EMBL-EBI

MIBBI, MIASE and all that

Nicolas Le Novère, EMBL-EBI

*“Perfection is achieved, not when there is nothing more to add,
but when there is nothing left to take away.”*

Antoine De Saint Exupery

What are standard reporting guidelines?

- List of the core set of information that has to be provided with a data-set, so that a user is able to make sensible use of it.
- “Minimal”: Only the information shared by all the providers/users of this type of data should be considered.
- “Standard”: The information should be provided in a form that can be fully interpreted by all.

_computational BIOLOGY

COMMENTARY

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

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

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

To fully understand the context, methods, data and conclusions that pertain to an experiment, one must have access to a range of background information. However, the current diversity of experimental designs and analytical techniques complicates the discovery and evaluation of experimental data; furthermore, the increasing rate of production of those data compounds the problem. Community opinion increasingly favors that a regularized set of the available metadata ('data about the data') per-

guidelines for reporting proteomics experiments and describing systems biology models are gaining broader support in their respective database communities[8,9]; and progress is being made toward the standardization of the reporting of clinical trials in the medical literature[10]. Such minimum information checklists promote transparency in experimental reporting, enhance accessibility to data and support effective quality assessment, increasing the general value of a body of work (and the com-

overlaps in scope and arbitrary decisions on wording and substructuring inhibit their use in combination. These issues present difficulties for checklist users, especially those who routinely combine information from several disciplines. Here we explore some of the issues arising from the development of checklists in relative isolation, discuss the potential benefits of greater coordination and describe the mechanisms we have put in place to facilitate such coordination. In summary, we present

MIBBI: Minimum Information for Biological and Biomedical Investigations

Project News

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

Site navigation

The MIBBI Portal

Access to Minimum Information guidelines for diverse bioscience domains

About us

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

The MIBBI Foundry

Towards the next generation of MI guidelines for the biosciences

Project news

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

Related resources

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

Discussion

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

MIBBI search

A Google™ Custom

http://www.mibbi.org

links

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

search

Go | Search

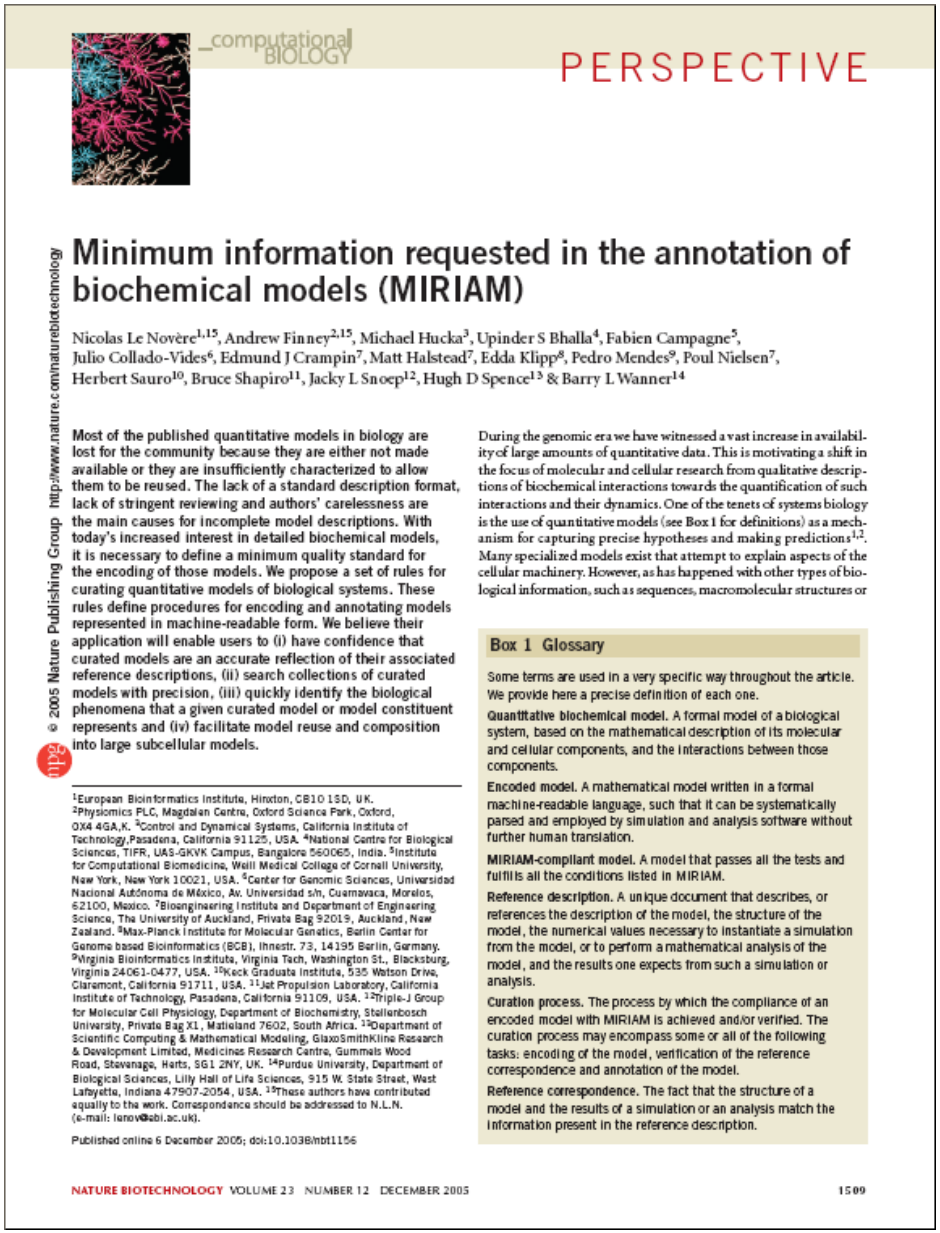
toolbox

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

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

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

- 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.

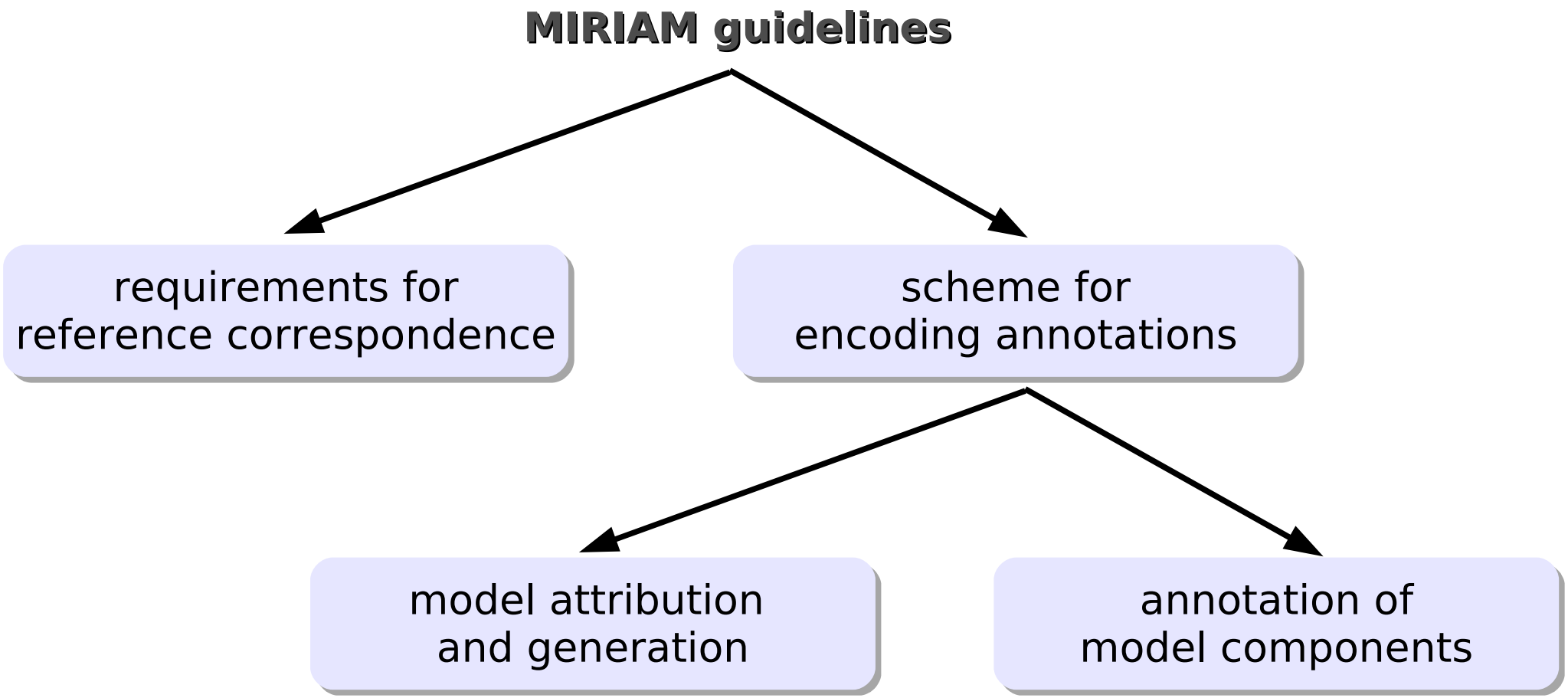
Encoded model. A mathematical model written in a formal machine-readable language, such that it can be systematically parsed and employed by simulation and analysis software without further human translation.

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.

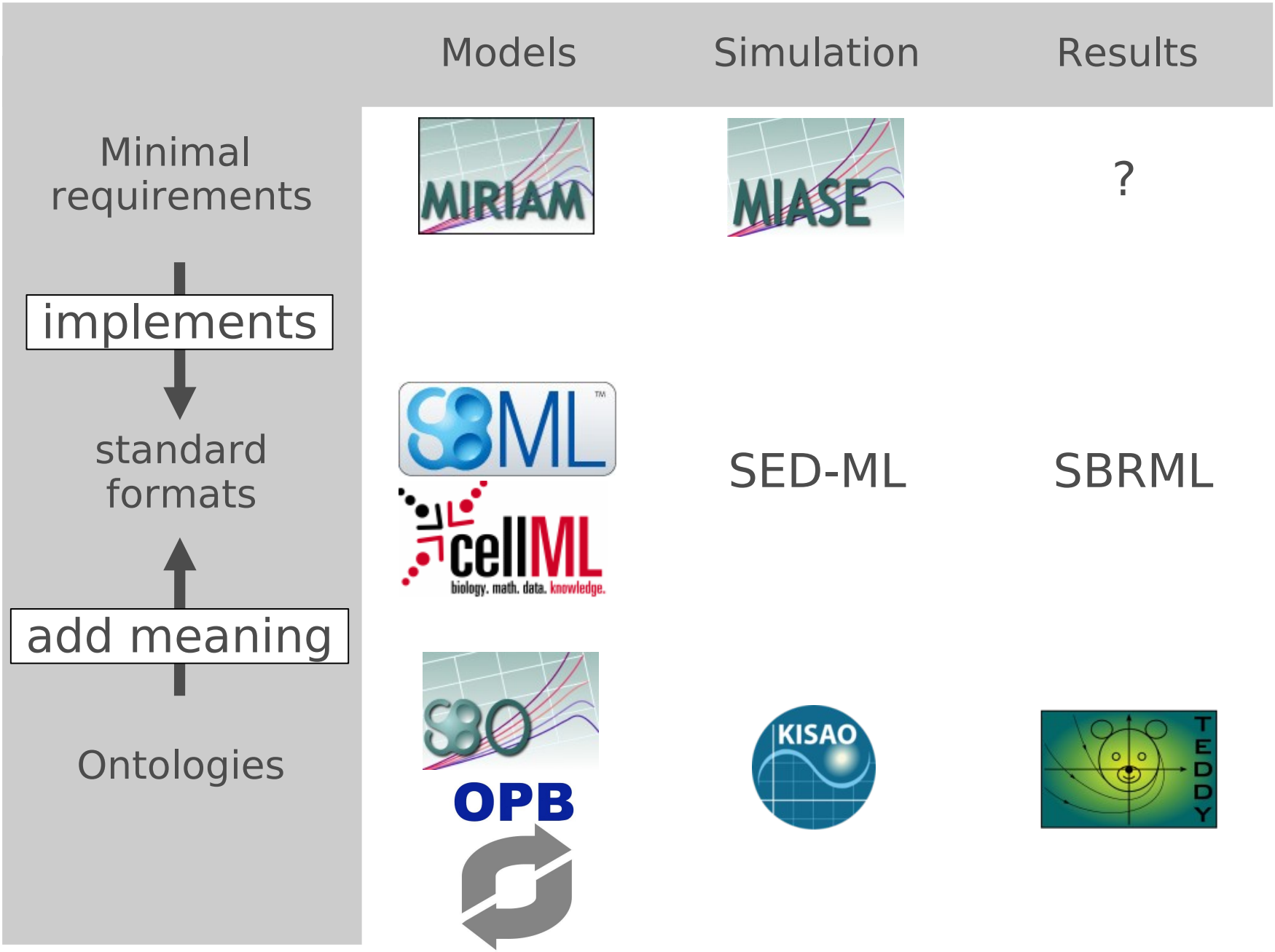
Reference correspondence. The fact that the structure of a model and the results of a simulation or an analysis match the information present in the reference description.


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

Spread of MIRIAM support

- Research applications: Yeast Metabolic Model (Herrgård M.J. et al (2008) , *Nat Biotechnol* 26: 1155-1160), : 2152 species, 1857 reactions, 4861 MIRIAM annotations. Human Metabolic Model: 4889 species, 8866 reactions, **66968** MIRIAM annotations
- Data resources using MIRIAM: BioModels Database (kinetic models), PSI consortium (protein interactions), Reactome (pathways), SABIO-RK (reaction kinetics)
- Software using MIRIAM: COPASI (Simulation), Snazzer (Network analysis, Simulations), Systems Biology Workbench (model design and simulation), The Virtual Cell (Simulation), SBMLsemantics, libSBMLannotation, SAINT
- Statistics of usage of MIRIAM Resources: Web Interface: 13 494 page requests in 2008, 245.15 MB data transferred, 1 539 unique users. WebServices: 578 620 requests

Mosaic of standards for computational modeling



What is encoded in model description?

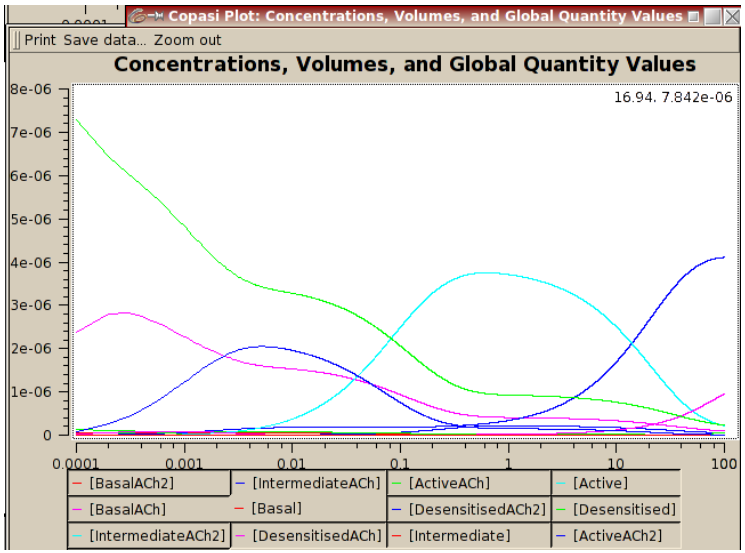
- List and structure of biological entities and processes
- Mathematical descriptions of the amounts and transformations

Missing:

- What to do with the models
- How to do it
- What to do with the results

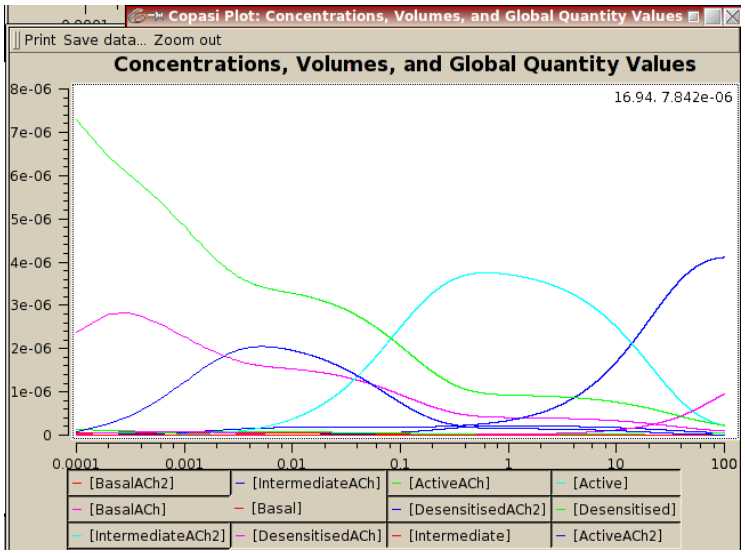
Simple post-processing

Edelstein et al 1996 (BIOMD0000000002)

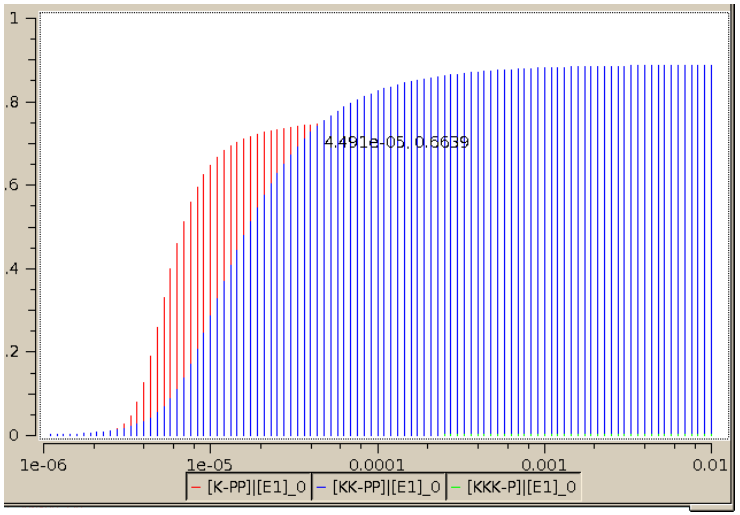


Iterative simulations, parameter scans

Edelstein et al 1996 (BIOMD0000000002)

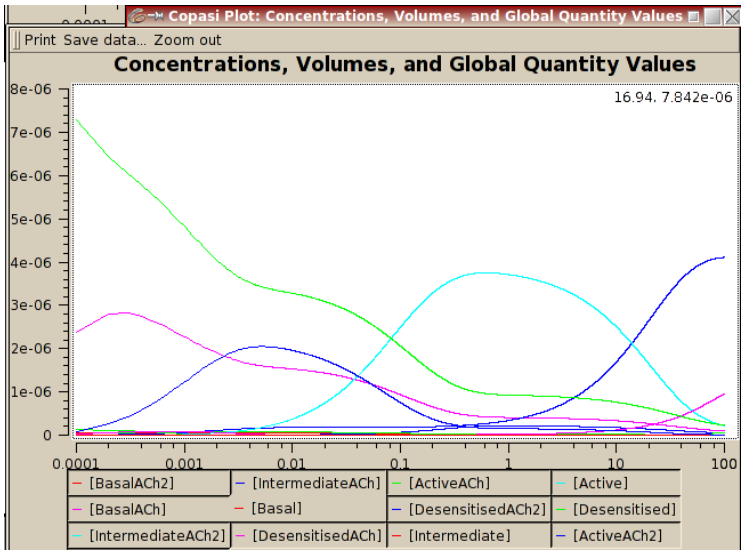


Huang & Ferrell (BIOMD0000000009)

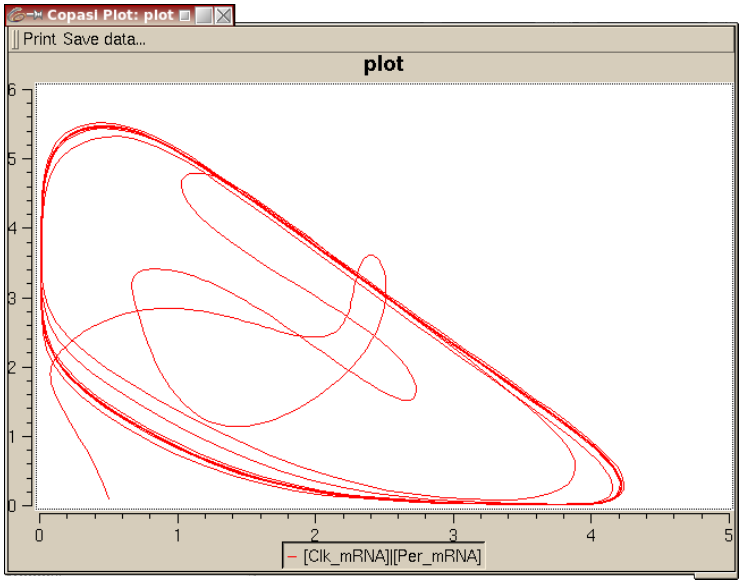


Non-trivial output, phase-plane, bifurcations

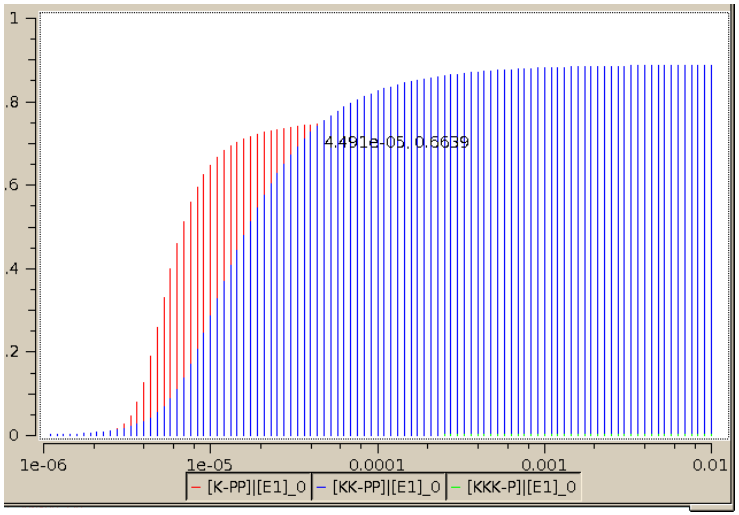
Edelstein et al 1996 (BIOMD0000000002)



Ueda, Hagiwara, Kitanol 2001 (BIOMD0000000022)

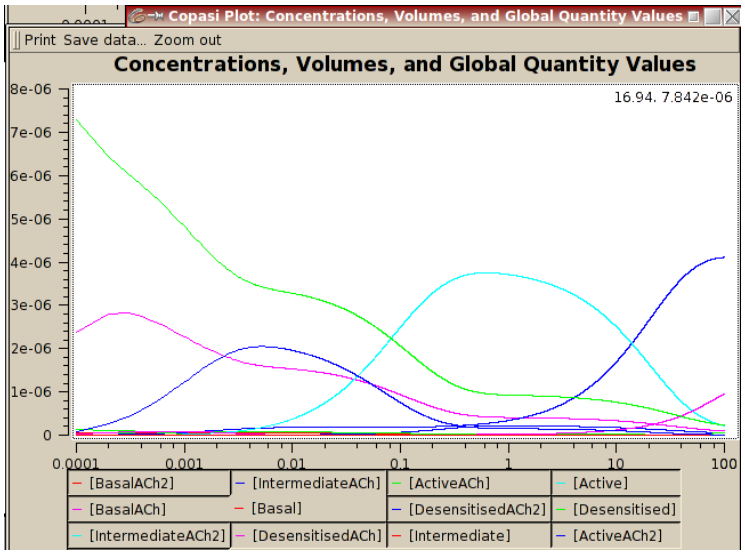


Huang & Ferrell (BIOMD0000000009)

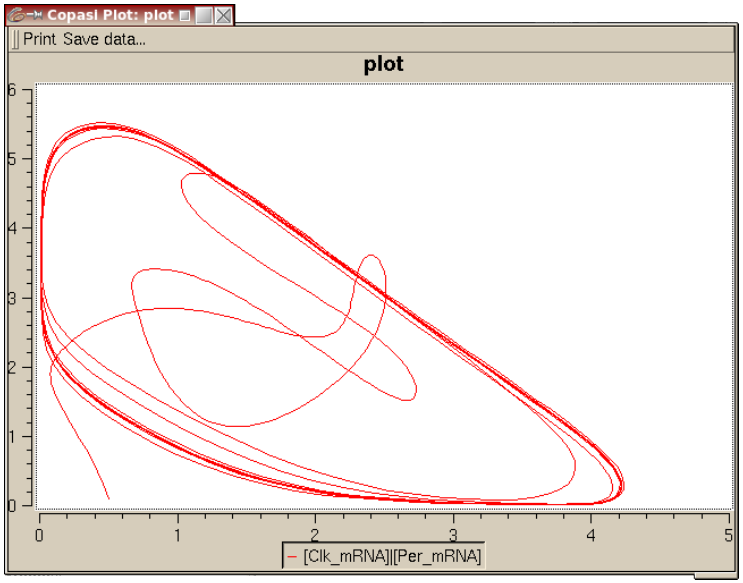


Complex post-processing

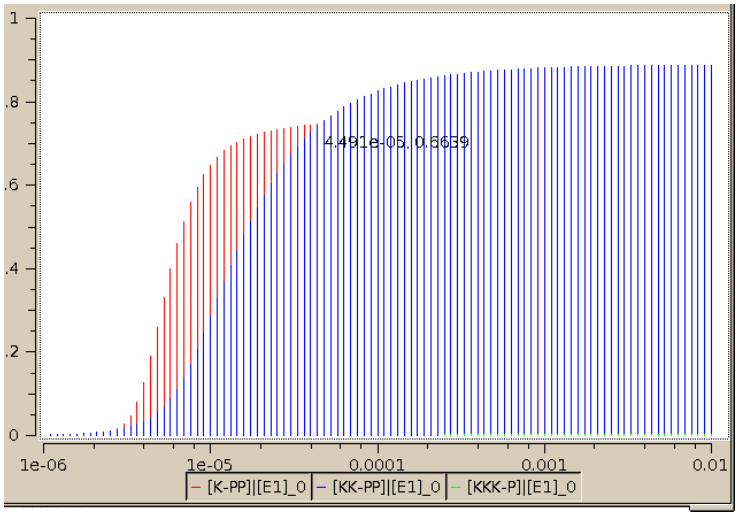
Edelstein et al 1996 (BIOMD0000000002)



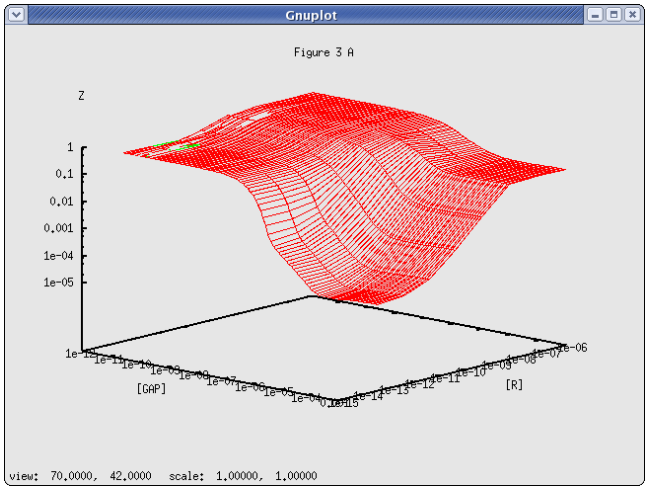
Ueda, Hagiwara, Kitanol 2001 (BIOMD0000000022)



Huang & Ferrell (BIOMD0000000009)



Bornheimer et al 2004 (BIOMD0000000086)



Minimum Information About a Simulation Experiment (MIASE)

MIASE aims at describing the information needed to run and repeat a numerical simulation experiment derived from a given quantitative model.

What information do-we need to achieve that goal?

- Information about the models to use
- Information about the simulation task to perform
- Information about result processing

More information at:

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

https://sourceforge.net/projects/miase

Email Archive: miase-discuss (read-only)

2008:	Jan (1)	Feb (35)	Mar (41)	Apr (4)	May (19)	Jun (26)	Jul (3)	Aug (2)	Sep (2)	Oct (1)	Nov	Dec (3)
2009:	Jan (49)	Feb (15)	Mar (17)	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec

miase-discuss | Ultimate | Show 25 | Change View

Topic	Topic Starter	Thread Posts	Last Post
[Miase-discuss] changes to example files XPath strings	Richard Adams <radams@st...>	2	2009-03-27 10:48
[Miase-discuss] sedml schema updated	Dagmar Köhn <dagmar@eb...>	1	2009-03-22 18:27
[Miase-discuss] listOfModels query	Richard Adams <radams@st...>	5	2009-03-22 23:54
[Miase-discuss] miase file content - manifest?	Richard Adams <radams@st...>	4	2009-03-06 12:03
Re: [Miase-discuss] SBML/sedML Patent Issue	Nicolas Le novère <lenov@eb...>	5	2009-03-04 21:06
[Miase-discuss] sed-ml logo	Dagmar Köhn <dagmar@eb...>	1	2009-02-09 15:07
[Miase-discuss] schema changes	Dagmar Köhn <dagmar@eb...>	3	2009-01-31 04:42