What it is
What it could be

What happened to Biology at the end of XXth century?

Annu. Rev. Genomics Hum. Genet. 2001. 2:343–72
Copyright © 2001 by Annual Reviews. All rights reserved

A New Approach to Decoding Life:
Systems Biology

Trey Ideker[1,2], Timothy Galitski[1], and Leroy Hood[1,2,3,4,5]
Institute for Systems Biology[1], Seattle, Washington 98105; Departments of

New Generation Computing, 18(2000)199-216
Ohmsha, Ltd. and Springer-Verlag

Invited Paper

Perspectives on Systems Biology

Hiroaki KITANO
Sony Computer Science Laboratories, Inc.

©Ohmsha, Lt

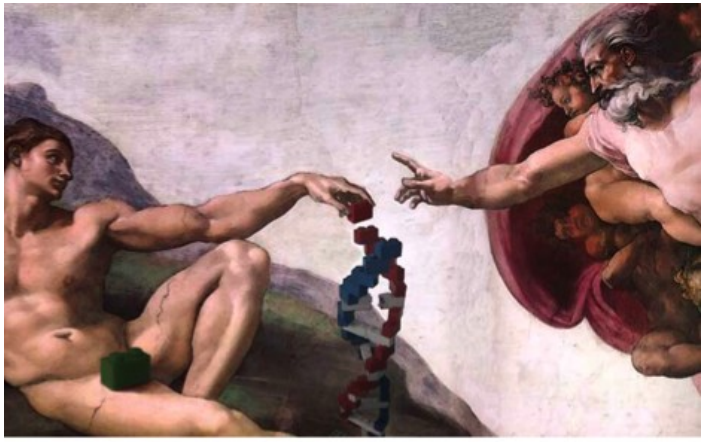
What happened to biology at the end of XXth century?

RESEARCH ARTICLE

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,[1] John I. Glass,[1] Carole Lartigue,[1] Vladimir N. Noskov,[1] Ray-Yuan Chuang,[1] Mikkel A. Algire,[1] Gwynedd A. Benders,[2] Michael G. Montague,[1] Li Ma,[1] Monzia M. Moodie,[1] Chuck Merryman,[1] Sanjay Vashee,[1] Radha Krishnakumar,[1] Nacyra Assad-Garcia,[1] Cynthia Andrews-Pfannkoch,[1] Evgeniya A. Denisova,[1] Lei Young,[1] Zhi-Qing Qi,[1] Thomas H. Segall-Shapiro,[1] Christopher H. Calvey,[1] Prashanth P. Parmar,[1] Clyde A. Hutchison III,[2] Hamilton O. Smith,[2] J. Craig Venter[1,2]*

2 JULY 2010 VOL 329 SCIENCE www.sciencemag.org



EXTREME GENETIC ENGINEERING

An Introduction to Synthetic Biology

etc group

january 2007

Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors

Cell

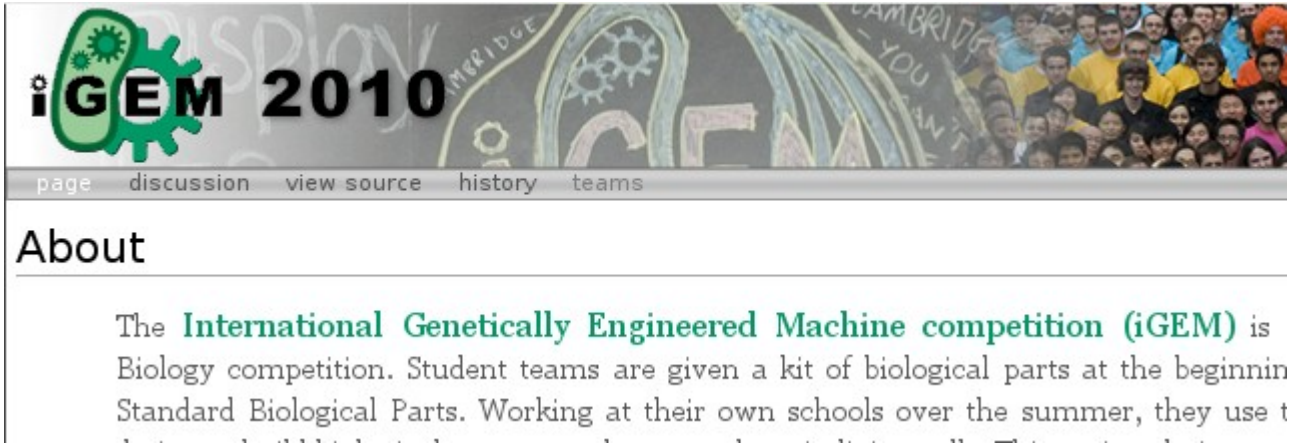
Kazutoshi Takahashi[1] and Shinya Yamanaka[1,2,*]
[1] Department of Stem Cell Biology, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan
[2] CREST, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan
*Contact: yamanaka@frontier.kyoto-u.ac.jp
DOI 10.1016/j.cell.2006.07.024

A synthetic oscillatory network of transcriptional regulators

Michael B. Elowitz & Stanislas Leibler

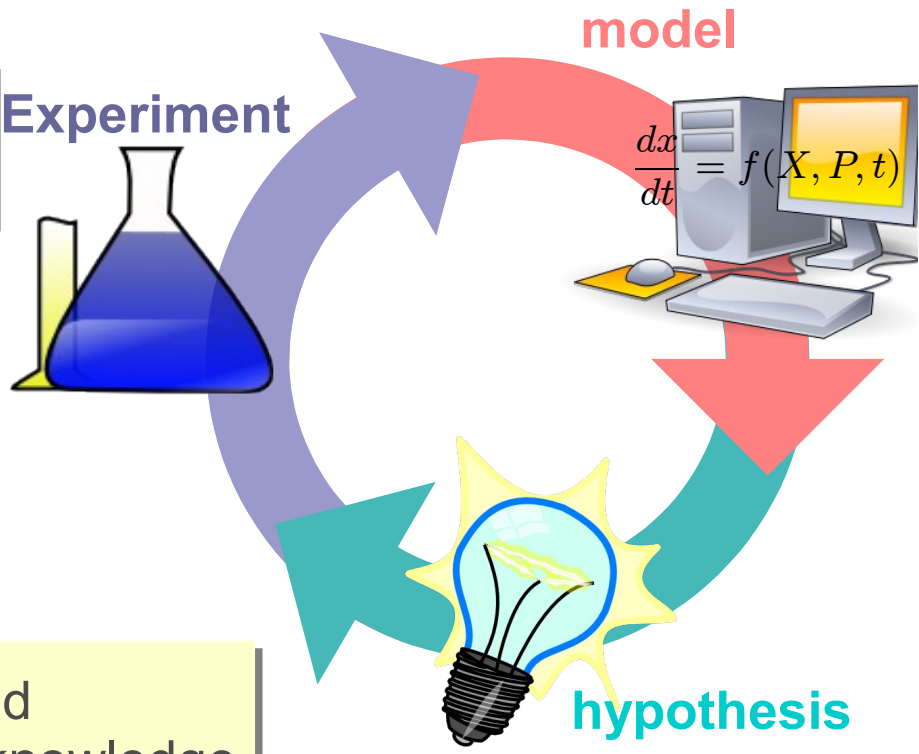
Departments of Molecular Biology and Physics, Princeton University, Princeton, New Jersey 08544, USA

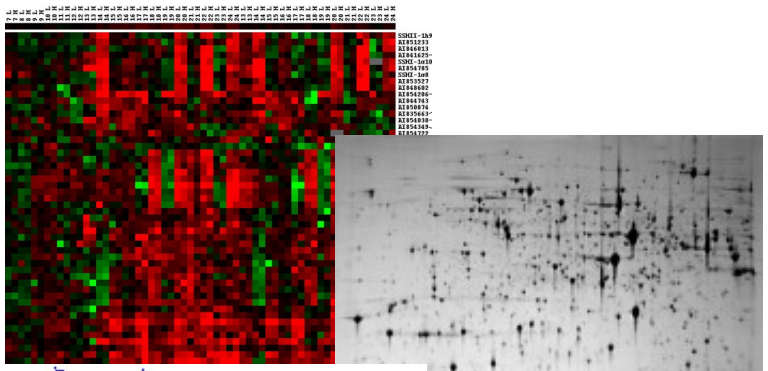
NATURE | VOL 403 | 20 JANUARY 2000 | www.nature.com

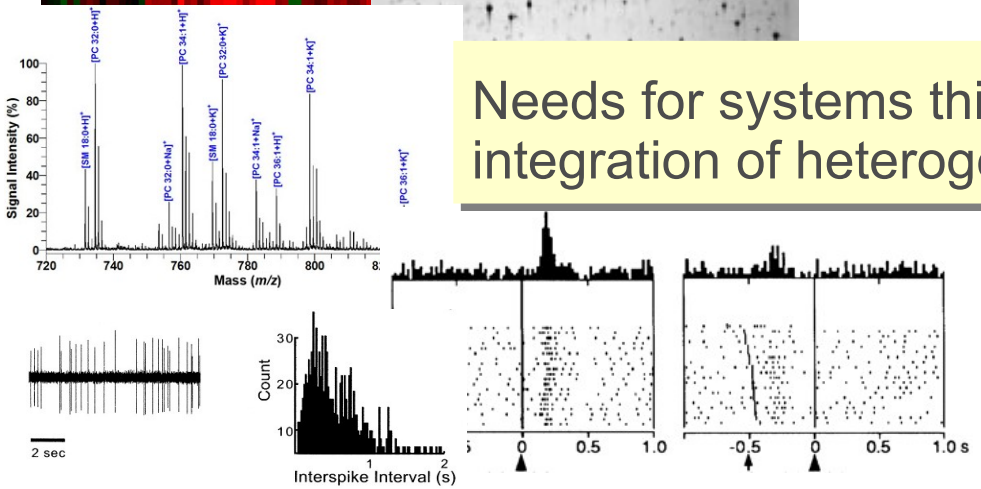


New way of doing biomedical research

Needs for interplay between models and reality tests

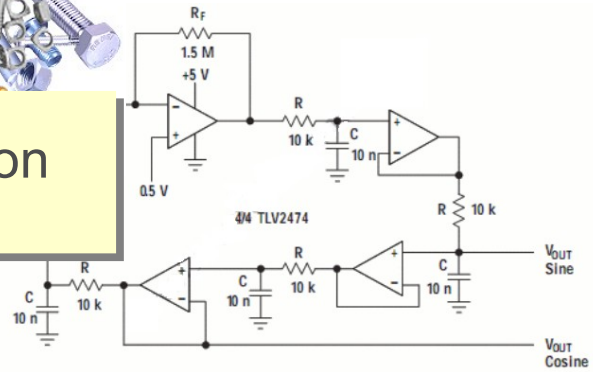






Needs for systems thinking and integration of heterogeneous knowledge

Needs for cooperation and standardisation



Computational modelling left the niches

- Metabolic networks Fung et al. A synthetic gene-metabolic oscillator. ***Nature*** 2005; Herrgård et al. A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. ***Nat Biotechnol*** 2008

- Signalling pathways Bray et al. Receptor clustering as a cellular mechanism to control sensitivity. ***Nature*** 1998; Bhalla ad Iyengar. Emergent properties of signaling pathways. ***Science*** 1998; Schoeberl et al. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. ***Nat Biotechnol*** 2002; Hoffmann et. The IκB-NF-κB signaling module: temporal control and selective gene activation. ***Science*** 2002; Smith et al. Systems analysis of Ran transport. ***Science*** 2002; Bhalla et al. MAP kinase phosphatase as a locus of flexibility in a mitogen-activated protein kinase signaling network. ***Science*** 2002; Nelson et al. Oscillations in NF-κB Signaling Control the Dynamics of Gene Expression. ***Science*** 2004; Werner et al. Stimulus specificity of gene expression programs determined by temporal control of IKK activity. ***Science*** 2005; Sasagawa et al. Prediction and validation of the distinct dynamics of transient and sustained ERK activation. ***Nat Cell Biol*** 2005; Basak et al. A fourth IkappaB protein within the NF-κB signaling module. ***Cell*** 2007; McLean et al. Cross-talk and decision making in MAP kinase pathways. ***Nat Genet*** 2007; Ashall et al. Pulsatile Stimulation Determines Timing and Specificity of NF-κB-Dependent Transcription. ***Science*** 2009; Becker et al. Covering a broad dynamic range: information processing at the erythropoietin receptor. ***Science*** 2010

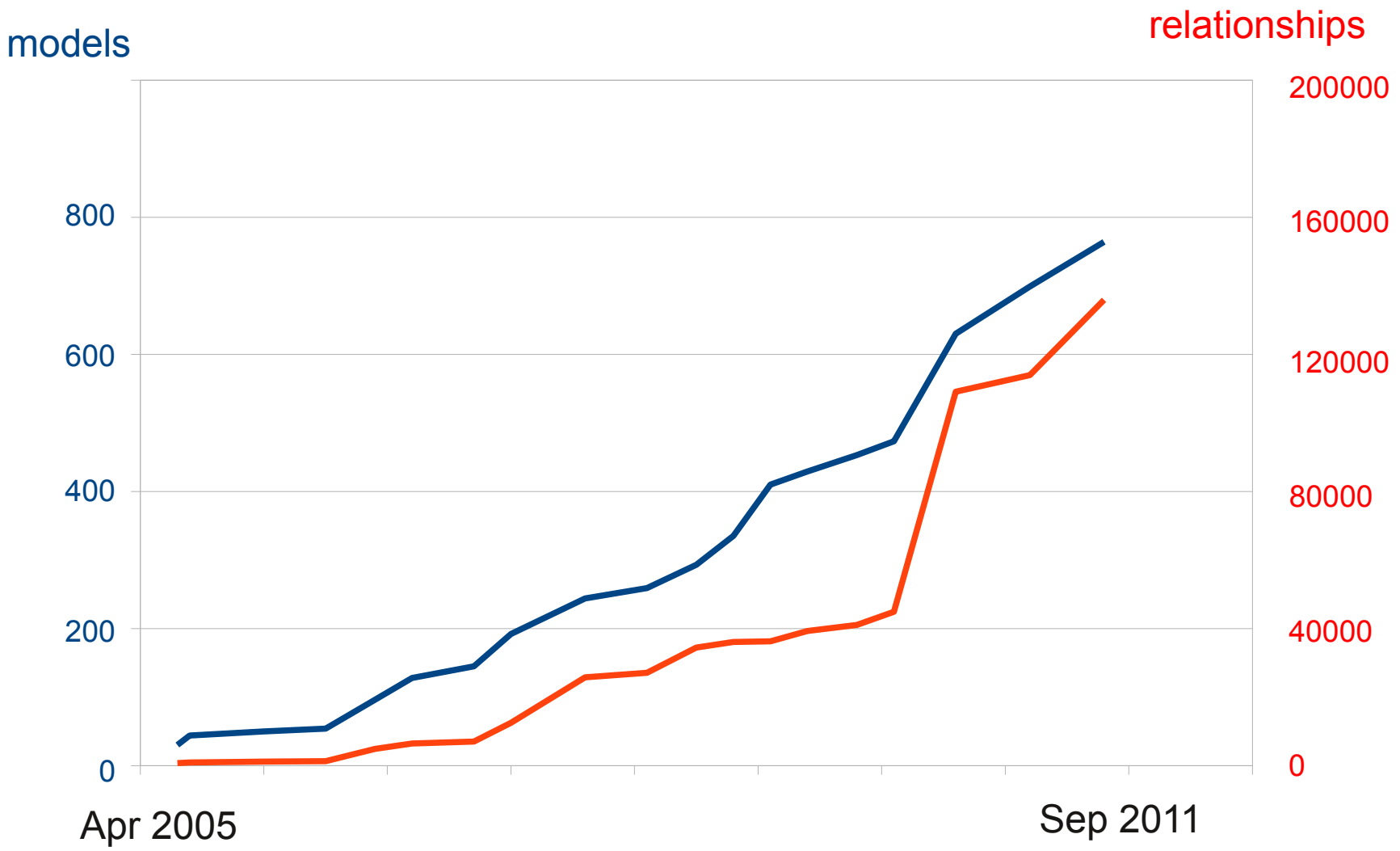
- Gene regulatory networks McAdams and Shapiro. Circuit simulation of genetic networks. ***Science*** 1995; Yue et al. Genomic cis-regulatory logic: Experimental and computational analysis of a sea urchin gene. ***Science*** 1998; Von Dassow et al. The segment polarity network is a robust developmental module. ***Nature*** 2000; Elowitz and Leibler. A synthetic oscillatory network of transcriptional regulators. ***Nature*** 2000; Shen-Orr et al, Network motifs in the transcriptional regulation network of Escherichia coli. ***Nat Genet*** 2002; Yao et al. A bistable Rb-E2F switch underlies the restriction point. ***Nat Cell Biol*** 2008; Friedland. Synthetic gene networks that count. ***Science*** 2009

- Pharmacometrics models Labrijn et al. Therapeutic IgG4 antibodies engage in Fab-arm exchange with endogenous human IgG4 in vivo. ***Nat Biotechnol*** 2009

- Physiological models Noble. Modeling the heart from genes to cells to the whole organ. ***Science*** 2002; Izhikevich and Edelman. Large-scale model of mammalian thalamocortical systems. ***PNAS*** 2008

- Infectious diseases Perelson et al. HIV-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time. ***Science*** 1996; Nowak. Population dynamics of immune responses to persistent viruses. ***Science*** 1996; Neumann et al. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. ***Science*** 1998

Computational models on the rise



BioModels Database growth since its creation

http://www.ebi.ac.uk/biomodels/

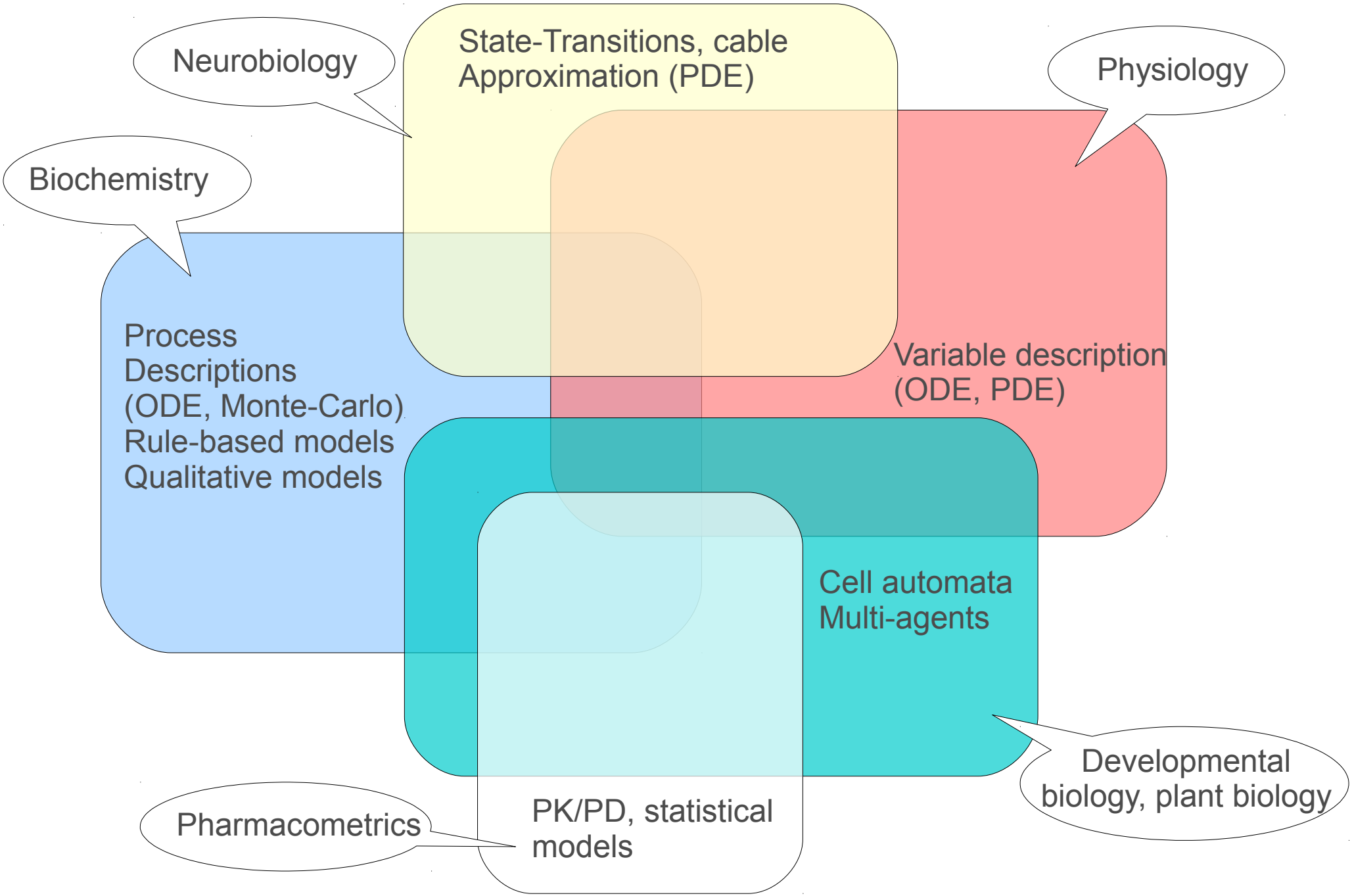
Interest from new stakeholders

- “Biologists”: computational models look “useful”, “serious”
- Publishers: computational are respectable, and can be published in high profile journals
- Funding agencies: Models could help with the major challenges (read “science that can be sold to citizen/electors”): Health, Food, Energy...
- Industries: Models could help with the major challenges (read “new opportunities to make money”): Pharmas, crops, biofuels ...

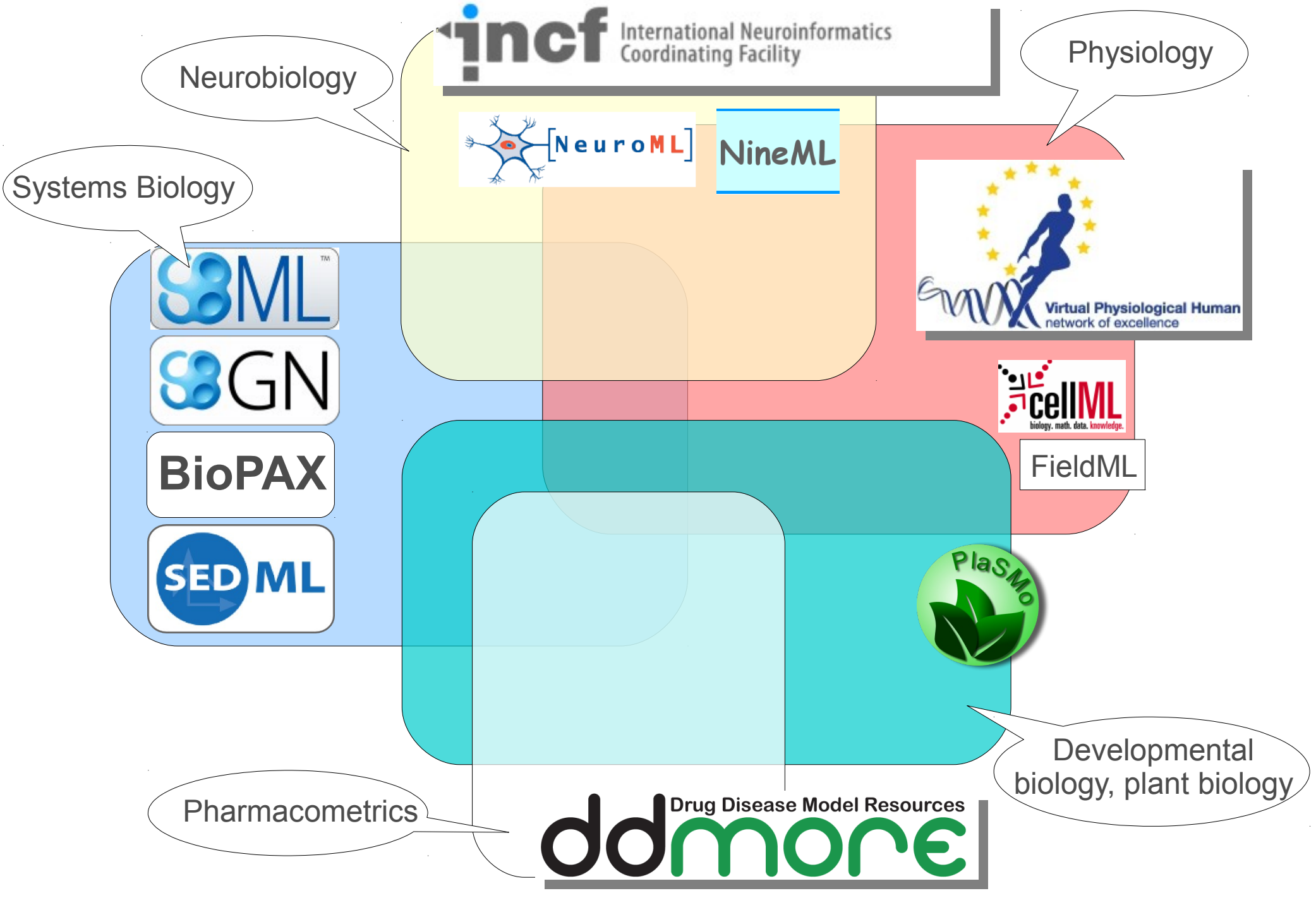
The matrix of standards for M&S in Sys Bio

	Model descriptions	Simulations and analysis	Numerical results
Minimal requirements	MIRIAM	MIASE	
Data-models	SBML, SBGN	SED ML	NuML?
Terminologies	SBO	KISAO	TEDDY

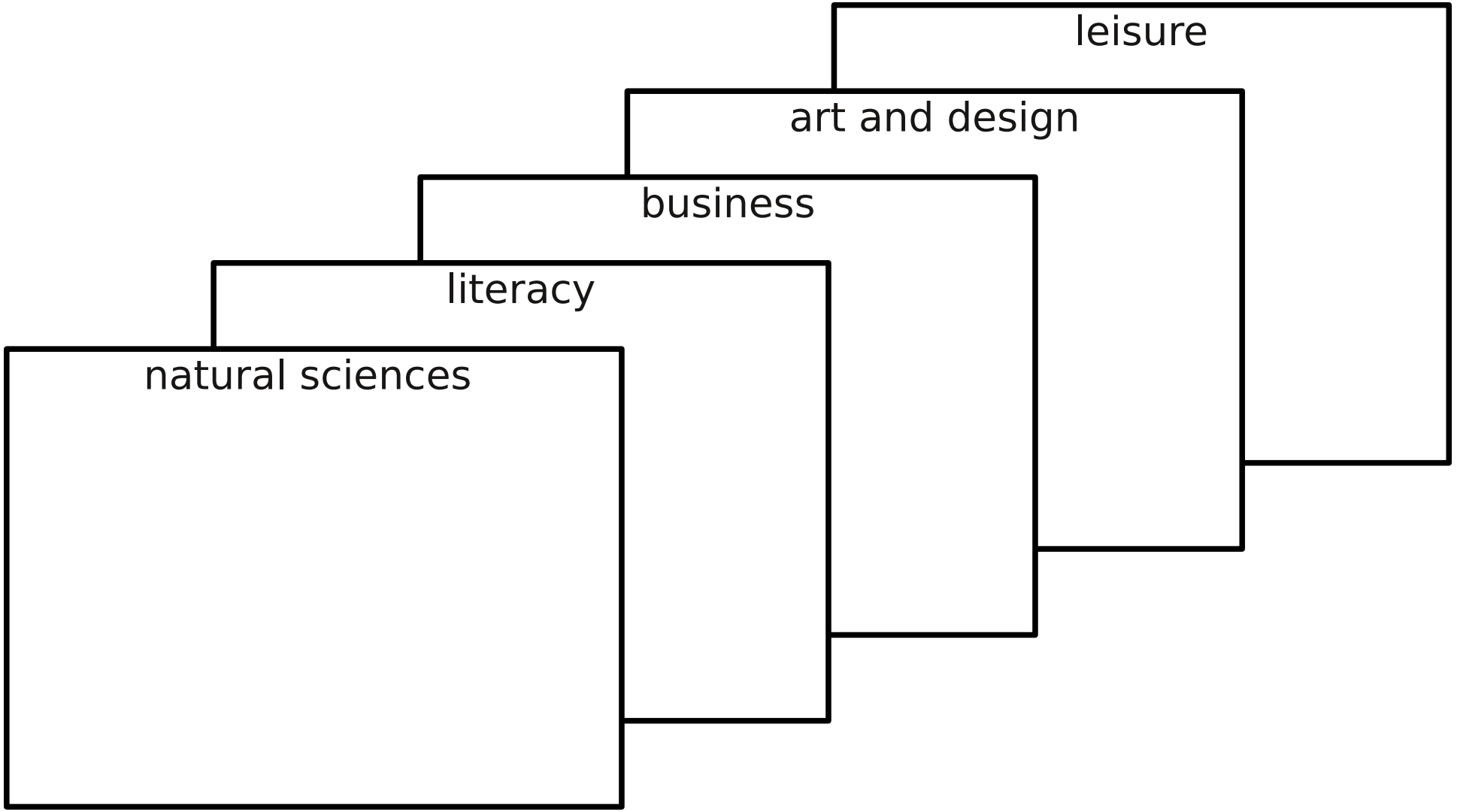
Dimension 3: Covering alternative modelling approaches



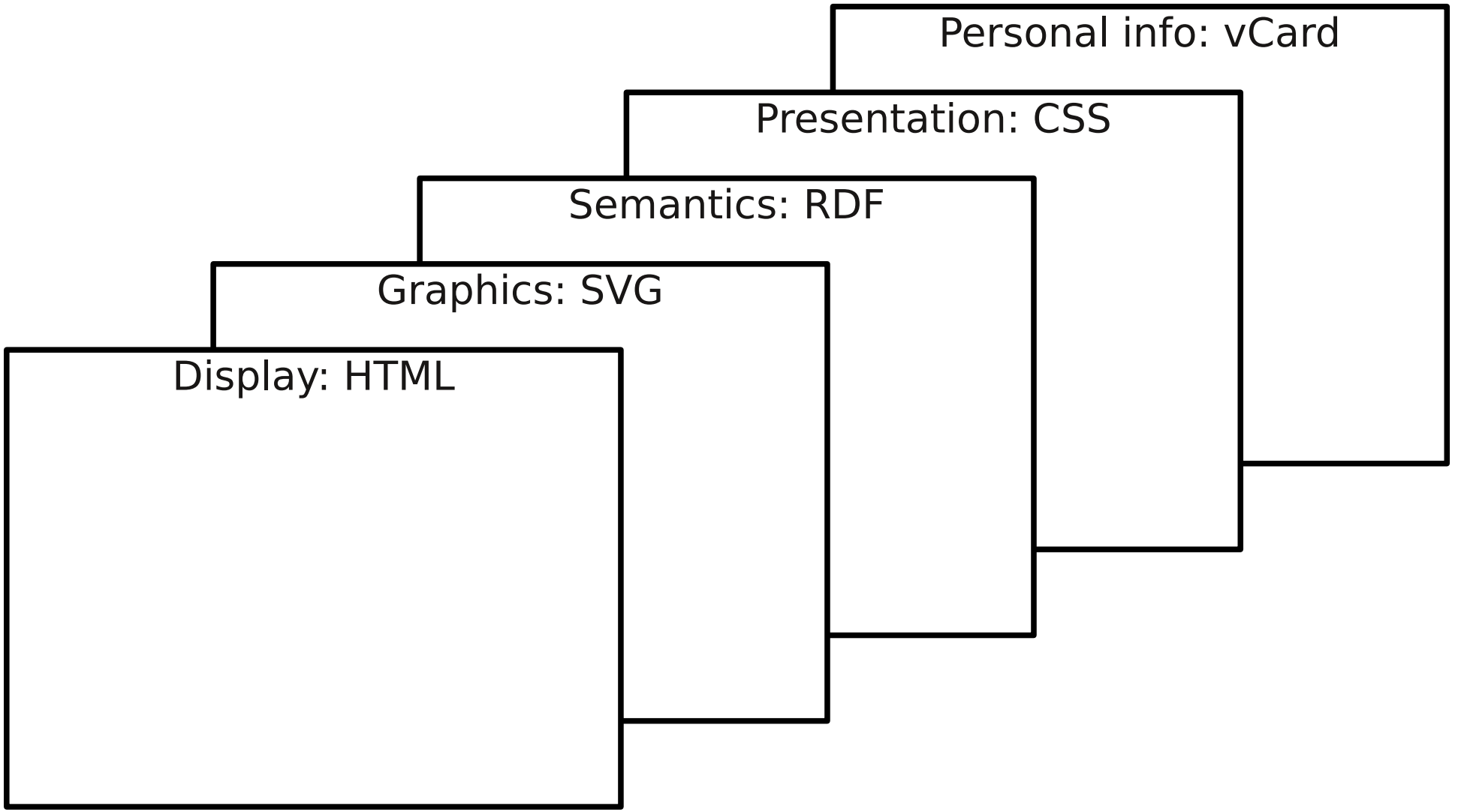
Parallel and redundant efforts



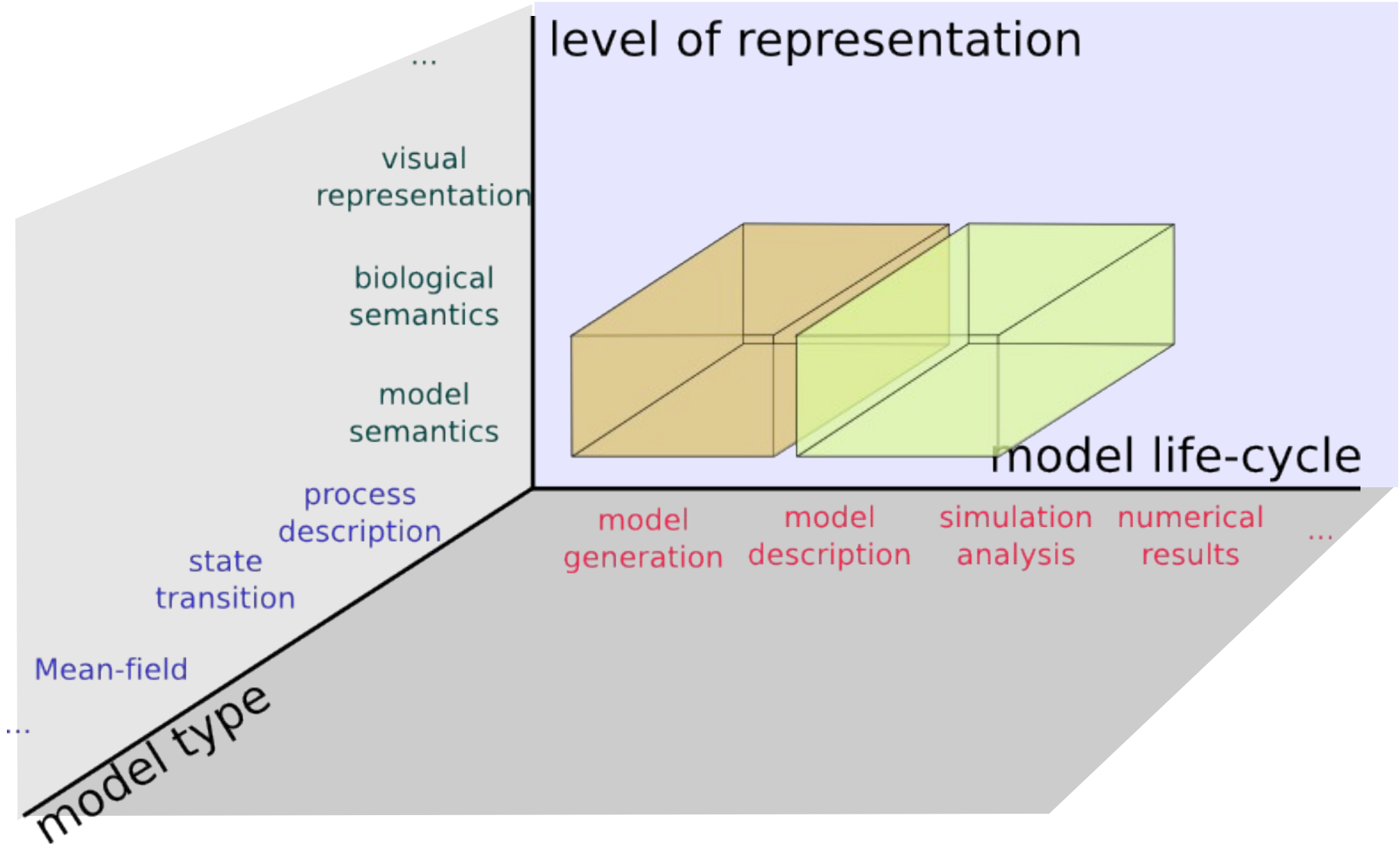
What if the world-wide web was built like this?



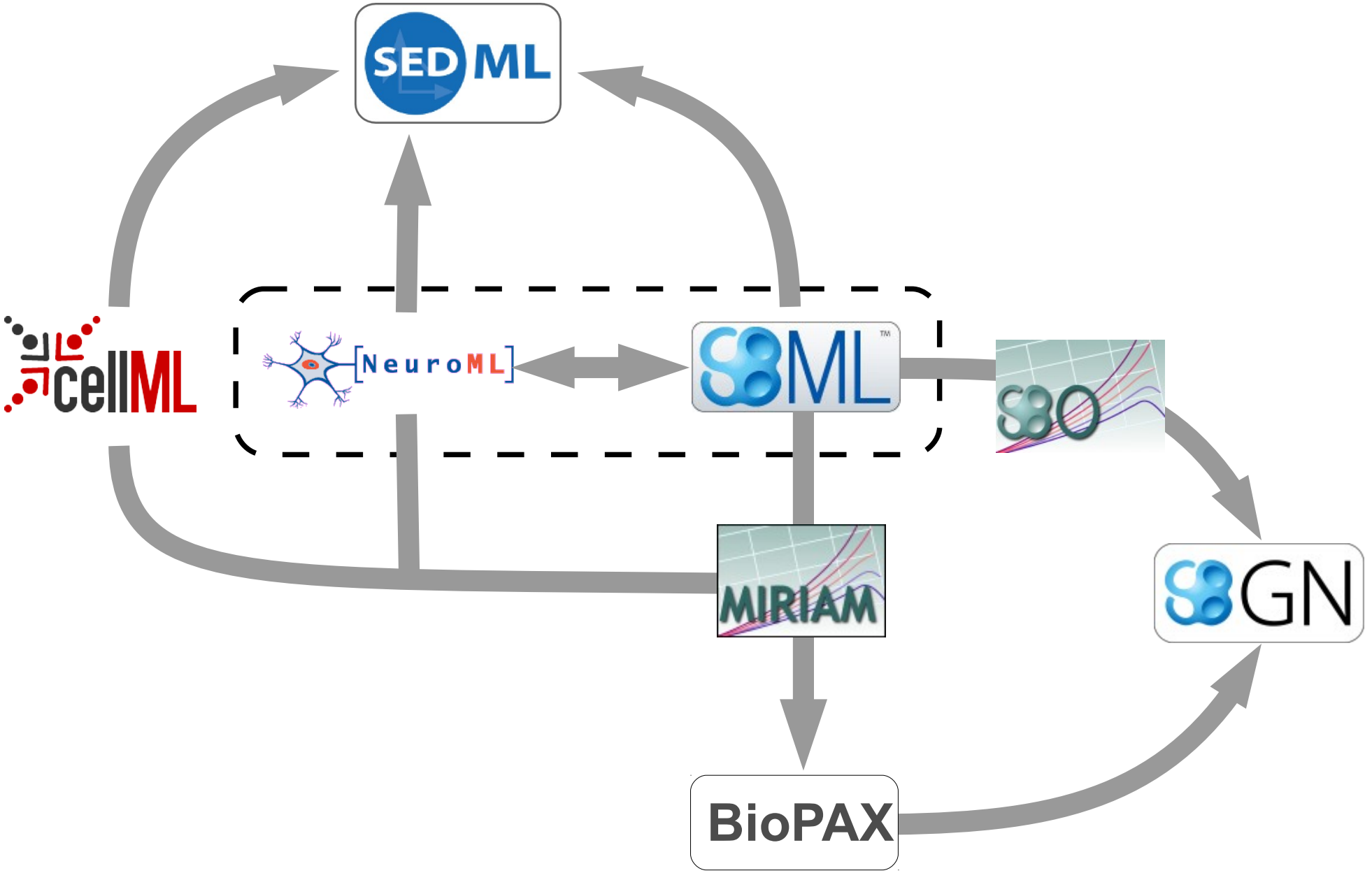
The correct way to do it



Covering the entire modeling in the life-sciences



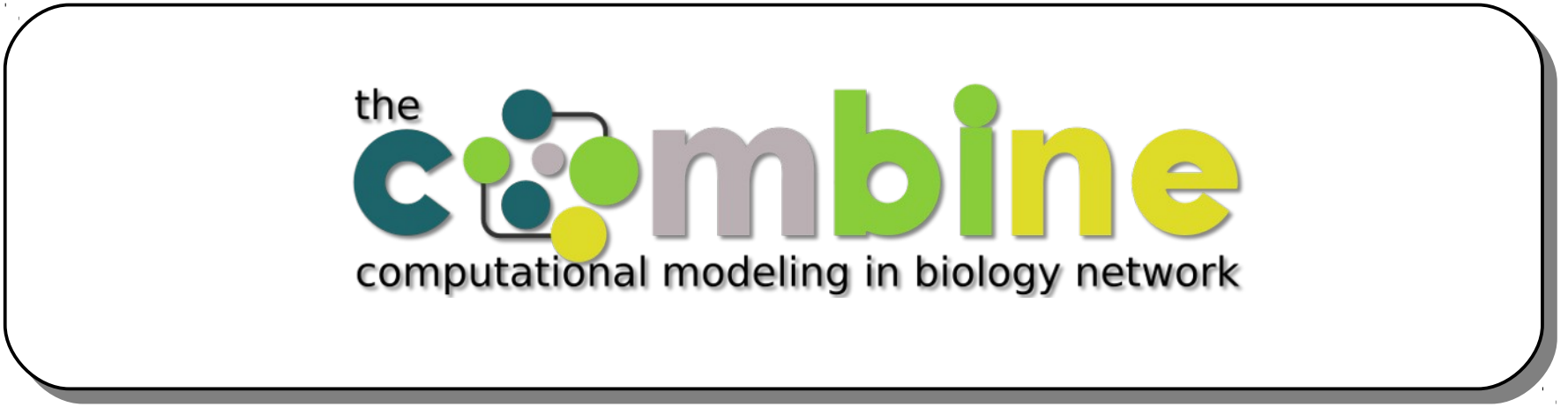
Existing standards interoperability



Threats to the whole enterprise

- Current efforts are largely dependent on key people. Their disengagement means stalling or disaggregation.
- Current funding structure is fragile. Many different grants, sometimes only supporting meetings, none of them infrastructure rolling funding, often tied to individuals.
- Current efforts, being developed under the umbrellas of specific institutions are not immune against intellectual property claims that would harm the community.
- Existing standards are developed with very different approaches, quality checks, and are based on completely different assumptions (e.g. implicit knowledge versus explicit mathematics).
- APIs, converters etc. need industry-grade support, incompatible with standard academic usages and possibilities

Overarching standardisation structure



The “WorldWide Web consortium” of modelling in biology

http://co.mbine.org/

Mission 1: Coordinating the standards

- **CORE STANDARDS:** Efforts fulfilling COMBINE criteria and aiming at following COMBINE rules and interoperate with other COMBINE standards

- **ASSOCIATED EFFORTS:** Standards that are not representation formats, but aiming at enrich or bridge the core standards

- **RELATED EFFORTS:** Formats developed by other communities, that complement or interoperate with COMBINE formats, and that we would like to see joining COMBINE or collaborating closely to COMBINE

Current core COMBINE standards

- Model semantics, Model structure, Process description:

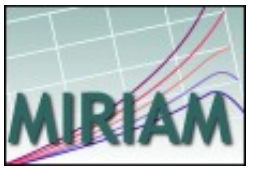
- Models semantics, Simulation and Analysis:

- Biological semantics, Model structure, Process description, Entity relationships:

- Visual representation, Model structure, Process description, Entity Relationships, Activity flow:

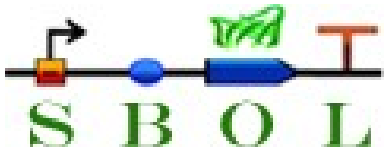
Current associated standardization efforts

- Concept and data reference:
  - MIRIAM Registry
  - Identifiers.org URIs



- Terminologies:
  - Modeling
  - Simulation
  - :

Related standardization efforts



PSI-ML

FieldML

NuML?

COMBINE does NOT aim to take over the development of the standard formats, but help coordinating and supporting this process

Mission 2: Coordinating meetings

- Annual COMBINE meetings
  - COMBINE 2010: October 6–9, Edinburgh, **81 attendees**
  - COMBINE 2011: September 3-7, Heidelberg, 82 registrations
  - COMBINE 2012: End of summer, beginning autumn , Toronto
  - COMBINE 2013: Date unknown, location unknown
    http://www.surveymonkey.com/s/combine-harmony-hosting-interest
- The Hackathons on Resources for Modeling in Biology
  - HARMONY 2011: April 18-22, New-York City, **59 attendees**
  - HARMONY 2012; Date unknown, location unknown
    http://www.surveymonkey.com/s/combine-harmony-hosting-interest

Mission 3: Developing Standard Operating Procedures

- Technical requirements
  - Who format covers what portion of the modeling space
  - Which technical solutions exist and must be used by the formats
  - How do formats interface
  - How to specify and document formats
  - ...
- Governance
  - How to initiate and maintain standardisation efforts
  - How to communicate with users and developers
  - How to develop a democratic and robust governance
  - ...

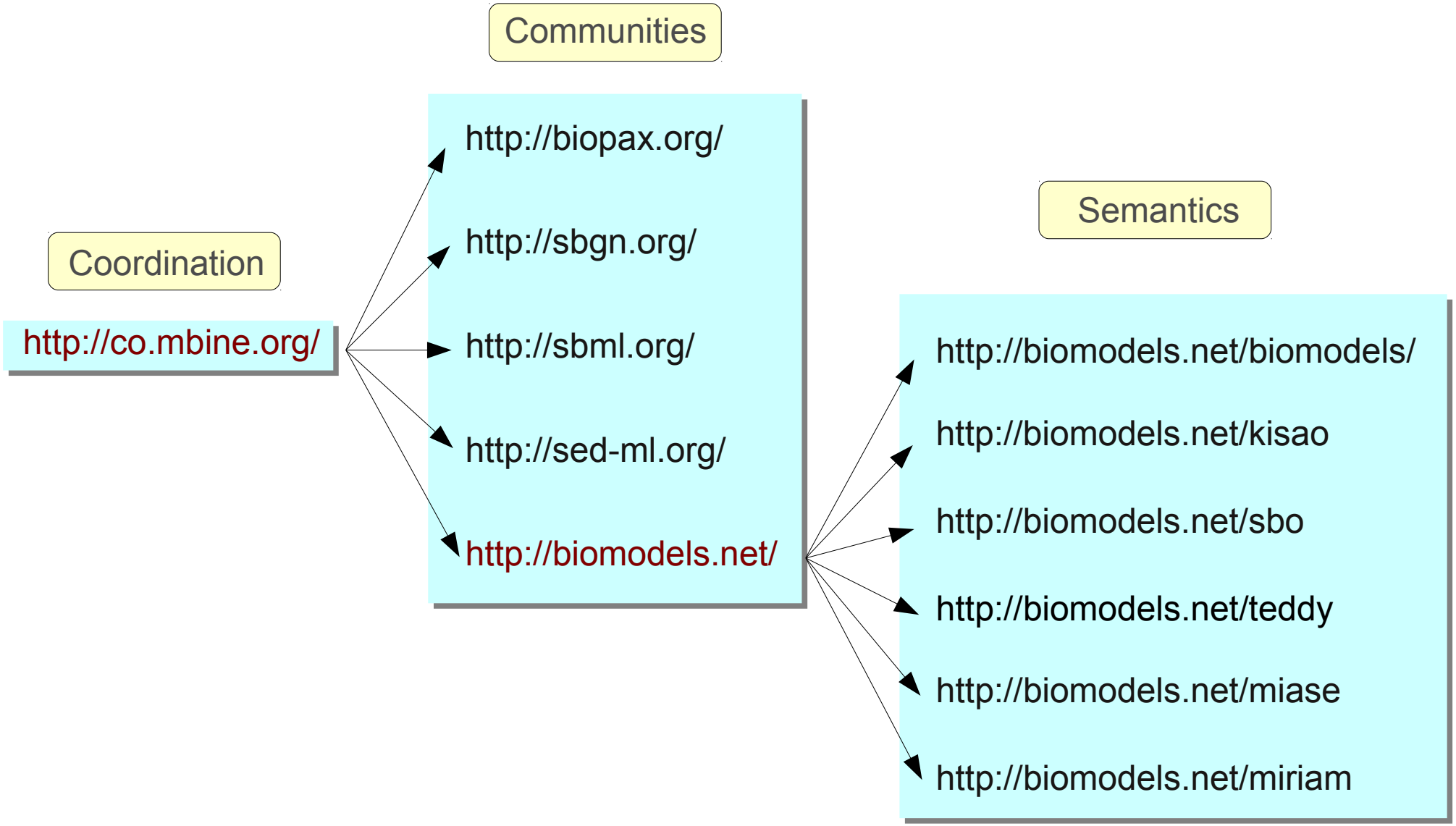
First tentative SOP: guidelines to develop a core COMBINE standard

- List the new development as a related standardization effort
- Join COMBINE community and attend meetings
- Comply with COMBINE criteria
  - Must cover aspects of modeling significantly different from the existing set of COMBINE standards
  - Must be described inprecise technical specification documents and formal specification languages
  - Specifications and other materials must be publicly available free of charge to everyone and be unencumbered by licensing restrictions
  - Development must be open. The entire COMBINE community must be able to participate without exclusion
  - Must be developed and used by more than a single team or organisation.
  - Development process must be led by democratically elected editorial boards
  - Mature software support must exist, including standard API implementations, and possibly validation tools
  - Development must be stable and active
- Decision by the coordinators (alt: vote of the community?)

Mission 4: Recognised voice

- COMBINE aims to become a “standardisation” body
  - This means a quality label. A “COMBINE standard” is a guarantee of stability, community endorsement, support etc.
  - COMBINE production can be used in SOPs at other organisations
  - COMBINE must be an actor on par with FGED, PSI, INCF etc.
- Single point of contact with user organisations including Industry
  - Tool developers (General platforms or specific tools)
  - Publishers
  - Pharmaceutical industry
- A point of contact for funding bodies
- A point of contact for legal entities, e.g. government and regulatory bodies

Where to find more information?



Acknowledgements

SBML editors: Frank Bergmann, *Andrew Finney, Stefan Hoops*, Michael Hucka, *Nicolas Le Novère,* Sarah Keating, Chris Myers, *Sven Sahle*, *Herbert Sauro*, Jim Schaff, Lucian Smith, *Darren Wilkinson*

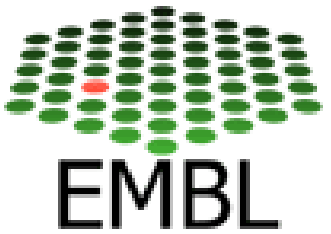
SBGN editors: Emek Demir, Nicolas Le Novère, Huaiyu Mi, Stuart Moodie, *Falk Schreiber*, *Anatoly Sorokin*, Alice Villéger

BioPAX editors: Peter D'Eustachio, Oliver Ruebenacker, Andrea Splendiani

SED-ML editors: Richard Adams, Franck Bergmann, Nicolas Le Novère, Andrew Miller, David Nickerson, Dagmar Waltemath

Metadata: Mélanie Courtot, Nick Juty, Camille Laibe, Anna Zhukova

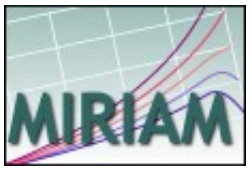
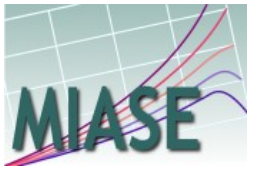
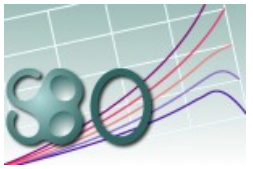
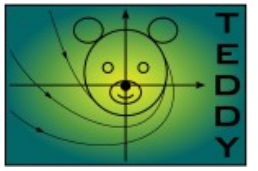
The whole community of Computational Systems Biology



Is the matrix of standards complete?

	Model descriptions	Simulations and analysis	Numerical results
Minimal requirements	MIRIAM	MIASE	
Data-models	SBML, SBGN, BioPAX	SED ML	NuML?
Terminologies	SBO	KISAO	TEDDY

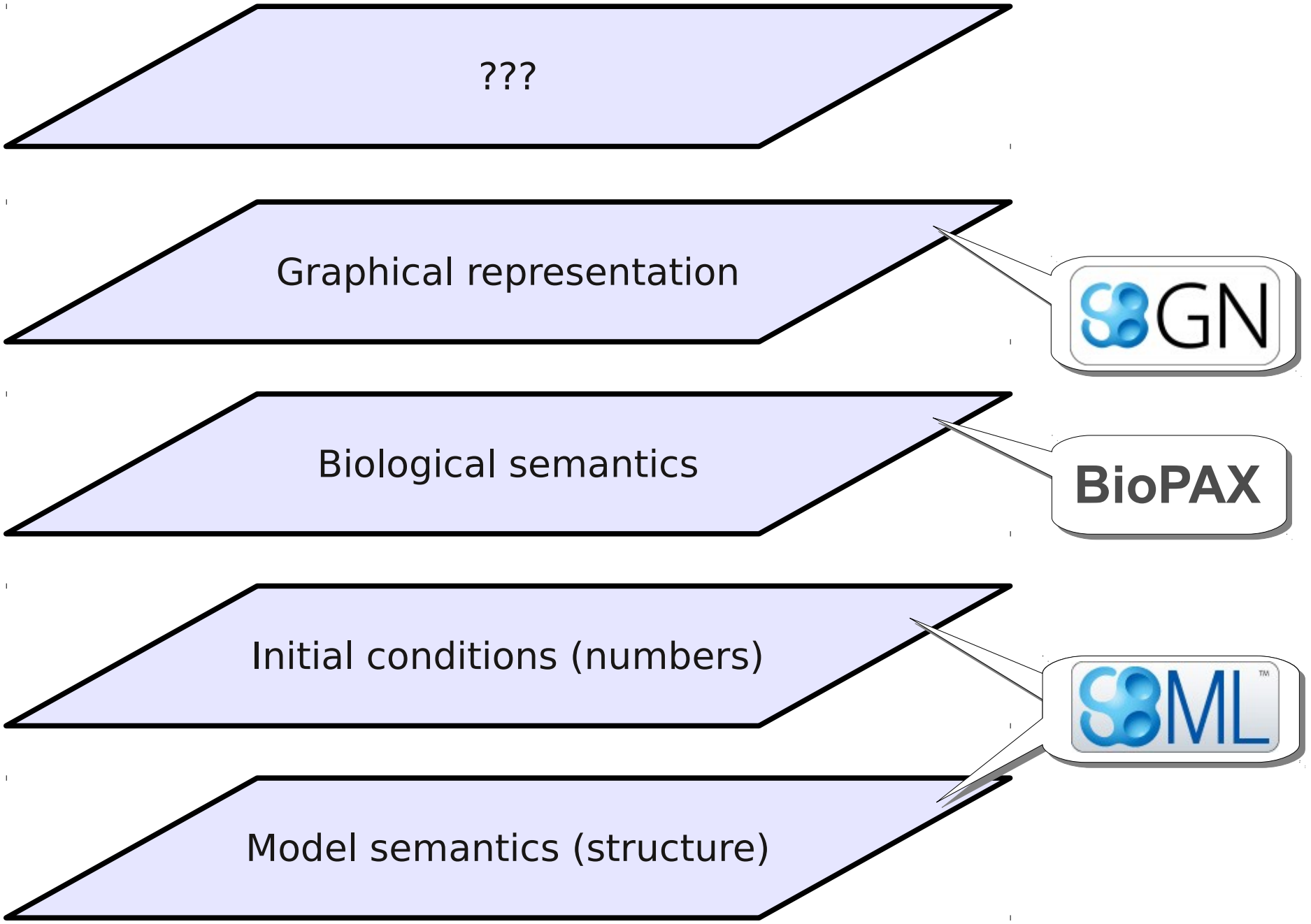
Is the matrix of standards complete?

	Model descriptions	Simulations and analysis	Numerical results
Minimal requirements			
Data-models			NuML?
Terminologies			

Dimension 1: Covering the entire model life-cycle

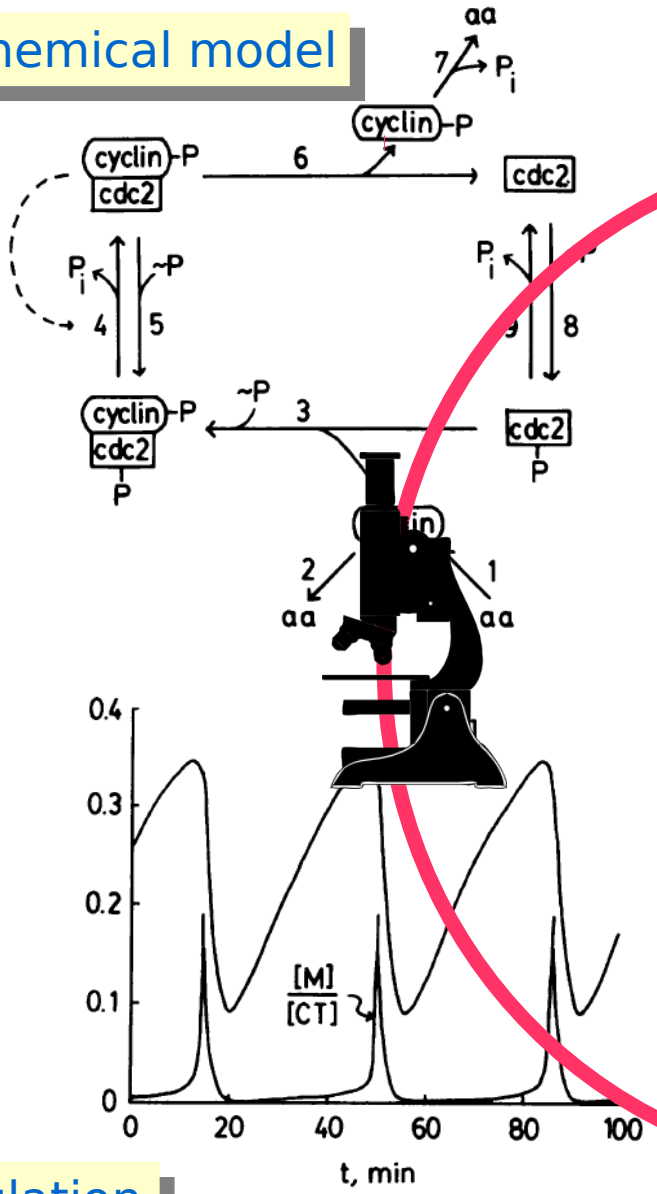
Model generation	Model structure	Parametrisation	Simulations and analysis	Numerical results
?	SBML	?	SED ML	NUML?

Dimension 2: Representing the levels of discourse



The models I am talking about

biochemical model



simulation

mathematical model

$$d[C2]/dt = k_6[M] - k_8[\sim P][C2] + k_9[CP]$$
$$d[CP]/dt = -k_3[CP][Y] + k_8[\sim P][C2] - k_9[CP]$$
$$d[pM]/dt = k_3[CP][Y] - [pM]F([M]) + k_5[\sim P][M]$$
$$d[M]/dt = [pM]F([M]) - k_5[\sim P][M] - k_6[M]$$
$$d[Y]/dt = k_1[aa] - k_2[Y] - k_3[CP][Y]$$
$$d[YP]/dt = k_6[M] - k_7[YP]$$

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

computational model

Tyson et al (1991) PNAS 88(1):7328-32