net/miriam
- MIASE: http://sourceforge.net/projects/miase
- SBO: http://biomodels.net/sbo
- BioModels Database: http://biomodels.net/biomodels

SBML's *Model*

model

- list of function definitions
- list of unit definitions
- list of compartment types (*Level 2 only*)
- list of species types (*Level 2 only*)
- list of compartments
- list of species
- list of parameters
- list of initial assignments
- list of rules
- list of constraints
- list of reactions
- list of events

SBML example

2 simple reactions:

$A \rightarrow B$

$B \rightarrow C$

Reaction r1 rate = $k_1[A]$

Reaction r2 rate = $k_2[B]$

1 compartment of volume 0.5 L

A, B, C in concentrations

A(0) = 3 M

B(0) = 0

C(0) = 0

$k_1 = 0.005$

$k_2 = 1 \times 10^{-5}$

Structure of SBML representation of the example

A → B
B → C

Reaction r1 rate = $k_1[A]$
Reaction r2 rate = $k_2[B]$
1 compartment of volume 0.5 L

A, B, C in concentrations
A(0) = 3 M, B(0) = 0, C(0) = 0
$k_1 = 0.005$, $k_2 = 1 \times 10^{-5}$

model

list of compartments

compartment id = "comp", initial volume = 0.5

list of species

species id = "A", initial concentration = 3

species id = "B", initial concentration = 0

species id = "C", initial concentration = 0

Structure of SBML representation of the example

A → B
B → C

Reaction r1 rate = $k_1[A]$
Reaction r2 rate = $k_2[B]$
1 compartment of volume 0.5 L

A, B, C in concentrations
A(0) = 3 M, B(0) = 0, C(0) = 0
$k_1 = 0.005$, $k_2 = 1 \times 10^{-5}$

list of reactions

reaction id = "r1"
list of reactants
species reference to "A"
list of products
species reference to "B"
kinetic law
math = " k1 • A • comp"
list of parameters
parameter id = "k1",
value = 0.005

reaction id = "r2"
list of reactants
species reference to "B"
list of products
species reference to "C"
kinetic law
math = " k2 • B • comp"
list of parameters
parameter id = "k2",
value = 1x10-5