EMBL-EBI

SBO SBML TEDDY

Controlled annotation of kinetics models

Nicolas Le Novère, EMBL-EBI

BioPAX MIRIAM SBGN

Storage and Annotation of Reaction Kinetics Data
May 21-23, 2007, EML research, Heidelberg, Germany

BioModels

;.... Curated Models
;.... Non-curated Models
;.... Search
;.... Simulate in JWS

Submit Your Model

;.... Curation tips
;.... Annotation tips

Sign-in

News
FAQ
Model of the month
Terms of Use
Related Software
Meetings
Contact
Quote

BIOMODELS.NET

Computational Neurobiology

SBML

BioModels Database

A Database of Annotated Published Models

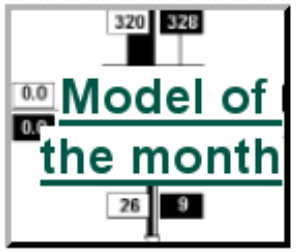
BioModels Database is a data resource that allows biologists to store, search and retrieve published mathematical models of biological interests. Models present in BioModels Database are annotated and linked to relevant data resources, such as publications, databases of compounds and pathways, controlled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

[Browse curated models] | [Browse non-curated models] | [Search] | [Simulate in JWS Online]

05th June 2007 - Eighth Release! [More] [Download All Models Under SBML L2 V1 Format]

17th April 2007 - BioModels Database ranked first data resource for Systems Biology [More] [Original article]

4th December 2006 - BioModels Database and DOQCS join forces [More]



Acknowledgements

BioModels Database is developed in collaboration by the teams of Nicolas Le Novère (EMBL-EBI, United-Kingdom), Michael Hucka (SBML Team, Caltech, USA) in collaboration with Upinder Bhalla (DOQCS, National Center for Biological Sciences, India), Herbert Sauro (Keck Graduate Institute, USA), Hiroaki Kitano (Systems Biology Institute, Japan), Hans Westerhoff and Jacky Snoep (JWS Online, Stellenbosch (ZA) and Manchester (UK) Universities and Stellenbosh University, ZA), as part of the BioModels.net initiative. BioModels Database development has benefitted from funds of the European Molecular Biology Laboratory (Le Novère team) and the National Institute of General Medical Sciences (SBML team).

Developers: Mélanie Courtot, Arnaud Henry, Camille Laibe, Chen Li (main developer), Lu Li, Nicolas Rodriguez (Alumni: Marco Donizelli)

Model curators and annotators: Harish Dharuri, Enuo He, Nicolas Le Novère, Lu Li, Rainer Machne, Bruce Shapiro (Alumni: Maria Schilstra)

External contributors: BioModels Database would not exists without the continuous support of many peopler, whether by ther contribution of models, of software, or by their constructive comments and criticisms. It is unfortunately impossible keep track and to to acknowledge all of them here, without risking to be unfair. Therefore, a big collective thank-you all.

BioModels Database is built thanks to many third-party free software such as libSBML, Jakarta Tomcat, Xindice, Xalan, Xerces, Jena and MySQL.

Main software used to curate models: CellDesigner, COPASI, Jarnac, MathSBML, XPP-Aut, SBMLodeSolver, SBMLeditor and SBMLmerge.

BioModels

;.... Curated Models
;.... Non-curated Models
;.... Search
;.... Simulate in JWS

Submit Your Model

;.... Curation tips
;.... Annotation tips

Sign-in

News
FAQ
Model of the month
Terms of Use
Related Software
Meetings
Contact
Quote

BIOMODELS.NET

BioModels Database

A Database of Annotated Published Models

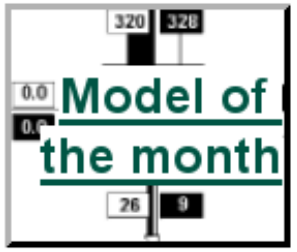
BioModels Database is a data resource that allows biologists to store, search and retrieve published mathematical models of biological interests. Models present in BioModels Database are annotated and linked to relevant data resources, such as publications, databases of compounds and pathways, controlled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

[Browse curated models] | [Browse non-curated models] | [Search] | [Simulate in JWS Online]

05th June 2007 - Eighth Release! [More] [Download All Models Under SBML L2 V1 Format]

17th April 2007 - BioModels Database ranked first data resource for Systems Biology [More] [Original article]

4th December 2006 - BioModels Database and DOQCS join forces [More]



Acknowledgements

BioModels Database is developed in collaboration by the teams of Nicolas Le Novère (EMBL-EBI, United-Kingdom), Michael Hucka (SBML Team, Caltech, USA) in collaboration with Upinder Bhalla (DOQCS, National Center for Biological Sciences, India), Herbert Sauro (Keck Graduate Institute, USA), Hiroaki Kitano (Systems Biology Institute, Japan), Hans Westerhoff and Jacky Snoep (JWS Online, Stellenbosch (ZA) and Manchester (UK) Universities and Stellenbosh University, ZA), as part of the BioModels.net initiative. BioModels Database development has benefitted from funds of the European Molecular Biology Laboratory (Le Novère team) and the National Institute of General Medical Sciences (SBML team).

Developers: Mélanie Courtot, Arnaud Henry, Camille Laibe, Chen Li (main developer), Lu Li, Nicolas Rodriguez (Alumni: Marco Donizelli)

Model curators and annotators: Harish Dharuri, Enuo He, Nicolas Le Novère, Lu Li, Rainer Machne, Bruce Shapiro (Alumni: Maria Schilstra)

External contributors: BioModels Database would not exists without the continuous support of many peopler, whether by ther contribution of models, of software, or by their constructive comments and criticisms. It is unfortunately impossible keep track and to to acknowledge all of them here, without risking to be unfair. Therefore, a big collective thank-you all.

BioModels Database is built thanks to many third-party free software such as libSBML, Jakarta Tomcat, Xindice, Xalan, Xerces, Jena and MySQL.

Main software used to curate models: CellDesigner, COPASI, Jarnac, MathSBML, XPP-Aut, SBMLodeSolver, SBMLeditor and SBMLmerge.

Kinetics models are biological data

nature

6 May 2005 Volume 435 Issue no 7038

In pursuit of systems

The study of functioning groups of molecules is an important frontier of biology at reductionist and holistic levels. Central to the long-term goals of scientific research, it brings its own challenges of infrastructure and evaluation.

What is the difference between a live cat and a dead one? One scientific answer is 'systems biology'. A dead cat is a collection of its component parts. A live cat is the emergent behaviour of the system incorporating those parts. There is certainly a vast distance to go before we can fully encompass such a system within scientific description. So how is systems biology already moving us towards the fullest possible description of a live cat?

By focusing on the behaviour of individual proteins and other biomolecules, much of what gives life its unique properties can be missed. To a systems biologist, the network of interactions formed by these components is more important than the molecules themselves. Properties such as robustness and evolvability, essential characteristics of life, then emerge from the topology of biological networks, independent of the constituents from which they are built.

Such a holistic view may sound dangerously soft-edged. Far from it. Systems biology couples the acquisition of comprehensive, high-definition data sets to the construction of quantitative models and computer simulations. Indeed, it is an explicit aim of both the Kyoto Encyclopedia of Genes and Genomes and the Alliance for Cellular Signaling to construct a fully functioning computer model of a cell.

The present state of experimental systems biology is both tantalizing and frustrating. To provide the level of detail required for us to know what is going on in a cell, microarray technologies will need to be faster and require smaller samples; ways to label and follow more biological molecules within a cell must be discovered; new spectroscopic tools to non-invasively measure multiple metabolite levels will need to be developed, and so on. *Nature* is committed to publishing studies that push back the technological frontier of what it is possible to know about important biological systems.

But technical wizardry and large data sets are only part of the systems-biology approach — a system is not fully understood until a quantitative model can be built. The role of modelling in biological research is controversial and can spark heated debates. What is clear, though, is that the wealth of experimental data emerging from systems biology would be uninterpretable without detailed models against which they can be compared. Advances in modelling and simulation are thus no less important than data collection.

Every discipline generates community infrastructures, and systems biology is no exception. In the past five years, systems-biology institutes, departments and initiatives have been springing up across the globe. New journals have been launched, including The Institution of Electrical Engineers' *Systems Biology* and Nature Publishing Group's *Molecular Systems Biology*. The latter, an author-pays, online-only journal, is a joint venture with the European Molecular Biology Organization and went live last month.

The exchange of models between researchers is imperative, so a welcome development last month was the launch of BioModels (www.ebi.ac.uk/biomodels), a curated database for the deposition of biological models. BioModels has built on the success of Systems Biology Markup Language (SBML) in providing a format for the presentation of models, allowing them to be implemented on different software platforms. *Nature* journals and *Molecular Systems Biology* support submissions involving SBML.

It is hoped that BioModels will form the basis of a universally accepted repository that can do for systems biology what GenBank and the Protein Data Bank have done for genetics and structural biology. *Nature* applauds such efforts and will encourage authors of papers containing suitable models to contribute them to BioModels.

Systems biology presents an intellectual challenge to scientists and journal editors alike. Papers in this field document a highly multidisciplinary endeavour. Reviewers of such papers are very good at dissecting the aspects that fall within their sphere of expertise, but are less insightful beyond. So it falls to editors to weigh their frequently conflicting opinions in taking balanced and clear-sighted decisions. As a multidisciplinary journal, *Nature* welcomes the particular challenges that systems biology presents. ■

Kinetics models are biological data

nature

6 May 2005 Volume 435 Issue no 7038

In pursuit of systems

The study of functioning groups of molecules is an important frontier of biology at reductionist and holistic levels. Central to the long-term goals of scientific research, it brings its own challenges of infrastructure and evaluation.

What is the difference between a live cat and a dead one? One scientific answer is 'systems biology'. A dead cat is a collection of its component parts. A live cat is the emergent behaviour of the system incorporating those parts. There is certainly a vast distance to go before we can fully encompass such a system within scientific description. So how is systems biology already moving us towards the fullest possible description of a live cat?

By focusing on the behaviour of individual proteins and other biomolecules, much of what gives life its unique properties can be missed. To a systems biologist, the network of interactions formed by these components is more important than the molecules themselves. Properties such as robustness and evolvability, essential characteristics of life, then emerge from the topology of biological networks, independent of the constituents from which they are built.

Such a holistic view may sound dangerously soft-edged. Far from it. Systems biology couples the acquisition of comprehensive, high-definition data sets to the construction of quantitative models and computer simulations. Indeed, it is an explicit aim of both the Kyoto Encyclopedia of Genes and Genomes and the Alliance for Cellular Signaling to construct a fully functioning computer model of a cell.

The present state of experimental systems biology is both tantalizing and frustrating. To provide the level of detail required for us to know what is going on in a cell, microarray technologies will need to be faster and require smaller samples; ways to label and follow more biological molecules within a cell must be discovered; new spectroscopic tools to non-invasively measure multiple metabolite levels will need to be developed, and so on. *Nature* is committed to publishing studies that push back the technological frontier of what it is possible to know about important biological systems.

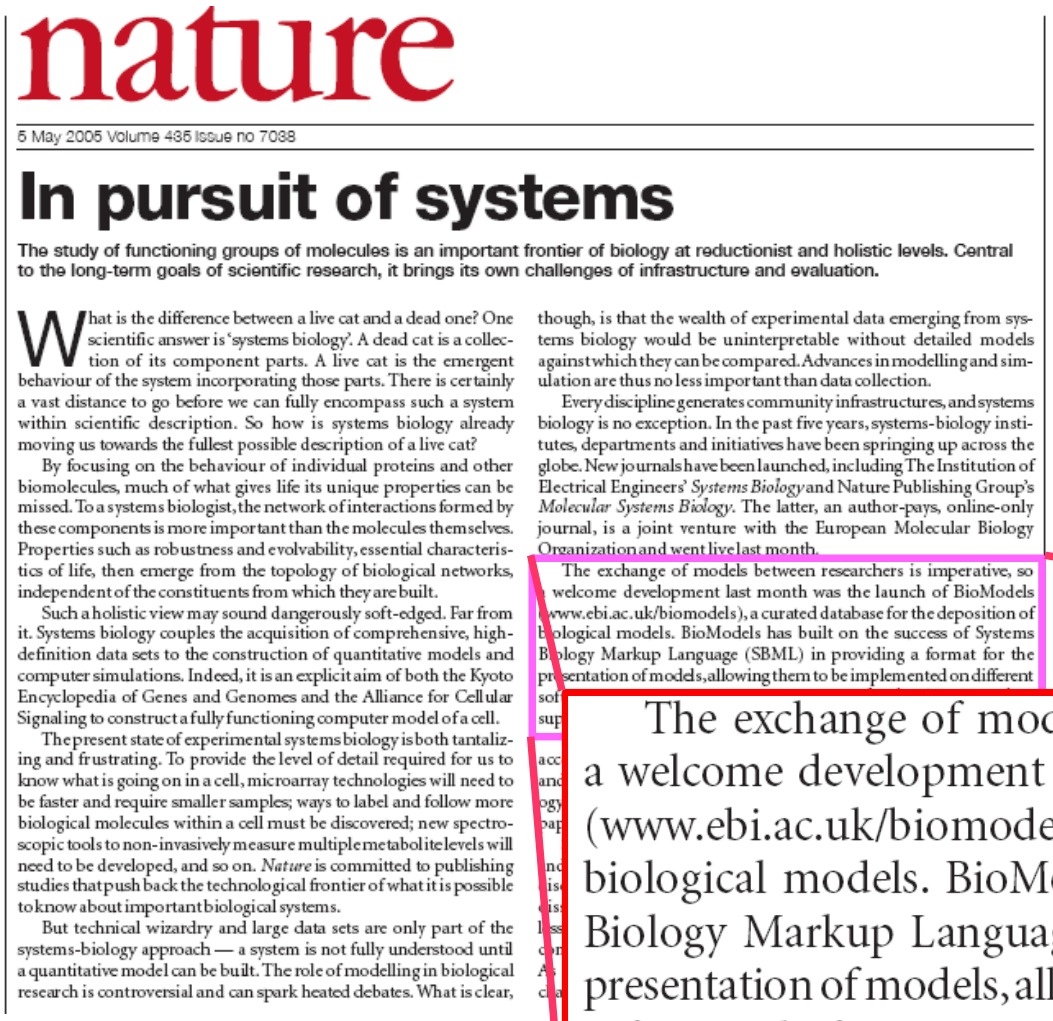
But technical wizardry and large data sets are only part of the systems-biology approach — a system is not fully understood until a quantitative model can be built. The role of modelling in biological research is controversial and can spark heated debates. What is clear, though, is that the wealth of experimental data emerging from systems biology would be uninterpretable without detailed models against which they can be compared. Advances in modelling and simulation are thus no less important than data collection.

Every discipline generates community infrastructures, and systems biology is no exception. In the past five years, systems-biology institutes, departments and initiatives have been springing up across the globe. New journals have been launched, including The Institution of Electrical Engineers' *Systems Biology* and Nature Publishing Group's *Molecular Systems Biology*. The latter, an author-pays, online-only journal, is a joint venture with the European Molecular Biology Organization and went live last month.

The exchange of models between researchers is imperative, so a welcome development last month was the launch of BioModels (www.ebi.ac.uk/biomodels), a curated database for the deposition of biological models. BioModels has built on the success of Systems Biology Markup Language (SBML) in providing a format for the presentation of models, allowing them to be implemented on different software platforms. *Nature* journals and *Molecular Systems Biology* support submissions involving SBML.

The exchange of models between researchers is imperative, so a welcome development last month was the launch of BioModels (www.ebi.ac.uk/biomodels), a curated database for the deposition of biological models. BioModels has built on the success of Systems Biology Markup Language (SBML) in providing a format for the presentation of models, allowing them to be implemented on different software platforms. *Nature* journals and *Molecular Systems Biology* support submissions involving SBML.

Kinetics models are biological data



nature

6 May 2005 Volume 435 Issue no 7038

In pursuit of systems

The study of functioning groups of molecules is an important frontier of biology at reductionist and holistic levels. Central to the long-term goals of scientific research, it brings its own challenges of infrastructure and evaluation.

What is the difference between a live cat and a dead one? One scientific answer is 'systems biology'. A dead cat is a collection of its component parts. A live cat is the emergent behaviour of the system incorporating those parts. There is certainly a vast distance to go before we can fully encompass such a system within scientific description. So how is systems biology already moving us towards the fullest possible description of a live cat?

By focusing on the behaviour of individual proteins and other biomolecules, much of what gives life its unique properties can be missed. To a systems biologist, the network of interactions formed by these components is more important than the molecules themselves. Properties such as robustness and evolvability, essential characteristics of life, then emerge from the topology of biological networks, independent of the constituents from which they are built.

Such a holistic view may sound dangerously soft-edged. Far from it. Systems biology couples the acquisition of comprehensive, high-definition data sets to the construction of quantitative models and computer simulations. Indeed, it is an explicit aim of both the Kyoto Encyclopedia of Genes and Genomes and the Alliance for Cellular Signaling to construct a fully functioning computer model of a cell.

The present state of experimental systems biology is both tantalizing and frustrating. To provide the level of detail required for us to know what is going on in a cell, microarray technologies will need to be faster and require smaller samples; ways to label and follow more biological molecules within a cell must be discovered; new spectroscopic tools to non-invasively measure multiple metabolite levels will need to be developed, and so on. *Nature* is committed to publishing studies that push back the technological frontier of what it is possible to know about important biological systems.

But technical wizardry and large data sets are only part of the systems-biology approach — a system is not fully understood until a quantitative model can be built. The role of modelling in biological research is controversial and can spark heated debates. What is clear, though, is that the wealth of experimental data emerging from systems biology would be uninterpretable without detailed models against which they can be compared. Advances in modelling and simulation are thus no less important than data collection.

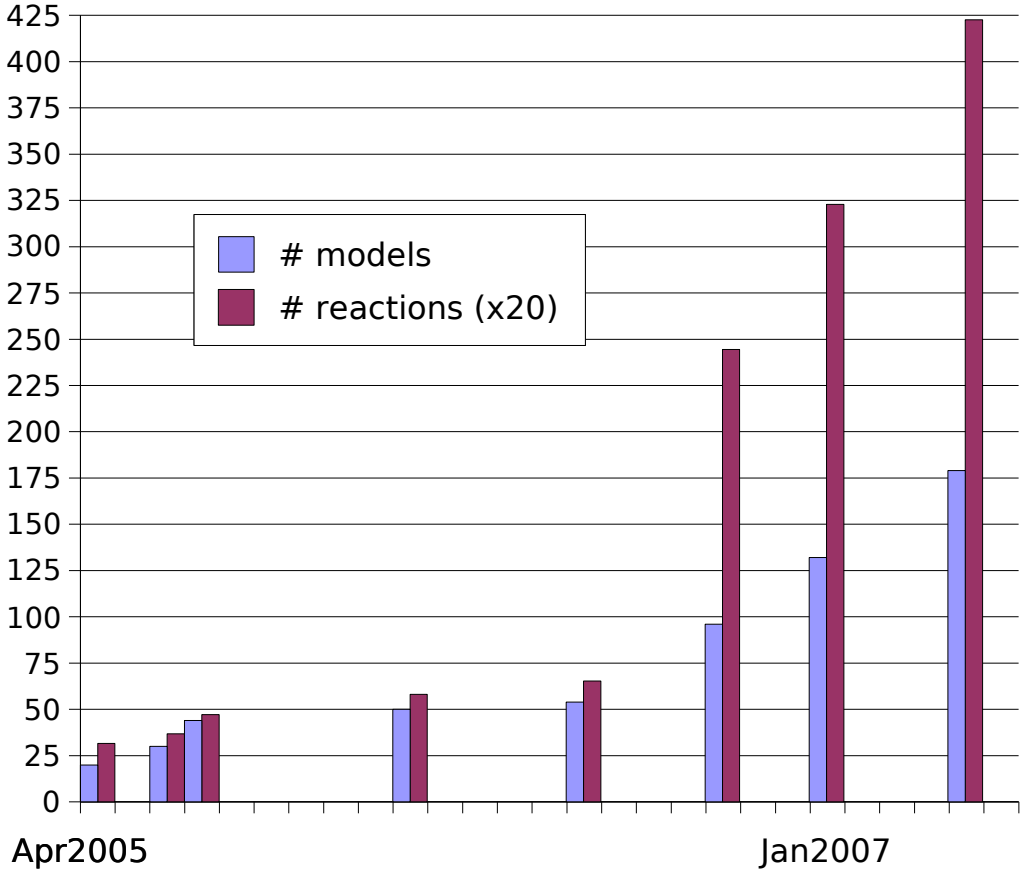
Every discipline generates community infrastructures, and systems biology is no exception. In the past five years, systems-biology institutes, departments and initiatives have been springing up across the globe. New journals have been launched, including The Institution of Electrical Engineers' *Systems Biology* and Nature Publishing Group's *Molecular Systems Biology*. The latter, an author-pays, online-only journal, is a joint venture with the European Molecular Biology Organization and went live last month.

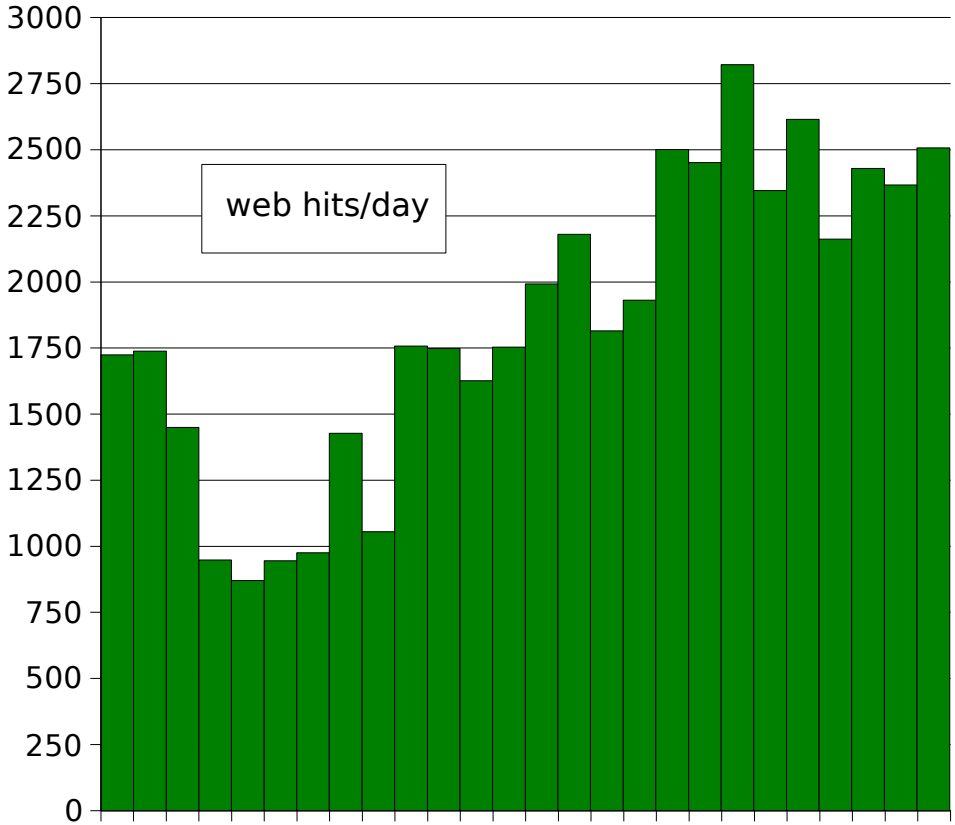
The exchange of models between researchers is imperative, so a welcome development last month was the launch of BioModels (www.ebi.ac.uk/biomodels), a curated database for the deposition of biological models. BioModels has built on the success of Systems Biology Markup Language (SBML) in providing a format for the presentation of models, allowing them to be implemented on different software platforms. *Nature* journals and *Molecular Systems Biology* support submissions involving SBML.

- Deposition advised by
  - Molecular Systems Biology
  - all PLoS journals
  - all BMC Journals

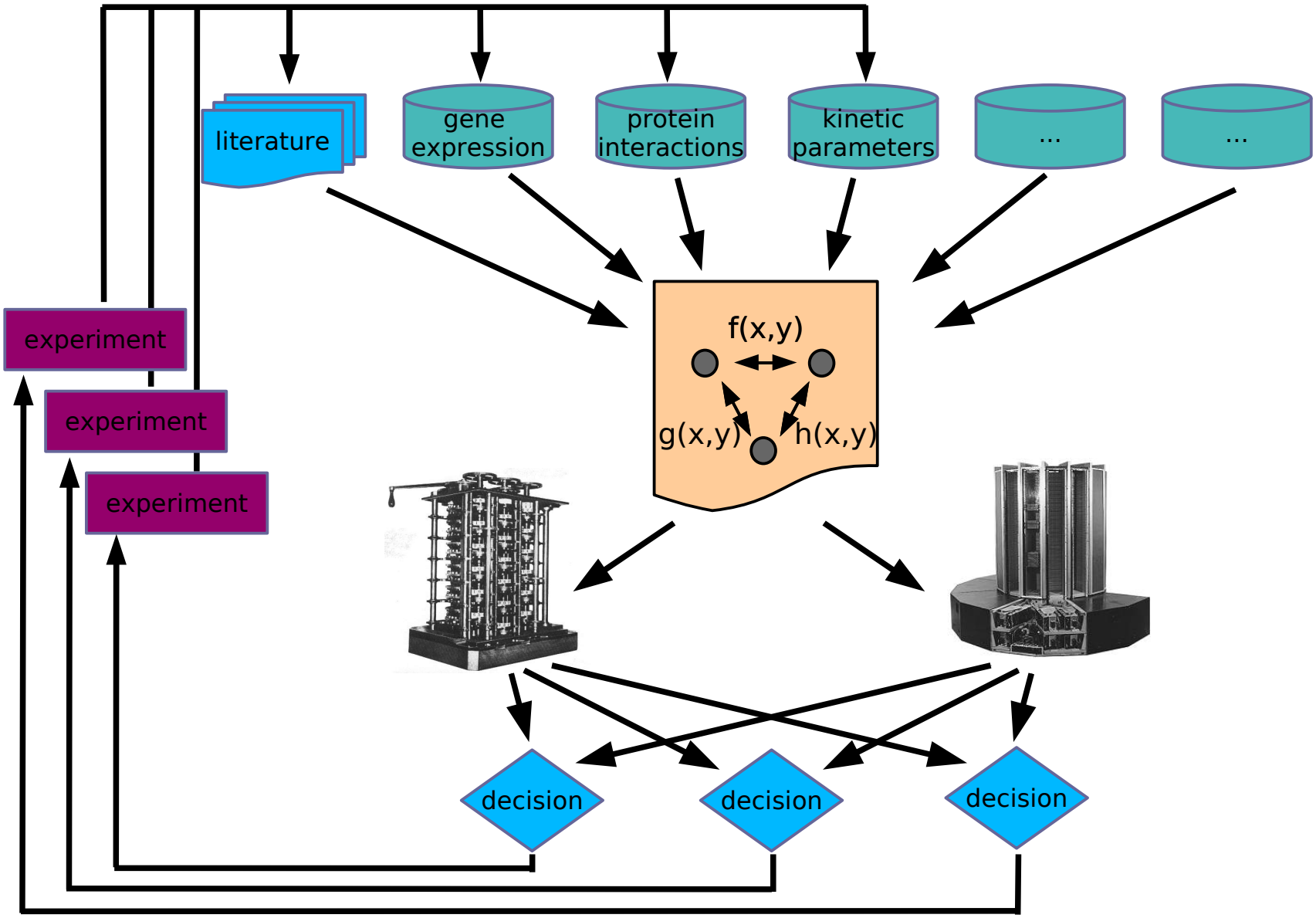
The exchange of models between researchers is imperative, so a welcome development last month was the launch of BioModels (www.ebi.ac.uk/biomodels), a curated database for the deposition of biological models. BioModels has built on the success of Systems Biology Markup Language (SBML) in providing a format for the presentation of models, allowing them to be implemented on different software platforms. *Nature* journals and *Molecular Systems Biology* support submissions involving SBML.

Steady-increase of BioModels DB

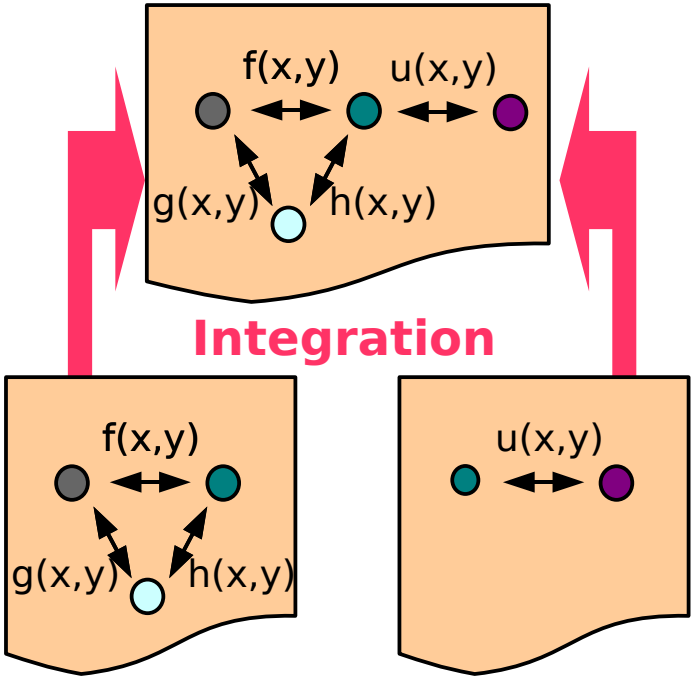




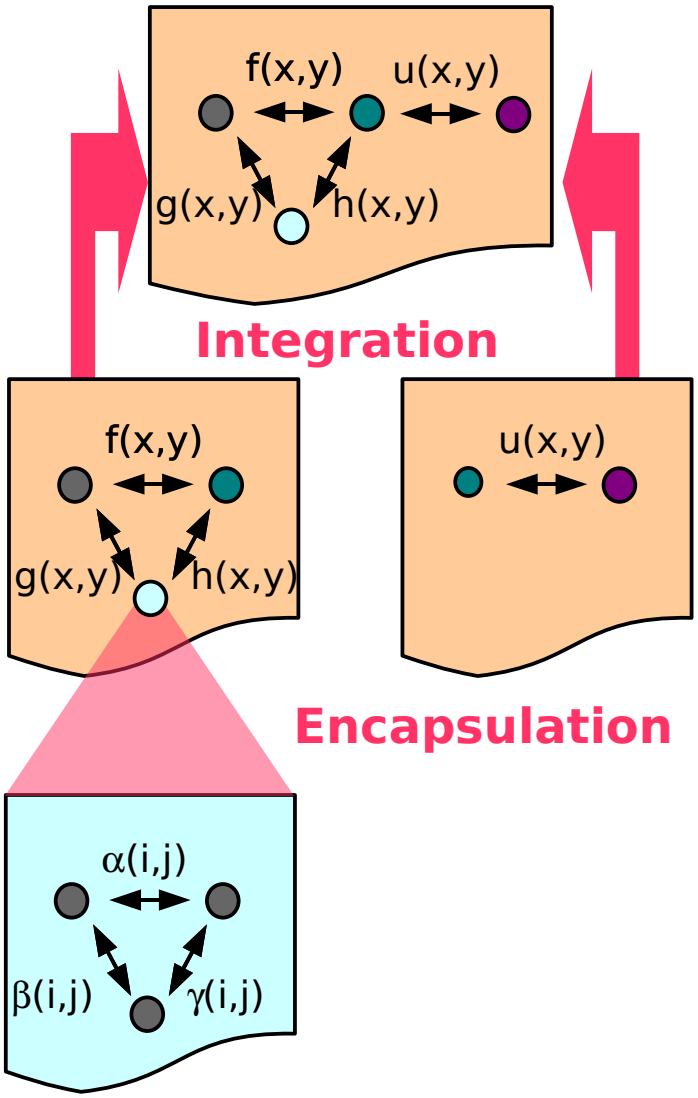
The model as an integrator of knowledge



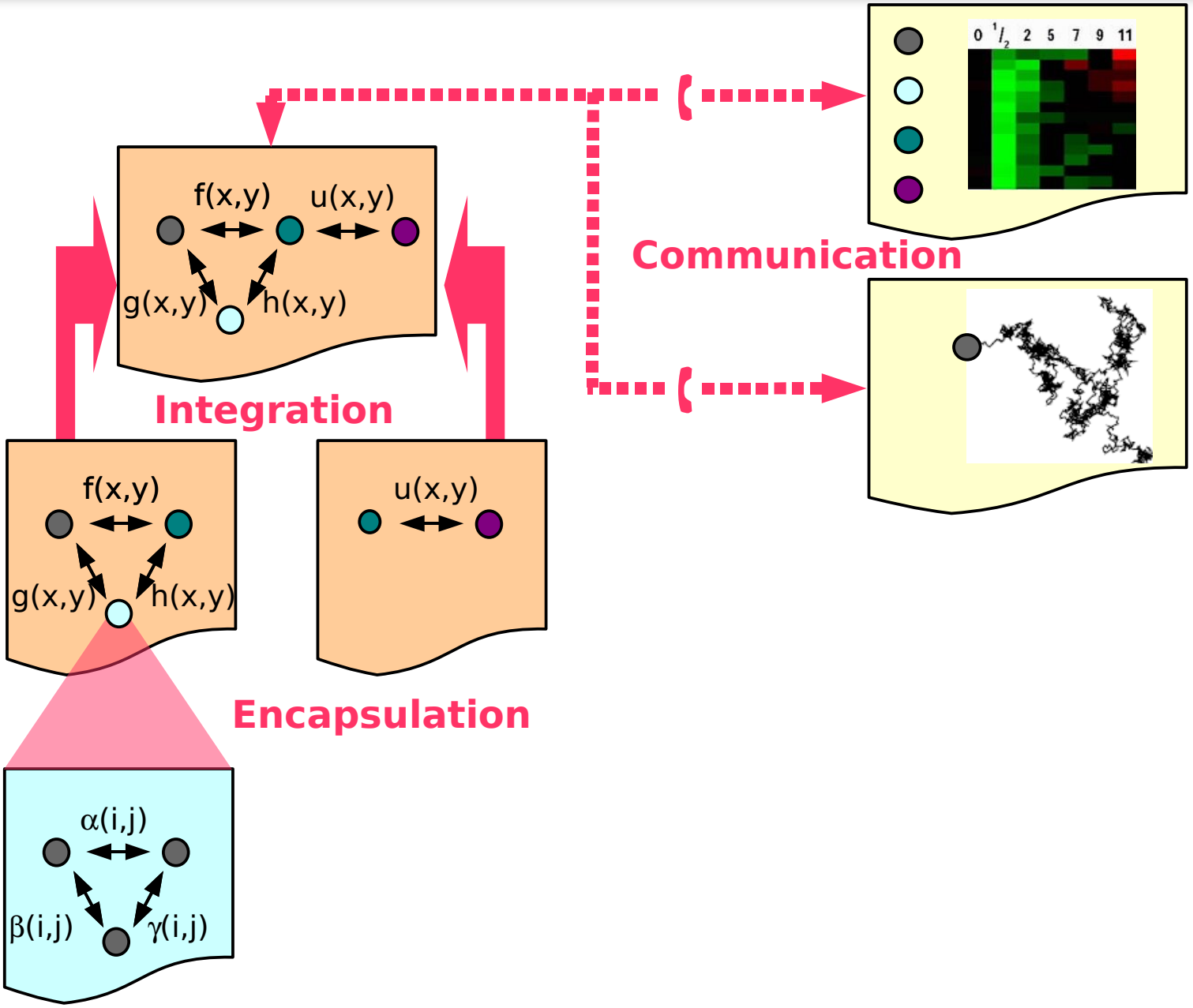
Technical bottlenecks



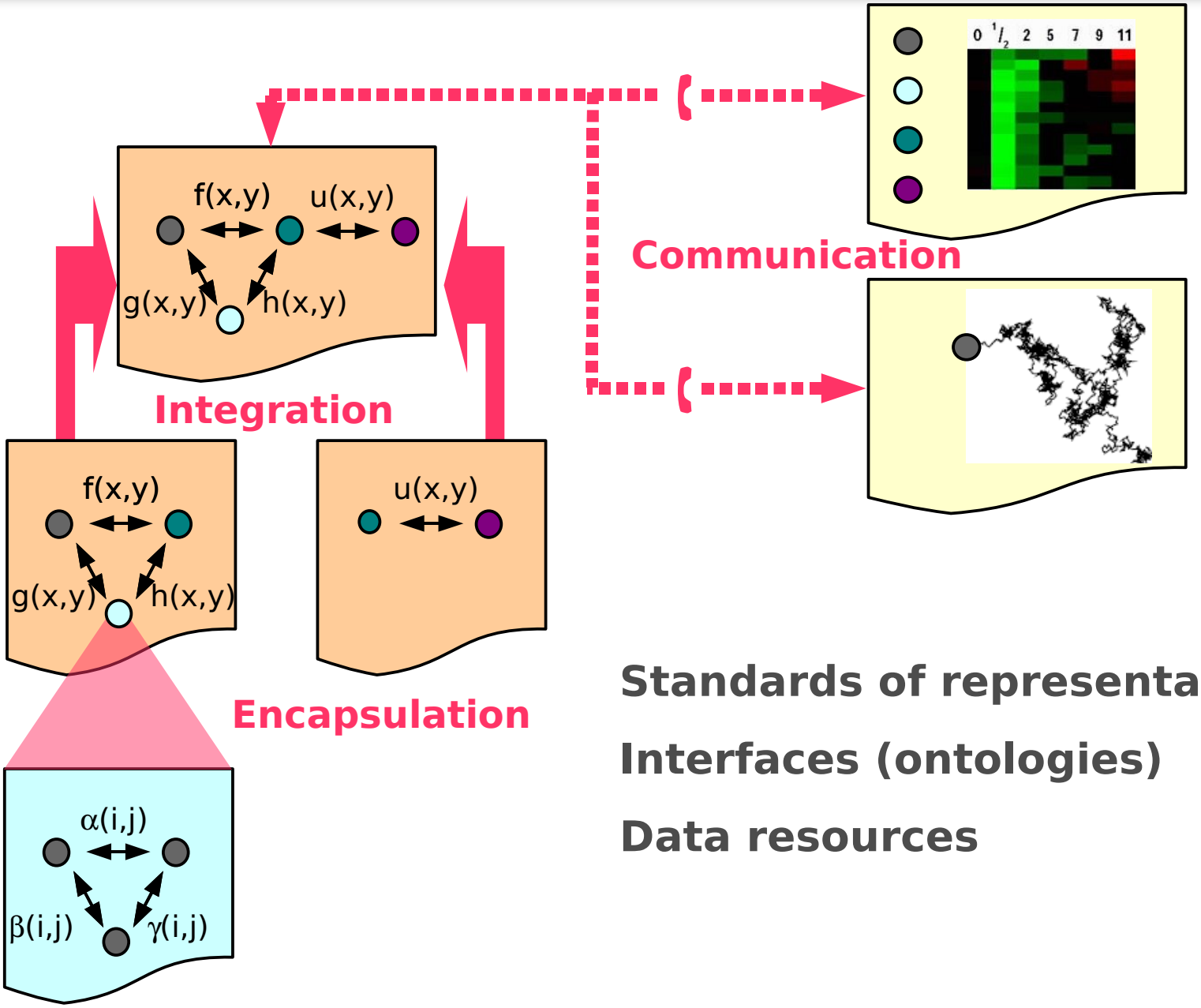
Technical bottlenecks



Technical bottlenecks



Technical bottlenecks



Standards of representation

Interfaces (ontologies)

Data resources

SBML Systems Biology Markup Language

Google Search this site

home • contacts • documents • downloads • FAQs • forums • Level 3 • models • news • online tools • wiki • workshops

The Systems Biology Markup Language (SBML) is a computer-readable format for representing **models of biochemical reaction networks**. SBML is applicable to metabolic networks, cell-signaling pathways, regulatory networks, and many others.

Internationally Supported and Widely Used

SBML has been evolving since mid-2000 through the efforts of an international group of software developers and users. Today, SBML is **supported by over 100 software systems**, including the following (where '*' indicates SBML support in development):

BALSA
BASIS
BIOCHAM
BioCharon
ByoDyn
BioCyc
BioGrid
BioModels
BioNetGen
BioPathwise
Bio Sketch Pad
BioSens
BioSPICE Dashboard
BioSpreadsheet
BioTapestry
BioUML
BSTLab
CADLIVE
CellDesigner
Cellerator
CellML2SBML
Cellware
CL-SBML
CLEML
COPASI
Cyto-Sim
Cytoscape
DBsolve
Dizzy
E-CELL
ecellJ
ESS
FluxAnalyzer
Fluxor
Gepasi
Gillespie2
HSMB
HybridSBML
INSILICO discovery
JACOBIAN
Jarnac
JDesigner
JigCell
JSim
JWS Online
Karyote*
KEGG2SBML
Kineticon
Kinsolver*
libSBML
MathSBML
MesoRD
MetaboLogica
MetaFluxNet
MMT2
Modesto
Moleculizer
Monod
Narrator
NetBuilder
Oscill8
PANTHER Pathway
PathArt
PathScout
Pathway Analyser
PathwayLab
Pathway Tools
PathwayBuilder
PATIKAweb
PaVESy
PET
PNK
PottersWheel
Reactome
ProcessDB
PROTON
pysbml
PySCeS
runSBML
SABIO-RK
SBML ODE Solver
SBML-PET
SBMLeditor
SBMLmerge
SBMLR
SBMLSim
SBMLToolbox
SBliD
SBToolbox
SBW
SCIpath
Sigmoid*
SigPath
SigTran
SIMBA
SimBiology
Simpathica
SimPheny*
SimWiz
SloppyCell
SmartCell
SRS Pathway Editor
StochSim
StochKit
STOCKS
TERANODE Suite
Trelis
Virtual Cell
WebCell
WinSCAMP
XPPAUT

A Free and Open Language

Advances in biotechnology are leading to larger, more complex quantitative models. The systems biology

BioNetGen@VCell Release

(October 6, 2006) **BioNetGen@VCell** is a new release of BioNetGen, a tool for automatically generating a reaction network from user-specified rules for biomolecular interactions on the level of protein domains.

read more

PottersWheel supports SBML

(October 4, 2006) **PottersWheel** 1.2 beta, a MATLAB systems biology toolbox, supports model creation, fitting data, and designing new experiments.

read more

SBML Level 2 Version 2 Released!

(September 25, 2006) The final version of the **SBML Level 2 Version 2** specification is now available!

read more

SBML Wikipedia entry

(September 18, 2006) There is now an updated entry for SBML in Wikipedia. Let us know your suggestions for improvements.

read more

SBML Tutorial at ICSB 2006

(September 8, 2006) Mike Hucka will be leading a tutorial on SBML this year at ICSB 2006 in Japan. The focus will be on the about-to-be-released SBML Level 2 Version 2.

read more

See older news items.

New: SBML Level 2 Version 3

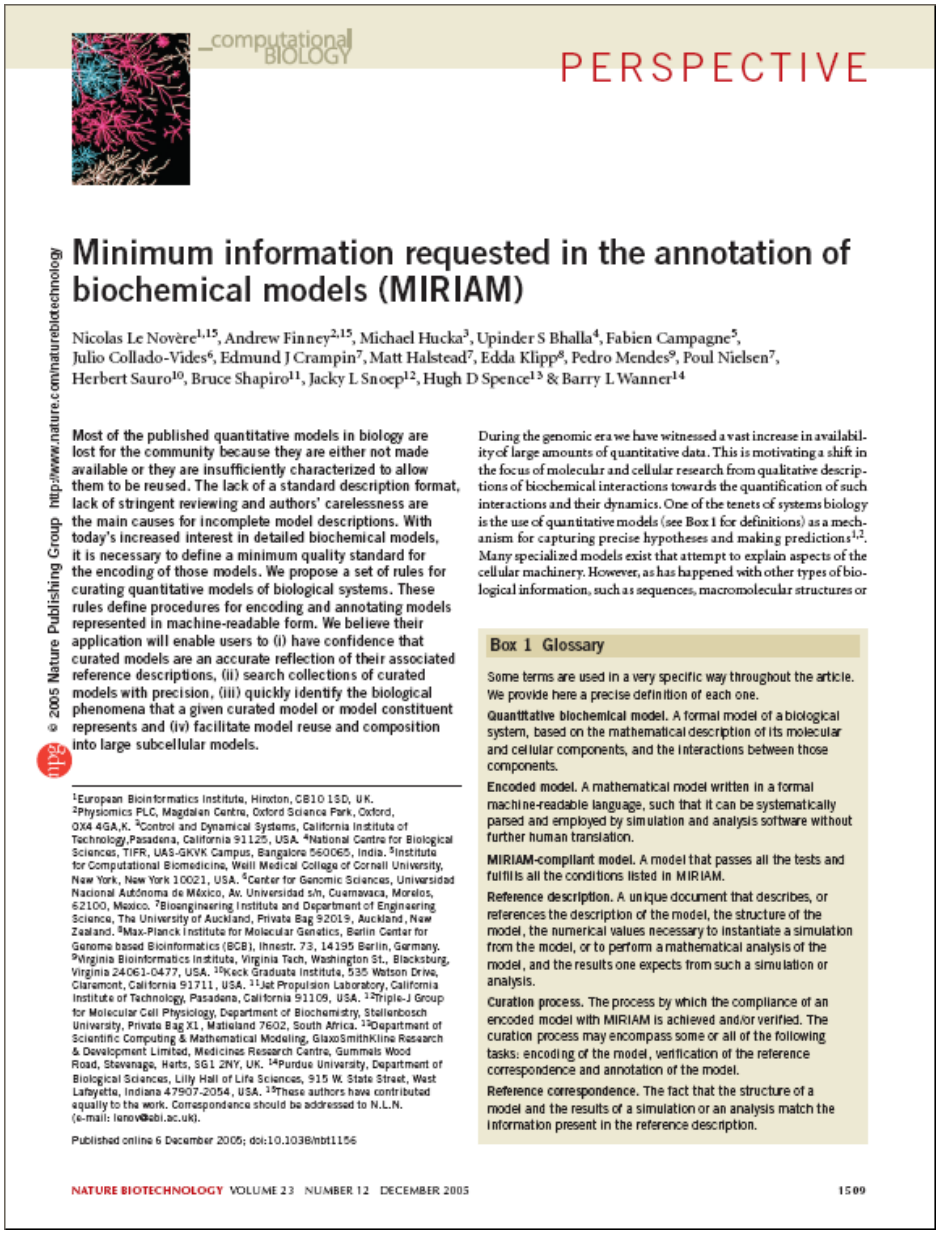
- Pre-released on March 20th 2007
- Simpler and cleaner (units ...)
- Generic entities (compartmentType, speciesType)
  ➔ path to generalised reactions
- Constraints and initialAssignments
- Controlled (MIRIAM) annotations (+ links to SBO)
- Backward compatible with Level 2 Version 1
- More detailed and bug-free specification ... 158 pages, 10pt, small margin.

Is SBML enough? What's missing?

- An SBML model lists participants, but does not identify them.
- An SBML model contains mathematical expressions, but does not tell-us what they “mean”, and how they are derived.
- An SBML model constructed for a certain modelling approach cannot be used straight-away within another modelling framework.

⇒ **SBML models cannot be easily searched**
SBML models cannot be easily converted
SBML models cannot be easily merged

- Proposed guidelines for curation of quantitative models
  - Specifically about encoding & annotation
  - Limited to models that can be simulated
- Effort arose from a meeting organized by Andrew Finney during ICSB 2004
- Not specific to SBML; applicable to any structured model format



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, Claramont, 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, Gunnels 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.

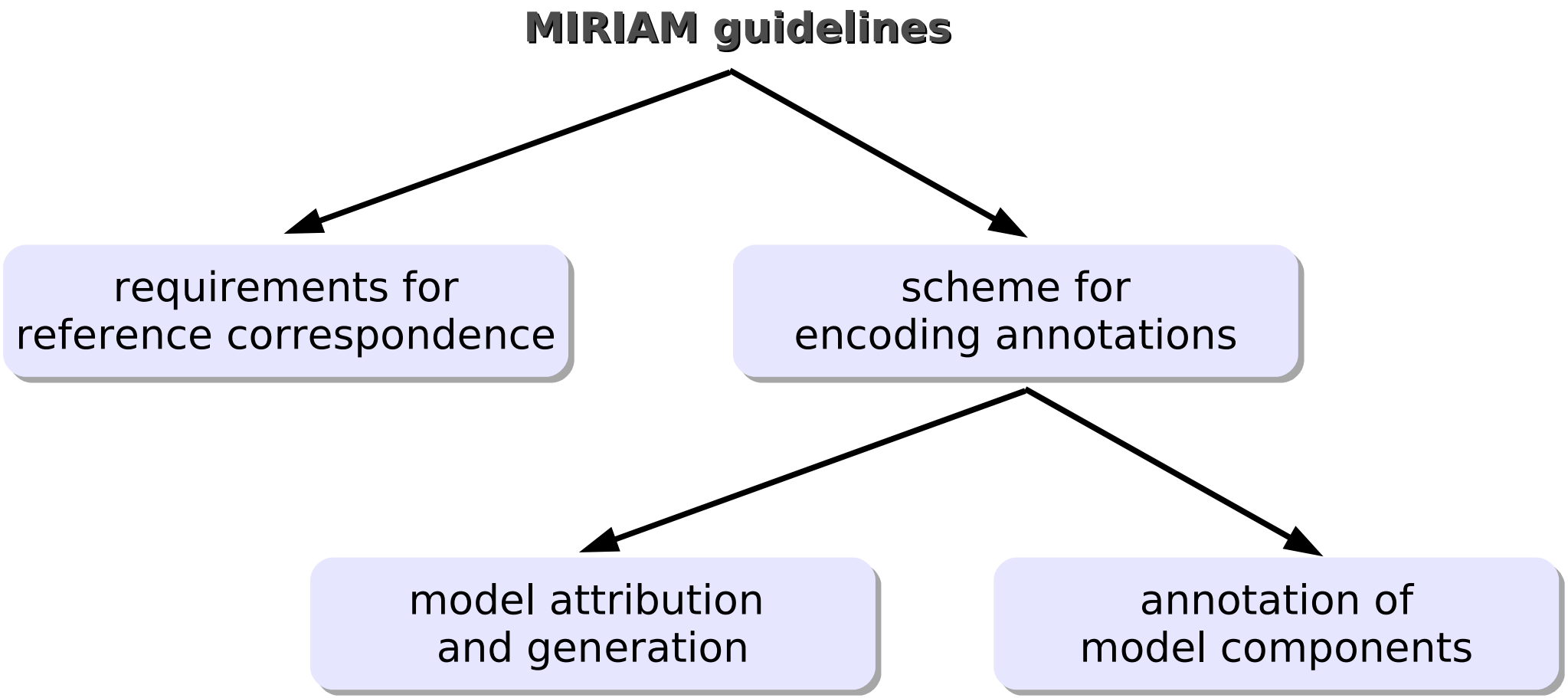
Encoded model. A mathematical model written in a formal machine-readable language, such that it can be systematically parsed and employed by simulation and analysis software without further human translation.

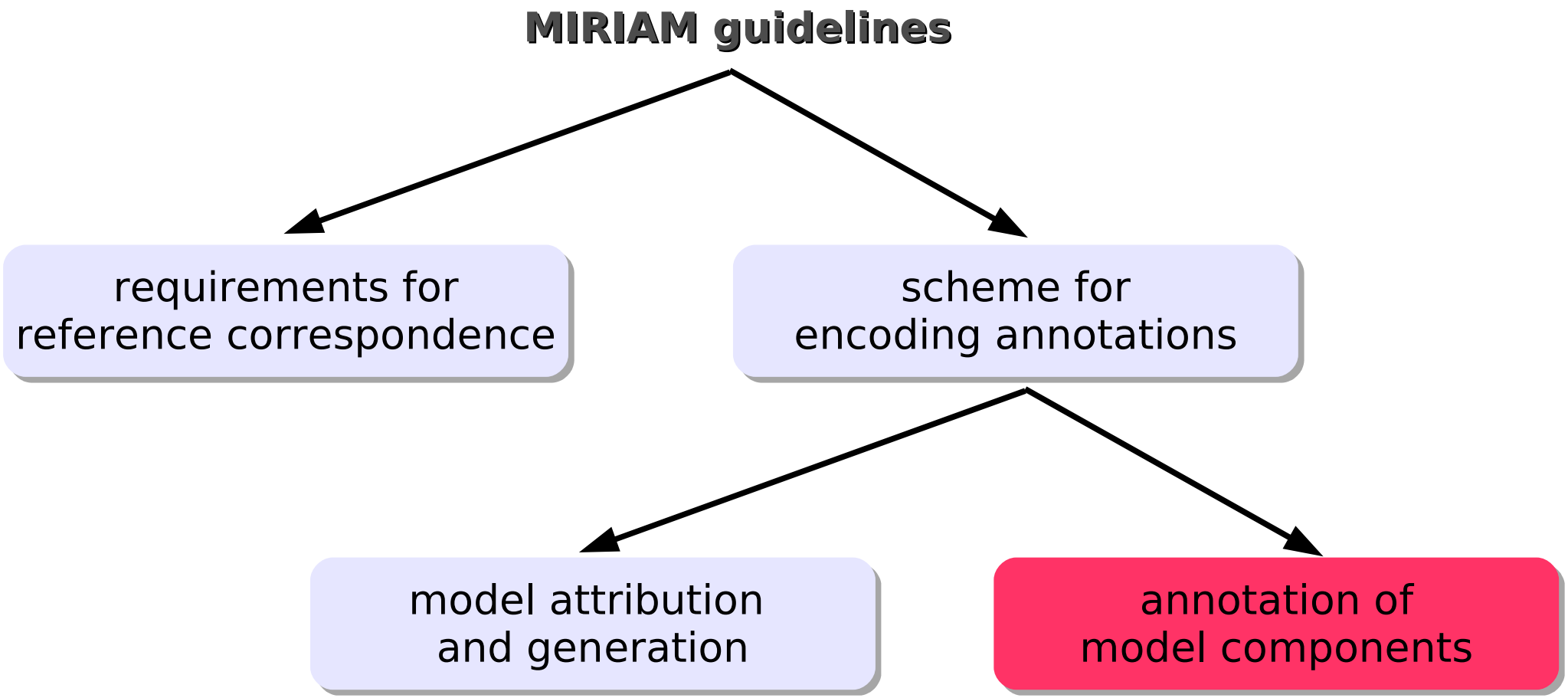
MIRIAM-compliant model. A model that passes all the tests and fulfills all the conditions listed in MIRIAM.

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.

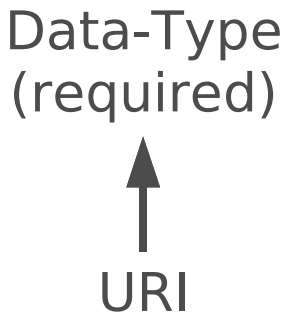

MIRIAM guidelines
requirements for reference correspondence
scheme for encoding annotations
model attribution and generation
annotation of model components


MIRIAM guidelines
requirements for reference correspondence
scheme for encoding annotations
model attribution and generation
annotation of model components

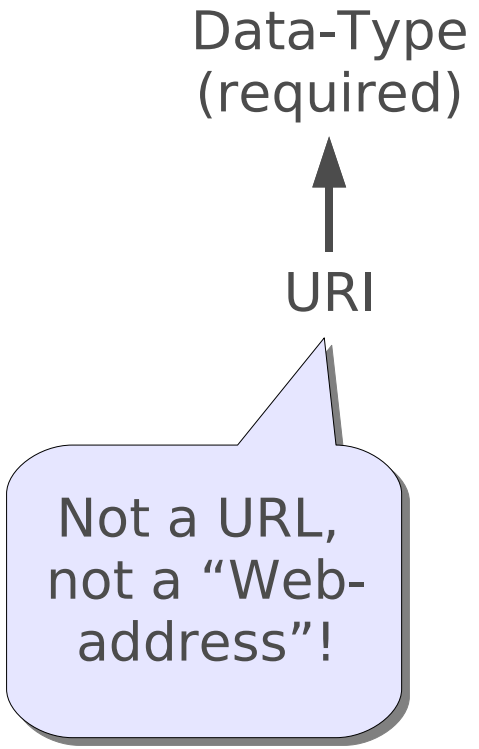
Characteristics of a useful identifier

- **Unique**
  an identifier must never be assigned to two different objects;
- **Perennial**
  the identifier is constant and its lifetime is permanent;
- **Standards compliant**
  must conform on existing standards, such as URI;
- **Resolvable**
  identifiers must be able to be transformed into locations of online resources storing the object or information about the object;
- **Free of use**
  everybody should be able to use and create identifiers, freely and at no cost.

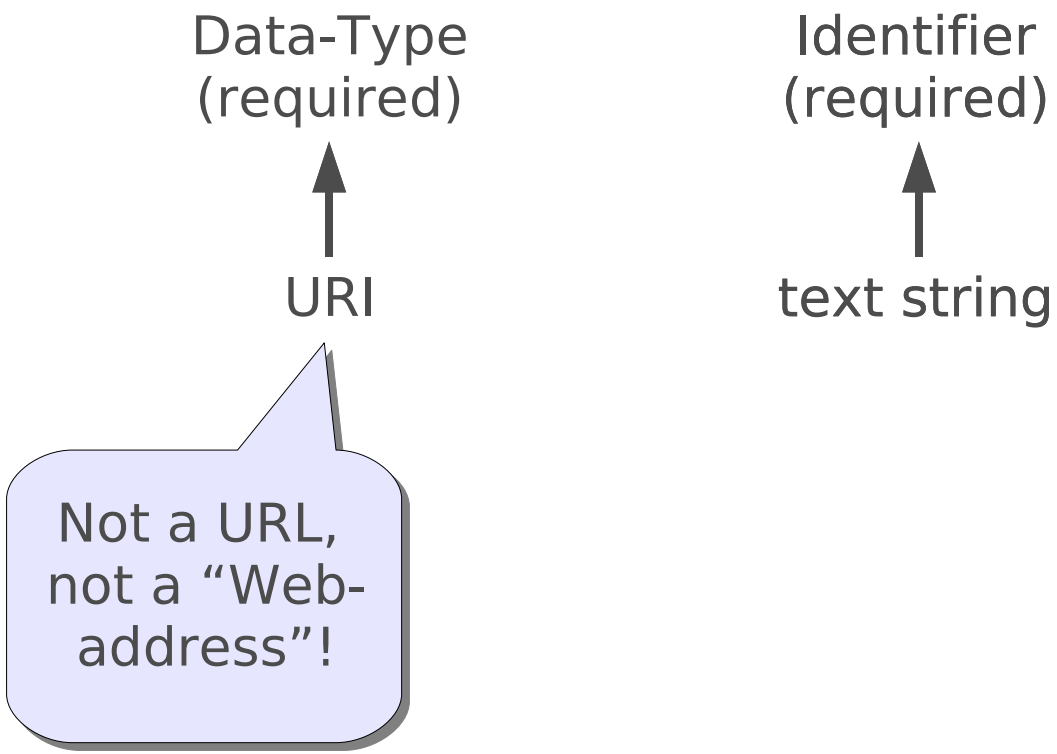
MIRIAM annotation



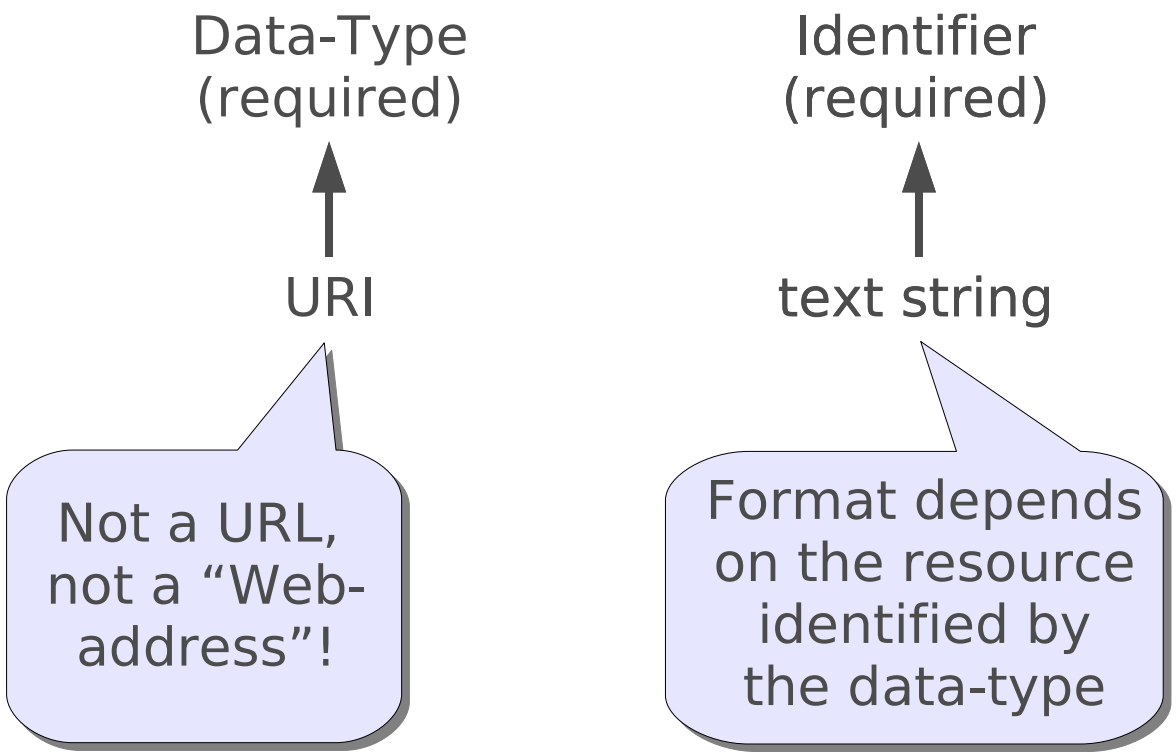
MIRIAM annotation



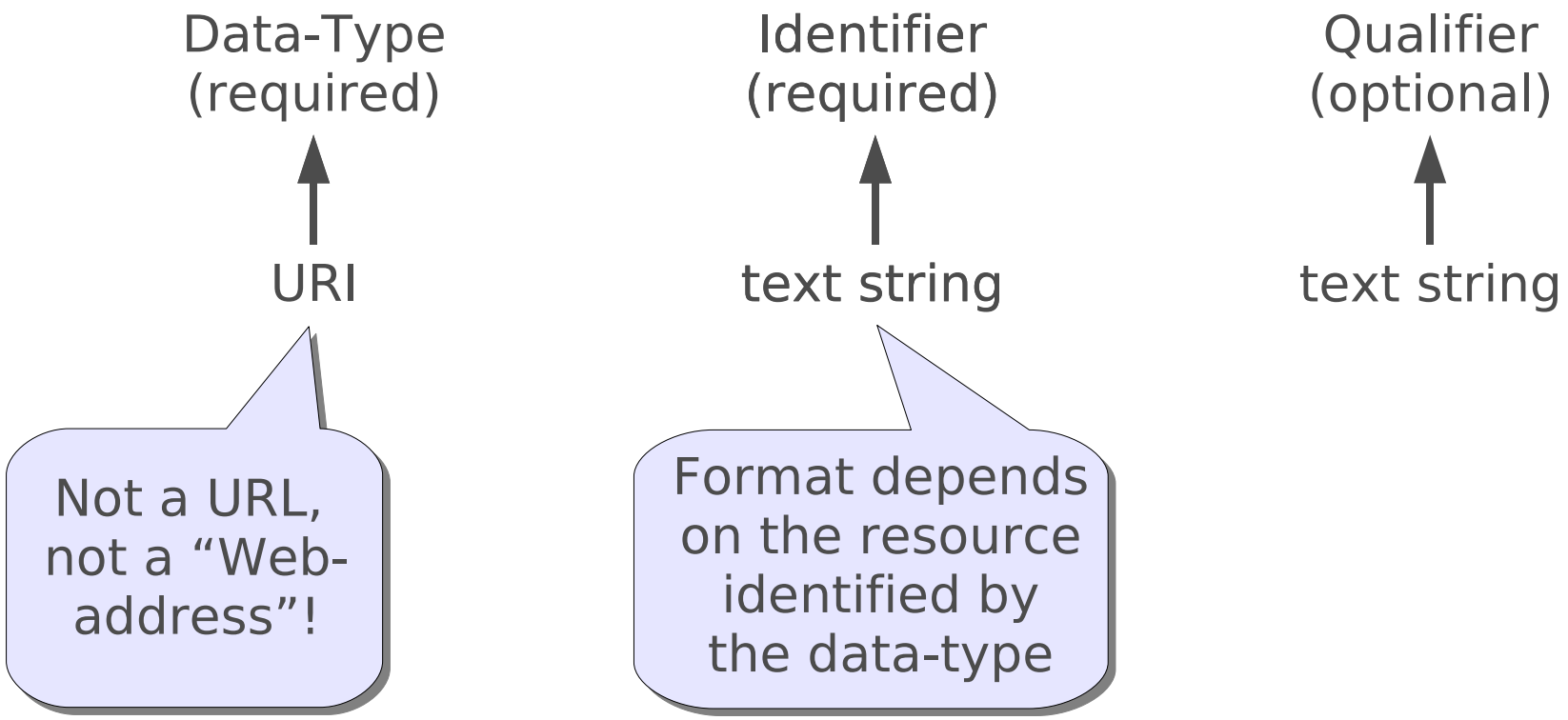
MIRIAM annotation



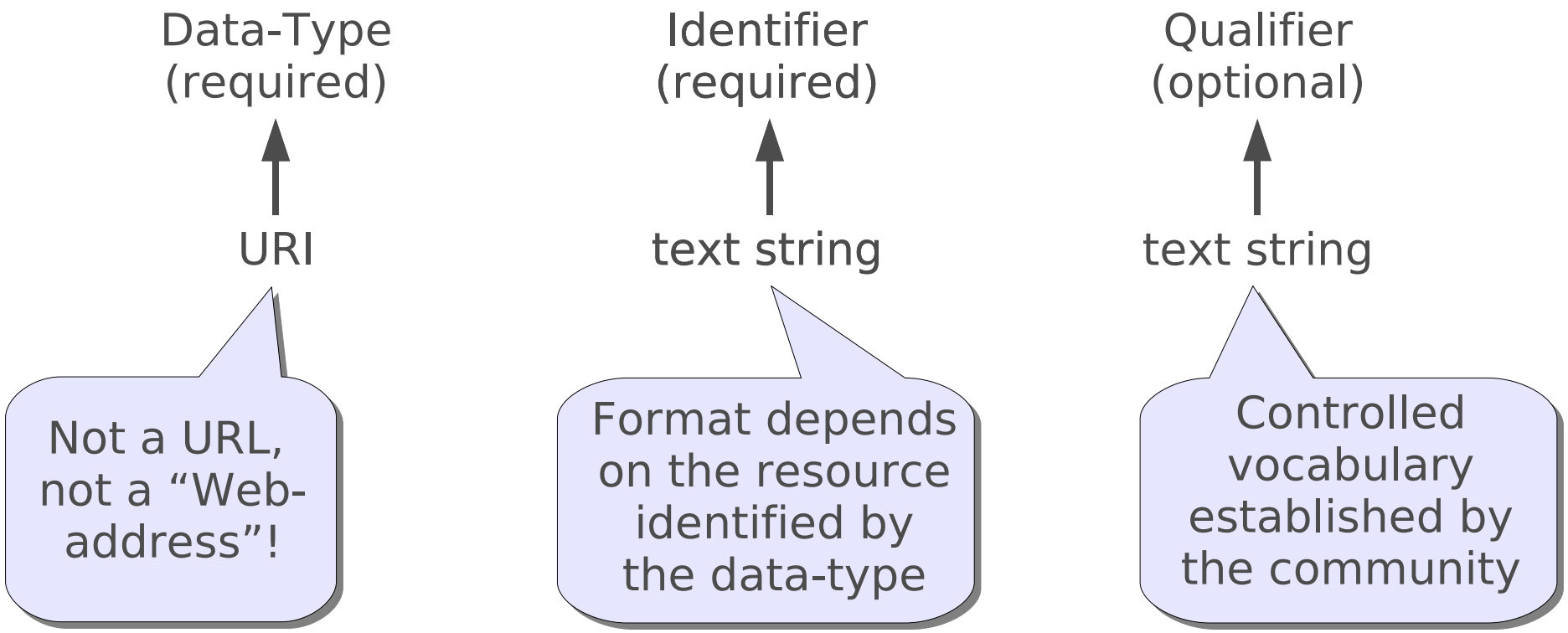
MIRIAM annotation



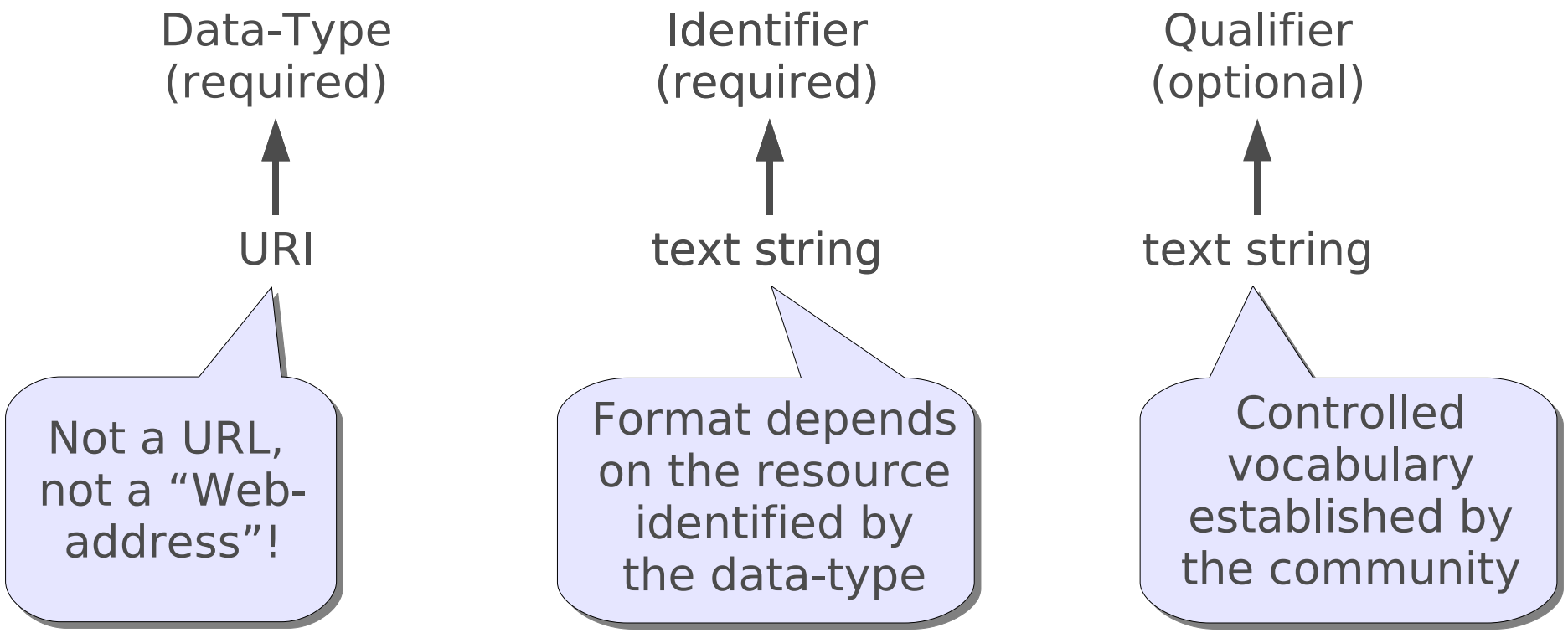
MIRIAM annotation



MIRIAM annotation



MIRIAM annotation



The data-type and the identifier can be combined in a single URI

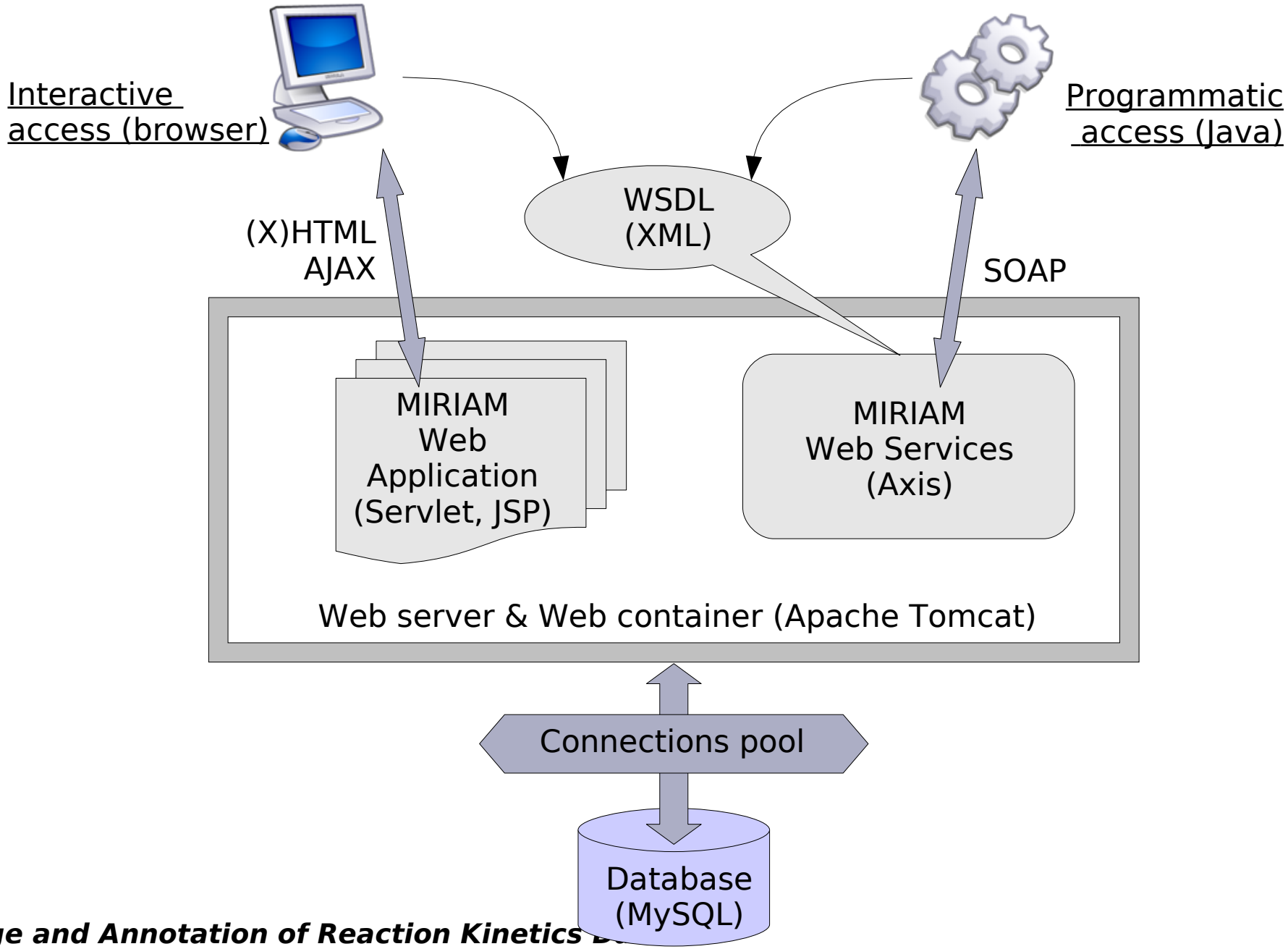
URL style: `http://www.MyResource.org/#MyIdentifier`

URN style: `urn:lsid:MyResource.org:MyIdentifier`

MIRIAM Resources

- **MIRIAM Database**
  Core element of the resource, storing all the information about the data-types and associated information;
- **MIRIAM Web Services**
  SOAP-based application programming interface (API) for querying MIRIAM Database
- **MIRIAM Library**
  Library to use MIRIAM Web Services
- **MIRIAM Web Application**
  Interactive web interface for browsing and querying MIRIAM Database, and submit or edit data-types.

MIRIAM Resources



MIRIAM

- Browse
- Query
- Submit new
- Export
- Sign In

- News
- Web Services
- MIRIAM Standard
- BioModels Qualifiers
- FAQ
- MIRIAM on SourceForge
- Contact

Browse the data-types

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

Name	URI	Definition
BIND	http://www.bind.ca/	BIND is a database of protein-protein interactions. This data-resource is not open-access.
BioModels Database	http://www.ebi.ac.uk/biomodels/	BioModels Database is a data resource that allows biologists to store, search and retrieve published mathematical models of biological interests.
ChEBI	http://www.ebi.ac.uk/chebi/	Chemical Entities of Biological Interest (ChEBI) is a freely available dictionary of molecular entities focused on 'small' chemical compounds.
CluSTr	http://www.ebi.ac.uk/clustr/	The CluSTr database offers an automatic classification of UniProt Knowledgebase and IPI proteins into groups of related proteins. The clustering is based on analysis of all pairwise comparisons (Smith-Waterman) between protein sequences.
DOI	http://www.doi.org/	The Digital Object Identifier System is for identifying content objects in the digital environment.
Ensembl	http://www.ensembl.org/	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.
Enzyme Nomenclature	http://www.ec-code.org/	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.
FlyBase	http://www.flybase.org/	FlyBase is the database of the Drosophila Genome Projects and of associated literature.
Gene Ontology	http://www.geneontology.org/	The Gene Ontology project provides a controlled vocabulary to describe gene and gene product attributes in any organism.
ICD	http://www.who.int/classifications/icd/	The International Classification of Diseases is the international standard diagnostic classification for all general epidemiological and many health management purposes.
IntAct	http://www.ebi.ac.uk/intact/	IntAct provides a freely available, open source database system and analysis tools for protein interaction data.
InterPro	http://www.ebi.ac.uk/interpro/	InterPro is a database of protein families, domains and functional sites in which identifiable features found in known proteins can be applied to unknown protein sequences.
KEGG Compound	http://www.genome.jp/kegg/compound/	KEGG compound contains our knowledge on the universe of chemical substances that are relevant to life.
KEGG Drug	http://www.genome.jp/kegg/drug/	KEGG DRUG contains chemical structures of drugs and additional information such as therapeutic categories and target molecules.
KEGG Glycan	http://www.genome.jp/kegg/glycan/	KEGG GLYCAN, a part of the KEGG LIGAND database, is a collection of experimentally determined glycan structures. It contains all unique structures taken from CarbBank, structures entered from recent publications, and structures present in KEGG

requirements for a MIRIAM-compliant resource

- **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 reaction 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

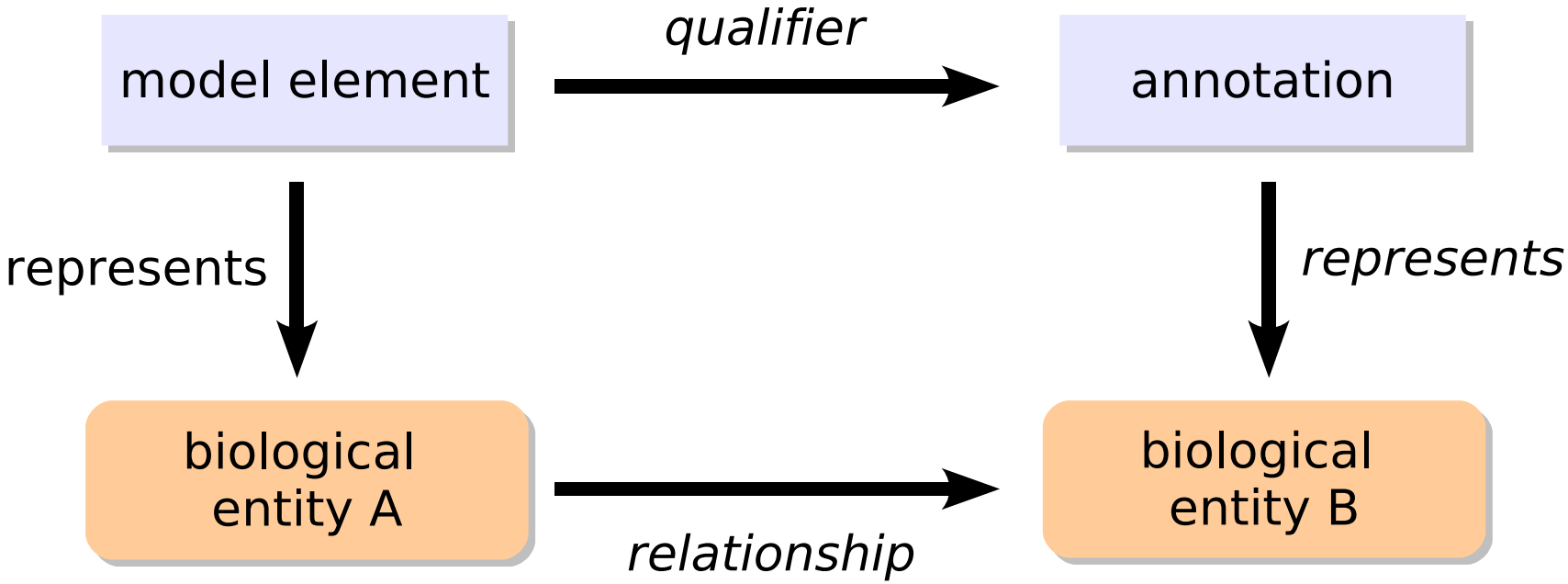
MIRIAM
- Browse
- Query
- Submit new
- Export
- Sign In
- News
- Web Services
- MIRIAM Standard
- BioModels Qualifiers
- FAQ
- MIRIAM on SourceForge
- Contact

MIRIAM

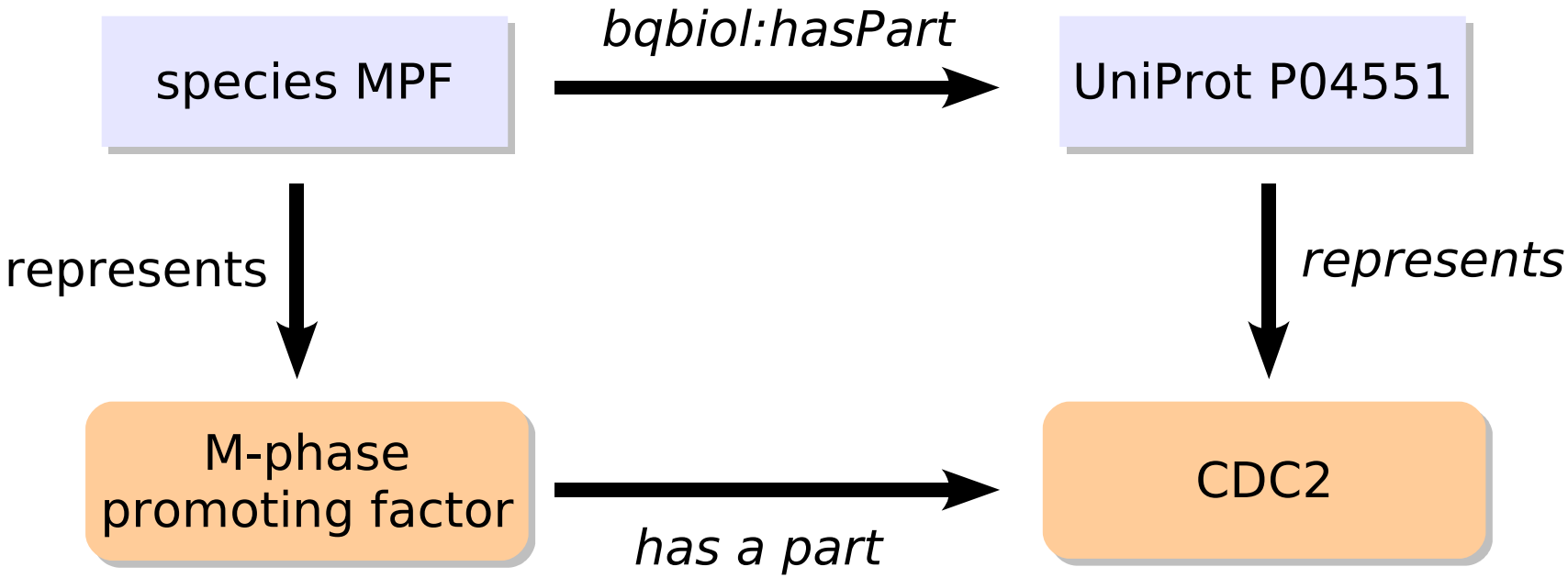
Data-type: ***Enzyme Nomenclature***

Name		
Identifier		MIR:00000004
Name		Enzyme Nomenclature
Synonyms		Enzyme Classification
		EC code
		EC
URIs		
Official URL		http://www.ec-code.org/
Official URN		urn:lsid:ec-code.org
Deprecated		http://www.ebi.ac.uk/IntEnz/
Information		
Definition		The Enzyme Classification contains the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzyme-catalysed reactions.
Identifier Pattern		^\d+|\d+\.(-|\d+)|\d+\.\d+\.(-|\d+)|\d+\.\d+\.\d+\.(-|\d+)$
Physical Locations		
Resource #1	**Data Entry**	http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&ec=**$id**
	Data Resource	http://www.ebi.ac.uk/intenz/
	Information	IntEnZ (Integrated relational Enzyme database)
	Institution	European Bioinformatics Institute, United Kingdom
Resource #2	**Data Entry**	http://www.genome.jp/dbget-bin/www_bget?ec:**$id**
	Data Resource	http://www.genome.jp/dbget-bin/www_bfind?enzyme
	Information	KEGG Ligand Database for Enzyme Nomenclature
	Institution	Kyoto University Bioinformatics Center, Japan
Resource #3	**Data Entry**	http://us.expasy.org/cgi-bin/nicezyme.pl?**$id**
	Data Resource	http://us.expasy.org/enzyme/
	Information	Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)
	Institution	Swiss Institute of Bioinformatics, Switzerland
Documentation		
URL(s)		http://www.chem.qmul.ac.uk/iubmb/enzyme/
		⊞ http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-view+MedlineFull+[medline-PMID:10812475]
Miscellaneous		
Date of creation		2006-08-14 19:38:06 GMT
Date of last modification		2007-03-09 16:00:07 GMT

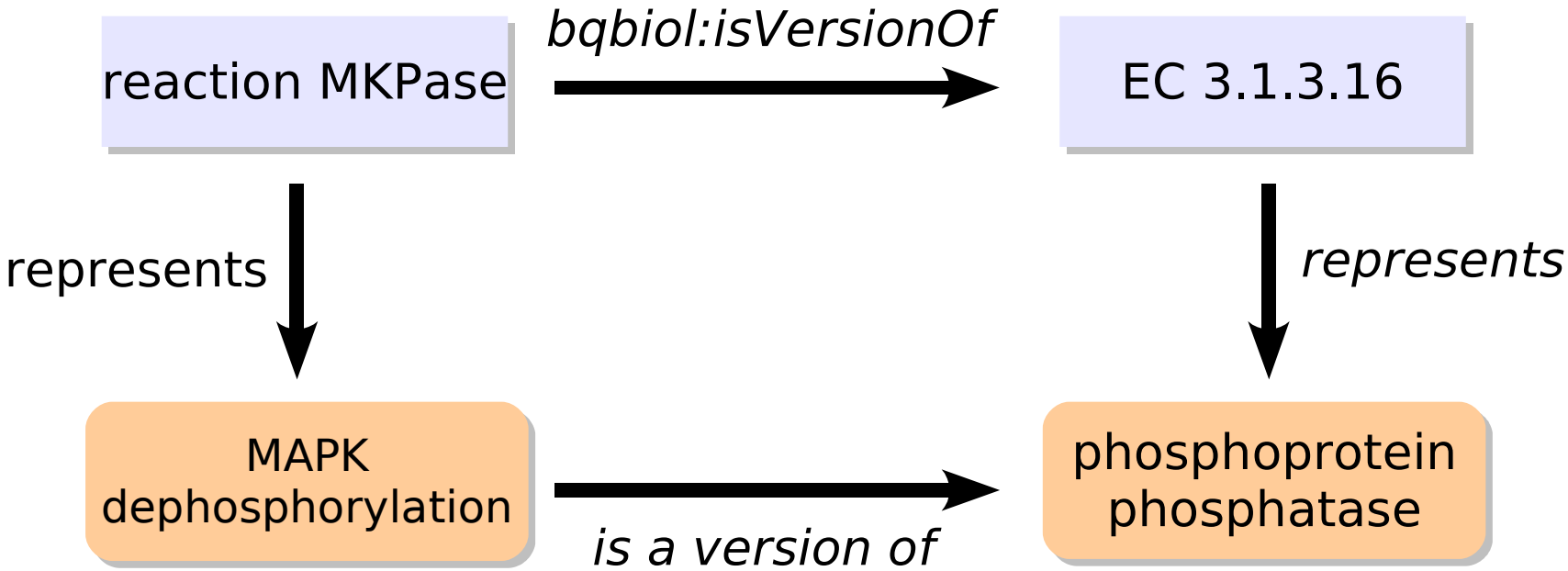
Qualification of annotation



Qualification of annotation



Qualification of annotation



Current 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.

SBML and MIRIAM

```
<species id="Ca_calmodulin" metaid="cacam">
  <annotation>
    <rdf:RDF
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:bqbiol="http://biomodels.net/biology-qualifiers/">
      <rdf:Description rdf:about="#cacam">
        <bqbiol:hasPart>
          <rdf:Bag>
            <rdf:li rdf:resource="http://www.uniprot.org/#P62158"/>
            <rdf:li rdf:resource="http://www.ebi.ac.uk/chebi/#CHEBI:29108"/>
          </rdf:Bag>
        </bqbiol:hasPart>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```


How to chose the right annotation?

- Choose the right level of abstraction
  - Precise enough to narrow down to the meaning of the component: not “vertebrate” for “mammal”; not “protein phosphorylation” for “MAPK phosphorylation”
  - Comprehensive enough to encompass the whole semantic of the component: not “rodent” for “mammal”; not “threonine phosphorylation” for “MAPK phosphorylation”
- Complete
  - If annotation of one part of a complex, annotate all the parts: Partial annotation can be misleading. The function of Cyclin/CDK is different of the function of CDK
  - (if annotating resource incomplete, file a curator request ...)
- Comply with usual identifier requirements
  - Unique, perennial, standard compliant, resolvable ...

Hidden assumptions

```
<reaction>
   <listOfReactants>
     <speciesReference species="S" />
   </listOfReactants>
   <listOfProducts>
     <speciesReference species="P" />
   </listOfProducts>
   <listOfModifiers>
     <speciesReference species="E" />
   </listOfModifiers>
   <kineticLaw>
     <listOfParameters>
       <parameter id="Km"/>
       <parameter id="kp"/>
     </listOfParameters>
     <math xmlns="http://www.w3.org/1998/Math/MathML">
       <apply>
         <divide/><apply>
                    <times/><ci>E</ci>
                            <ci>kp</ci>
                            <ci>S</ci>
                  </apply>
                  <apply>
                    <plus/><ci>Km</ci>
                           <ci>S</ci>
                  </apply>
       </apply>
     </math>
   </kineticLaw>
 </reaction>
```

$$\overset{E}{S \rightarrow P}$$

Import in a discrete simulator

Henri-Michaelis-Menten

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_p} E+P \; ; \; k_1 = k_{-1}/K_m$$

Van Slyke-Cullen

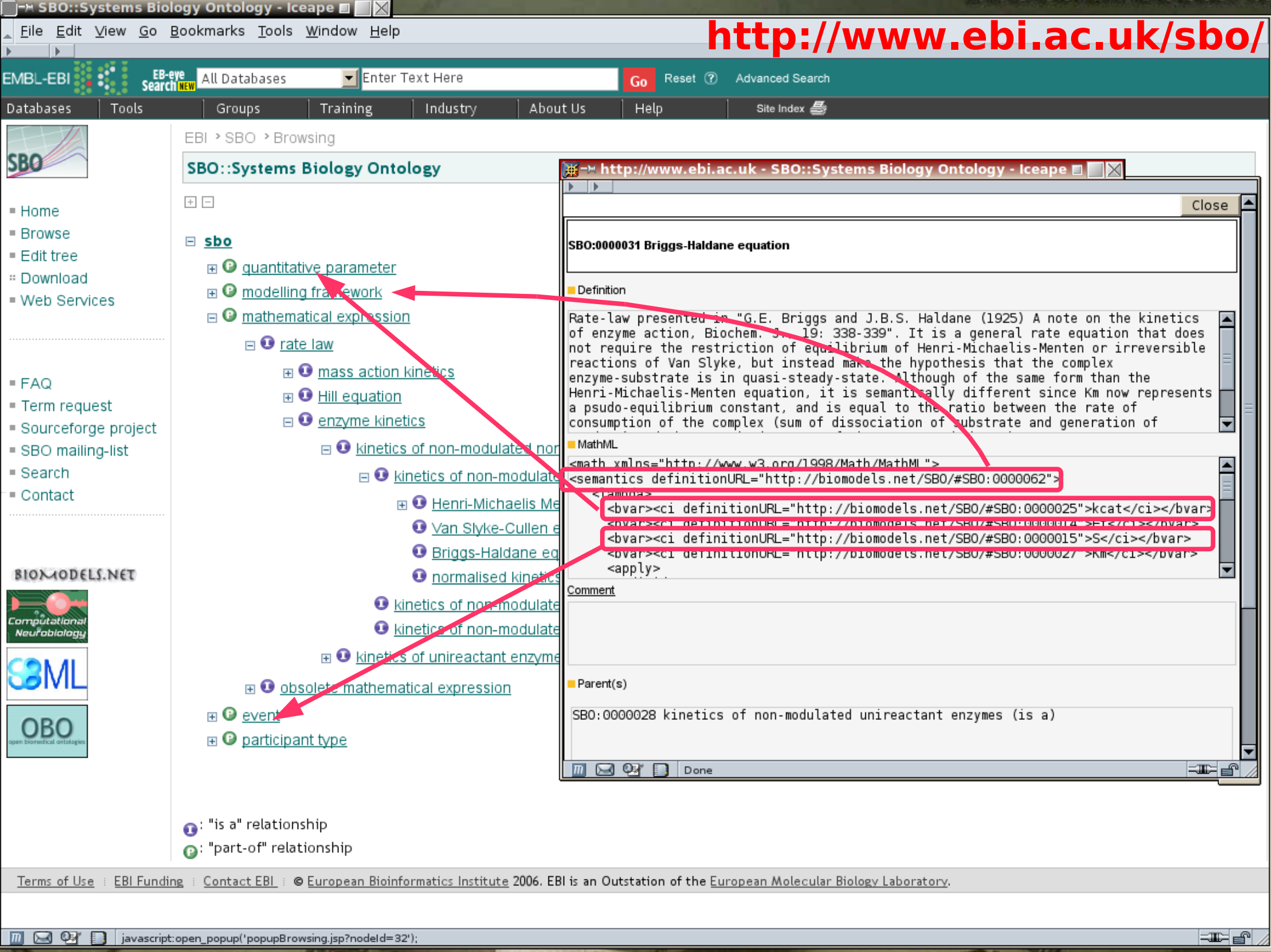
$$E+S \underset{k_{-1}}{\overset{k_1}{\rightarrow}} ES \xrightarrow{k_p} E+P \; ; \; k_1 = k_p/K_m$$

Briggs-Haldane

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_p} E+P \; ; \; k_1 = (k_{-1}+k_p)/K_m$$

Systems Biology Ontology vocabularies

- Types and roles of reaction participants, including terms like "substrate", "catalyst" etc., but also "macromolecule", or "channel"
- Parameter used in quantitative models. This vocabulary includes terms like "Michaelis constant" , "forward unimolecular rate constant"etc. A term may contain a precise mathematical expression stored as a MathML lambda function. The variables refer to other parameters.
- Mathematical expressions. Examples of terms are "mass action kinetics", "Henri-Michaelis-Menten equation" etc. A term may contain a precise mathematical expression stored as a MathML lambda function. The variables refer to the other vocabularies.
- Modelling framework to precise how to interpret the rate-law. E.g. "continuous modelling", "discrete modelling" etc.
- Event type, such as "catalysis" or "addition of a chemical group".


SBO::Systems Biology Ontology - Iceape
File Edit View Go Bookmarks Tools Window Help
http://www.ebi.ac.uk/sbo/
EMBL-EBI
EB-eye Search NEW
All Databases
Enter Text Here
Go
Reset
Advanced Search
Databases Tools Groups Training Industry About Us Help Site Index
EBI › SBO › Browsing
SBO::Systems Biology Ontology
Home
Browse
Edit tree
Download
Web Services
FAQ
Term request
Sourceforge project
SBO mailing-list
Search
Contact
BIOMODELS.NET
Computational Neurobiology
SBML
OBO
sbo
quantitative parameter
modelling framework
mathematical expression
rate law
mass action kinetics
Hill equation
enzyme kinetics
kinetics of non-modulated nor
kinetics of non-modulate
Henri-Michaelis Me
Van Slyke-Cullen e
Briggs-Haldane eq
normalised kinetics
kinetics of non-modulate
kinetics of non-modulate
kinetics of unireactant enzyme
obsolete mathematical expression
event
participant type
: "is a" relationship
: "part-of" relationship
http://www.ebi.ac.uk - SBO::Systems Biology Ontology - Iceape
Close
SBO:0000031 Briggs-Haldane equation
Definition
Rate-law presented in "G.E. Briggs and J.B.S. Haldane (1925) A note on the kinetics of enzyme action, Biochem. J., 19: 338-339". It is a general rate equation that does not require the restriction of equilibrium of Henri-Michaelis-Menten or irreversible reactions of Van Slyke, but instead make the hypothesis that the complex enzyme-substrate is in quasi-steady-state. Although of the same form than the Henri-Michaelis-Menten equation, it is semantically different since Km now represents a psudo-equilibrium constant, and is equal to the ratio between the rate of consumption of the complex (sum of dissociation of substrate and generation of
MathML
<math xmlns="http://www.w3.org/1998/Math/MathML">
<semantics definitionURL="http://biomodels.net/SBO/#SBO:0000062">
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000025">kcat</ci></bvar>
<bvar><ci definitionURL="http://biomodels.net/SBO/#SBO:0000015">S</ci></bvar>
<apply>
Comment
Parent(s)
SBO:0000028 kinetics of non-modulated unireactant enzymes (is a)
Done
Terms of Use | EBI Funding | Contact EBI | © European Bioinformatics Institute 2006. EBI is an Outstation of the European Molecular Biology Laboratory.
javascript:open_popup('popupBrowsing.jsp?nodeId=32');

SBML and SBO

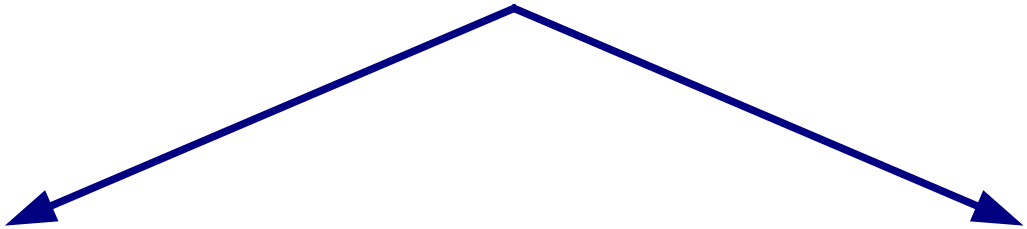
```
<reaction sboTerm="SBO:0000172">
  <listOfReactants>
    <speciesReference species="S" sboTerm="SBO:0000015"/>
  </listOfReactants>
  <listOfProducts>
    <speciesReference species="P" sboTerm="SBO:0000011"/>
  </listOfProducts>
  <listOfModifiers>
    <speciesReference species="E" sboTerm="SBO:0000014"/>
  </listOfModifiers>
  <kineticLaw sboTerm="SBO:0000031">
    <listOfParameters>
      <parameter id="K1" sboTerm="SBO:0000008"/>
      <parameter id="kp" sboTerm="SBO:0000025"/>
    </listOfParameters>
    <math xmlns="http://www.w3.org/1998/Math/MathML">
      <apply>
        <divide/><apply>
                   <times/><ci>E</ci>
                           <ci>kp</ci>
                           <ci>S</ci>
                 </apply>
                 <apply>
                   <plus/><ci>K1</ci>
                          <ci>S</ci>
                 </apply>
      </apply>
    </math>
  </kineticLaw>
</reaction>
```


SBML and SBO

```
<reaction sboTerm="SBO:0000172">
  <listOfReactants>
    <speciesReference species="S" sboTerm="SBO:0000015"/>
  </listOfReactants>
  <listOfProducts>
    <speciesReference species="P" sboTerm="SBO:0000011"/>
  </listOfProducts>
  <listOfModifiers>
    <speciesReference species="E" sboTerm="SBO:0000014"/>
  </listOfModifiers>
  <kineticLaw sboTerm="SBO:0000031">
    <listOfParameters>
      <parameter id="K1" sboTerm="SBO:0000008"/>
      <parameter id="kp" sboTerm="SBO:0000025"/>
    </listOfParameters>
    <math xmlns="http://www.w3.org/1998/Math/MathML">
      <apply>
        <divide/><apply>
                   <times/><ci>E</ci>
                           <ci>kp</ci>
                           <ci>S</ci>
                 </apply>
                 <apply>
                   <plus/><ci>K1</ci>
                          <ci>S</ci>
                 </apply>
      </apply>
    </math>
  </kineticLaw>
</reaction>
```

SBML and SBO

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



continuous simulator

$$v = \frac{C \cdot V \cdot [A]}{(U + [A])}$$

discrete simulator

$$v1 = \frac{(k_{-1} + V)}{U} \cdot [A] \cdot [C]$$

$$v2 = k_{-1} \cdot [D]$$

$$v3 = V \cdot [D]$$

SBML and SBO

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

functional compartment

simple chemical

simple chemical

enzyme

catalysis

substrate

product

catalyst

Briggs-Haldane equation

Km

kcat

Systems Biology Ontology

Assignee: (?) Any | Status: (?) Any | Category: (?) Any | Group: (?) Any

Show only: Submitter username (show mine) : | Summary keyword:

Sort By: (?) ID | Descending | Browse

List of suggested SBO term creations or modification.

Request ID	Summary	Open Date	Priority	Status	Assigned To	Submitted By
1682477	create types of simple chemicals needed by SBGN	* 2007-03-16 23:17	5	Open	lenov	lenov
1672292	types of RNAs	* 2007-03-02 09:55	5	Closed	lenov	lenov
1599415	transporter	* 2006-11-19 23:38	5	Closed	nobody	lenov
1598503	Acid dissociation constant	* 2006-11-17 17:02	5	Closed	nobody	golebiewskim
1598488	pK	* 2006-11-17 16:47	5	Closed	nobody	golebiewskim
1598430	Equilibrium constant	* 2006-11-17 16:00	5	Closed	nobody	golebiewskim
1598416	EC50 and IC50	* 2006-11-17 15:44	5	Closed	nobody	golebiewskim
1598404	Catalytic efficiency	* 2006-11-17 15:16	5	Closed	nobody	golebiewskim
1597770	move enzyme	* 2006-11-16 14:44	5	Closed	nobody	lenov

TErminology for the Description of DYnamics (https://sourceforge.net/projects/teddyontology)

- Dynamical behaviour of biological variables, including terms like “oscillation”, “bistability”, “steady-state”, “equilibrium” etc.
- Characteristics used to describe dynamics, such as “period” , “limit-cycle”, “Hopf bifurcation”etc.
- Functional role of model elements. Examples of terms are “negative feedback”, “integrator” etc.

An international collaboration

- EBI
  - **Nicolas Le Novère**
  - Mélanie Courtot
  - Marco Donizelli
  - Arnaud Henry
  - Christian Knuepfer
  - Chen Li
  - Lu Li
  - Camille Laibe
  - Nicolas Rodriguez
  - Alexander Broicher
- SBML team
  - **Michael Hucka**
  - **Andrew Finney**
  - Benjamin Borstein
  - Harish Dharuri
  - Enuo He
  - Sarah Keating
  - Maria Schilstra
  - Bruce Shapiro
- NCBS
  - Upinder Bhalla
  - Harsha Rani
- University of Washington
  - Herbert Sauro
- Vienna TBI
  - Rainer Machne
  - James Lu
- Systems Biology Institute
  - Hiroaki Kitano
  - Akira Funahashi
- JWS Online
  - Jacky Snoep
  - Hans Westerhoff
- 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.

European Bioinformatics Institute

British outstation of the European Molecular Biology Laboratory

- Databases
  - Sequences, structures
  - Transcriptomics, Proteomics pathways, models
  - Controlled vocabularies and dictionaries

- Research groups
  - Structural Genomics (Thornton)
  - Molecular Evolution (Goldman)
  - Text-Mining (Rebholz-Schumman)
  - Computational Systems Biology (Le Novère)
  - Statistical array analysis (Huber)
  - Genomic analysis of regulatory systems (Luscombe)
  - Systems Biology of ES cells (Bertone)