

Evolution of Computational Models in Systems Biology

what's the matter – what's there – what's next

MARTIN SCHARM

Department of Systems Biology & Bioinformatics, University of Rostock

<http://sems.uni-rostock.de>



Research Stay in Manchester, 2015



CombineArchive Toolkit

Sharing in silico experiments is essential for the advance of research in computational biology. The COMBINE archive is a digital container format to ease the management of numerous files and to enable the exchange of reproducible modeling results. We developed the CombineArchive Toolkit, consisting of a library, a web interface and a desktop application. It support scientists in creating, exploring, modifying and sharing COMBINE archives.



BivEs

Armed with our method for difference detection, BivEs is able to detect and communicate the differences in computational models. The differences are exported in several machine- and human-readable formats, closely syntactic to be integrated in other tools.

BudHat

BudHat showcases how BivEs improves the understanding of a model's changes. BudHat calls BivEs for the comparison two versions of a computational model and displays the obtained results in the web browser.



Version Control for Computational Models

With thousands of models available, a framework to track the differences between models and their versions is essential to compare and combine models. Focusing on SBML and CellML, we developed an algorithm to accurately detect and describe differences between versions of a model with respect to (i) the model's encoding, (ii) the structure of biological networks, and (iii) mathematical expressions.



Ontology of Differences

Changes in model versions are manifold and appear on different layers. We develop an ontology of differences occurring in model versions. It will support researchers in analyzing differences, discussing typical changes, identifying major changes and proving statistics.

2MT

2MT is our web-based platform to demonstrate the capabilities of SEEMS-related tools. It exemplifies how our model management solutions can be used in existing tools.

Masynoms

Containing SBML- and CellML models, linked semantic annotations (e.g. from bio-ontologies), simulation diagnostics, graphical representations and other available types of model-related data, our graph database Masynoms can now be queried for complete simulation experiments.



More

- Our retrieval engine for models applies bio-formation retrieval techniques to retrieve relevant models from MASYNOMS. The proposed ranking and retrieval techniques focus on the processing of model meta-information.

de.NBI Infrastructure

We will provide data management and support for systems biology projects, with a focus on provenance and reproducibility of experimental and modeling studies. de.NBI-FY2010 is part of a large German Network for Bioinformatics Infrastructure. **WE ARE HIRING!**

SBGN-ED

SBGN is a markup language to describe models and exchange information about biological systems graphically. We will further develop methods and tools for SBGN-compliant visualization of model-related information. **WE ARE HIRING!**



combine

HARMONY, April 19-23, 2015
HARMONY is a hackathon-type meeting of the COMBINE Community with a focus on development of the standards, interoperability and infrastructure. Instead of general discussions or oral presentations, the time is devoted to hands-on hacking and interaction between people focused on practical development of software and standards. The HARMONY 2015 is hosted at the Leuzea Wiltberg and it is hosted by the groups of Falk Schreiber and Dagmar Waltemath.

Events

Volkswagen Stiftung

VW Summer School, March 9-13, 2015
During the 2015 Witold Cell summer school we aim to develop a standard-compliant, open version of the whole-cell model. Eleven tutors and 40 students will hack and code, model and simulate, layout and annotate the whole-cell model using openly available software and COMBINE standards. This event is funded by the Volkswagen Stiftung.



Workshop on Reproducible and Citable Data and Models

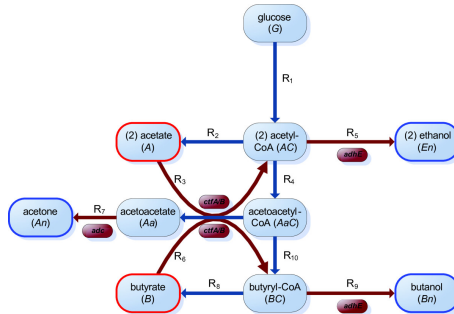
September 14-16, 2015
Computational biologists and experimentalists will learn about standards, citable data, about how to make scientific results sustainable, available through open repositories, and about how to find and reuse other people's works in a mixture of lectures and hands-on exercises. The workshop is funded by the ERASYS-APP program.

Models as graphs

The increasing diversity of model-related data that is necessary to perform a simulation study leads to new challenges in model storage. We developed a concept for graph-based storage of models and model-related data. Graphs reflect the model's structure much better, enable linking of model-related data on the storage layer, and allow an efficient search.



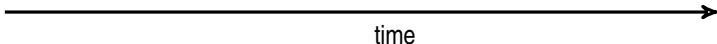
metabolic and gene regulation network model in *C. acetobutylicum*



Haus et. al. 2011

Papoutsakis ● → ● → ★

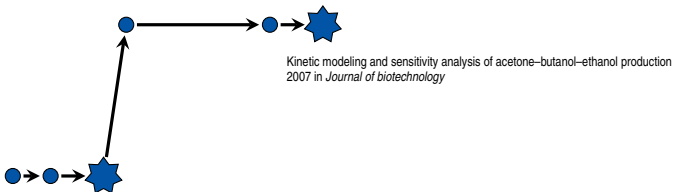
Equations and calculations for fermentations of butyric acid bacteria
1984 in *Biotechnology and bioengineering*



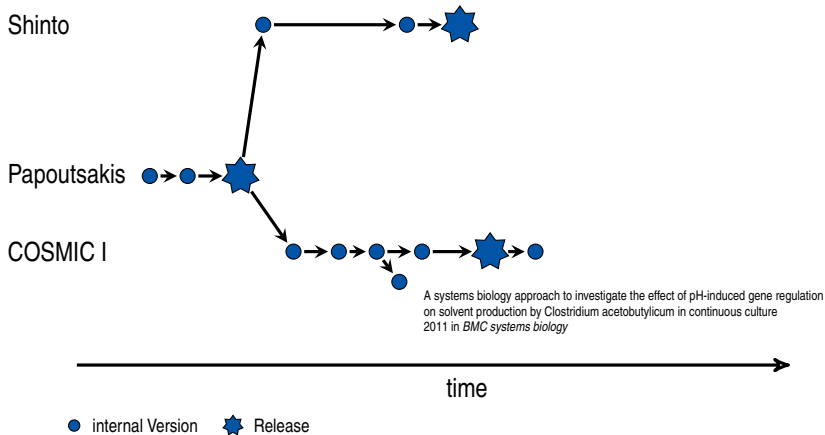
● internal Version ★ Release

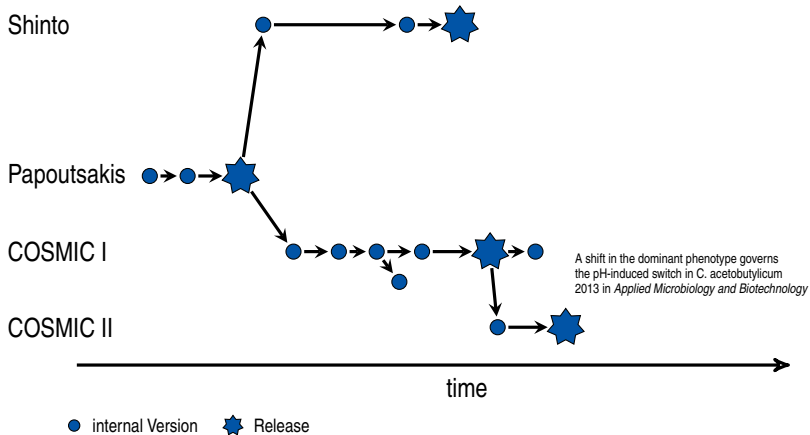
Shinto

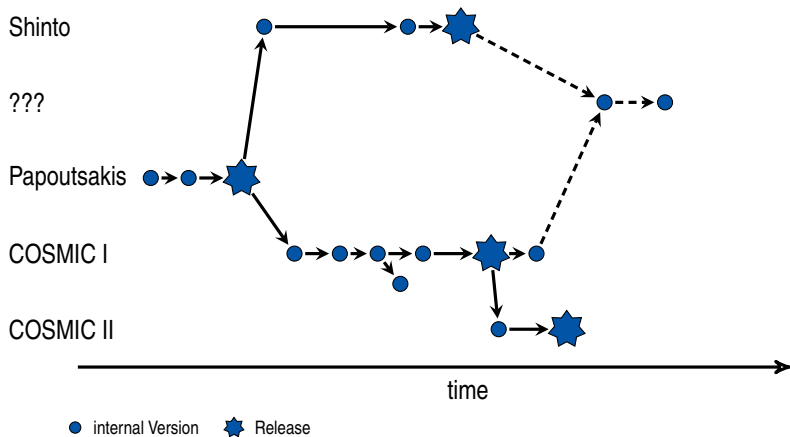
Papoutsakis

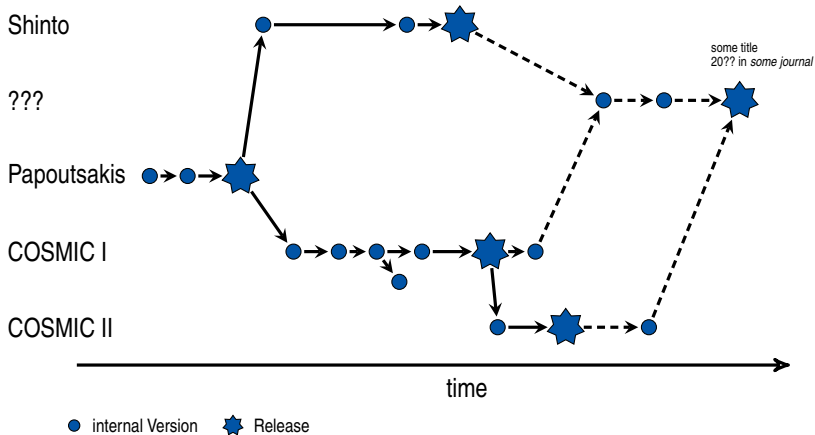


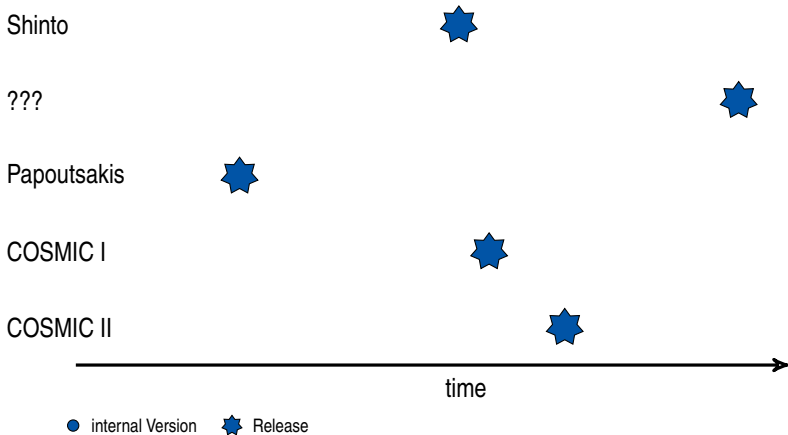
● internal Version ★ Release

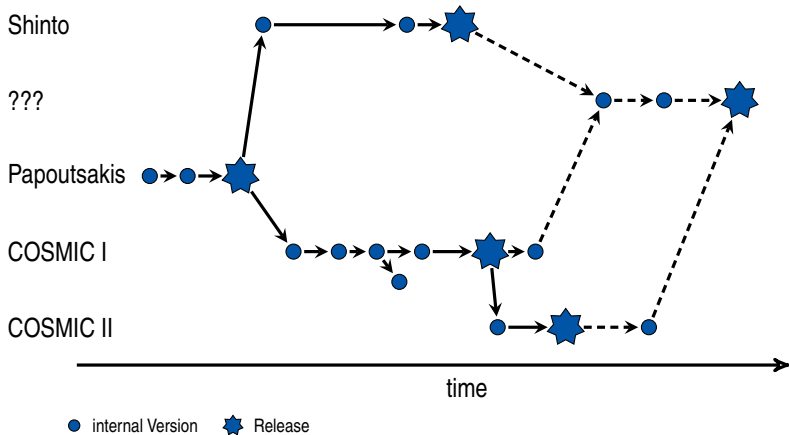


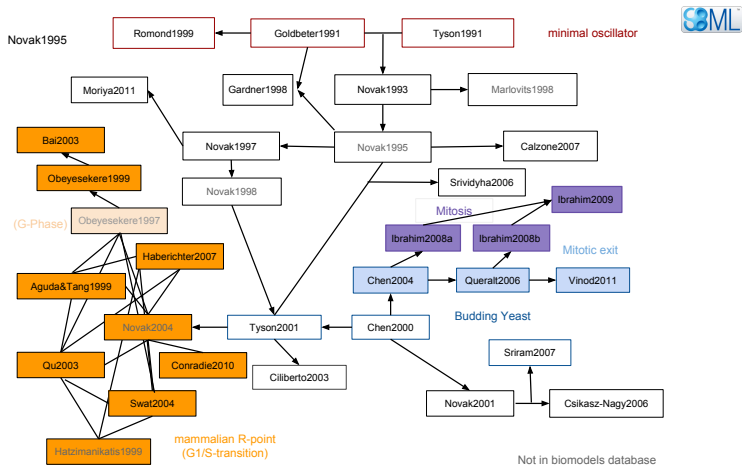


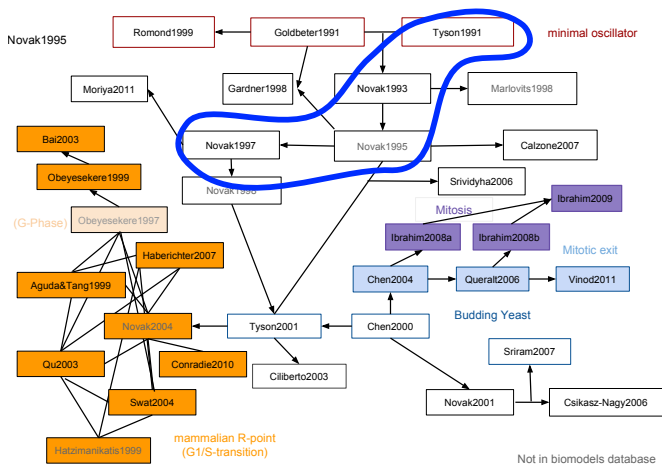


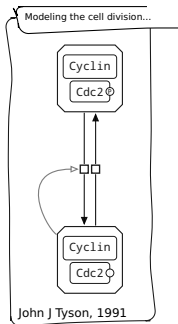


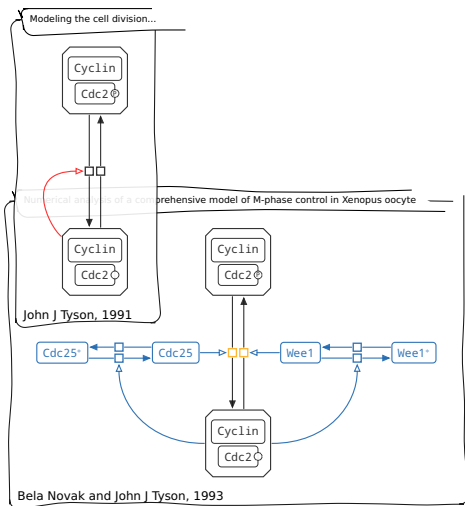


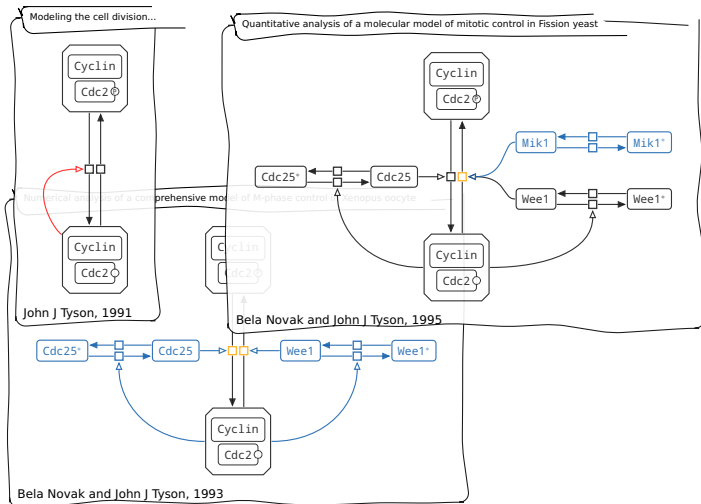


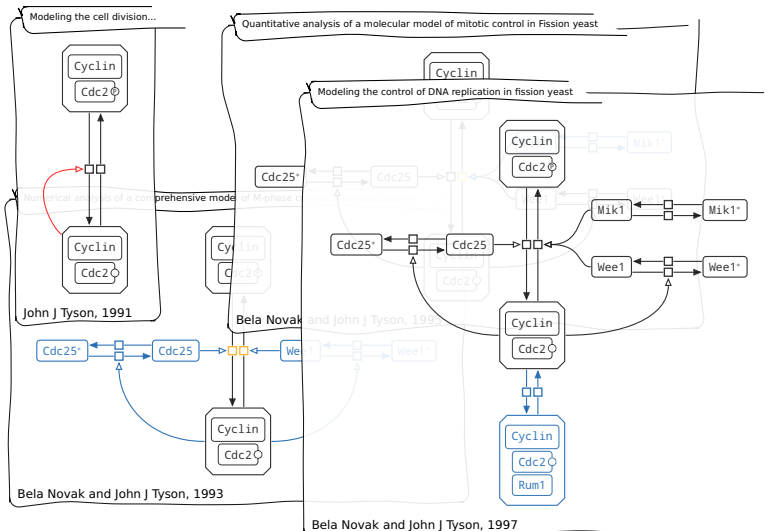


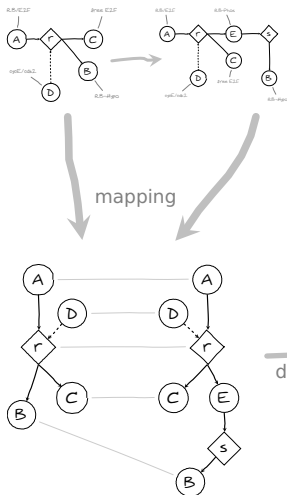












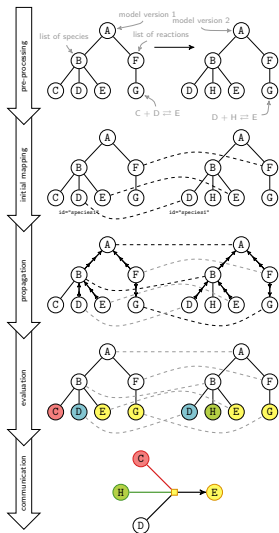


Biochemical Model Version Control System

- compares models encoded in standardised formats (currently:  and )
- maps hierarchically structured content
- constructs a diff (in XML format)
- is able to interpret this diff

```

<XML>
  Diff
  |
  |--- moves
  |   |
  |   |--- product of r: C
  |
  |--- deletes
  |   |
  |   |--- product of r: B
  |
  |--- inserts
  |   |
  |   |--- species: E
  |   |--- product of r: E
  |   |--- reaction s
  |
  </XML>
  
```

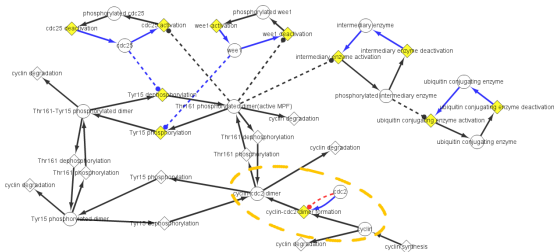


Identifying, Interpreting, and Communicating Changes in XML-encoded Models of Biological Systems

Scharm et. al. 2015, under revision at BIOINFORMATICS

Reactions

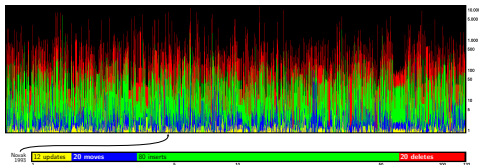
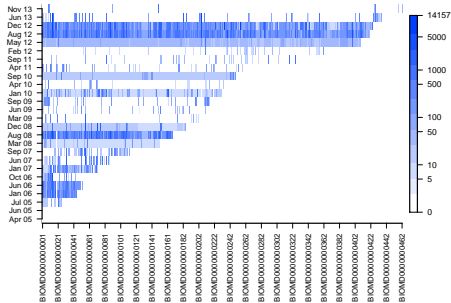
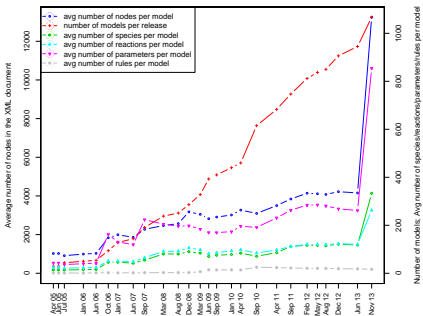
R19 (wee1 deactivation)	<ul style="list-style-type: none"> • <code>wee1 → wee1_p</code> • Modifiers: <code>dimer_p(unknown)</code>
R17 (cdc25 activation)	<ul style="list-style-type: none"> • <code>cdc25 → cdc25_p</code> • Modifiers: <code>dimer_p(unknown)</code>
R18 (cdc25 deactivation)	<ul style="list-style-type: none"> • <code>cdc25_p → cdc25</code>
R13 (Tyr15 dephosphorylation)	<ul style="list-style-type: none"> • <code>p_dimer_p → dimer_p</code> • Modifiers: <code>cdc25(unknown)</code>
R12 (Tyr15 phosphorylation)	<ul style="list-style-type: none"> • <code>dimer_p → p_dimer_p</code> • Modifiers: <code>wee1(unknown)</code>
R24 (ubiquitin conjugating enzyme deactivation)	<ul style="list-style-type: none"> • <code>Ubc_star → UBE</code>
R23 (ubiquitin conjugating enzyme activation)	<ul style="list-style-type: none"> • <code>UBE → Ubc_star</code> • Modifiers: <code>IE_p(unknown)</code>
R22 (intermediary enzyme deactivation)	<ul style="list-style-type: none"> • <code>IE_p → IE</code>
R21 (intermediary enzyme activation)	<ul style="list-style-type: none"> • <code>IE → IE_p</code> • Modifiers: <code>dimer_p(unknown)</code>
R20 (wee1 activation)	<ul style="list-style-type: none"> • <code>wee1_p → wee1</code>
R3 (cyclin-cdc2 dimer formation)	<ul style="list-style-type: none"> • <code>cyclin + cdc2 → dimer</code> • Modifiers: <code>cdc2(unknown)</code>



```

<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<BiVeS type="fullDiff">
  <update/>
  <delete>
    [...]
    <node id="6" oldChIDNo="1">
      <oldParent"/>
      <oldPath"/>
      <oldTag"/>
      <attribute id="1" name="specie" value="cdc2" triggeredBy="6"/>
      <oldVAlue"/>
    </delete>
    <insert>
      [...]
      <node id="12" newChIDNo="2">
        <newParent"/>
        <newPath"/>
        <newVAlue"/>
        <attribute id="13" name="specie" value="cdc2" triggeredBy="12"/>
        <attribute id="14" name="metaid" value="B18337" triggeredBy="12"/>
      </insert>
    [...]
  </BiVeS>


```



Scharm et. al. 2015, under revision at BIOINFORMATICS

Ad sponsored by COMBINE initiative

COMBINE 2015: October 12-16 in Salt Lake City

 **combine** the computational modeling in biology network



- Day 1: invited talks by Fred Adler, Mike Hucka, Richard Normann, Sharon Crook, Miriah Meyer, Huaiyu Mi, Tara Deans, and Anil Wipat
- Days 2-5: contributed talks and discussions.

Provenance for models of biological systems



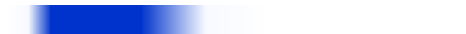
Woche

1 2 3 4 5 6 7 8 9 10

Setup / Aligning Visions



Learning & evaluating tools/workflows
Identifying research gap

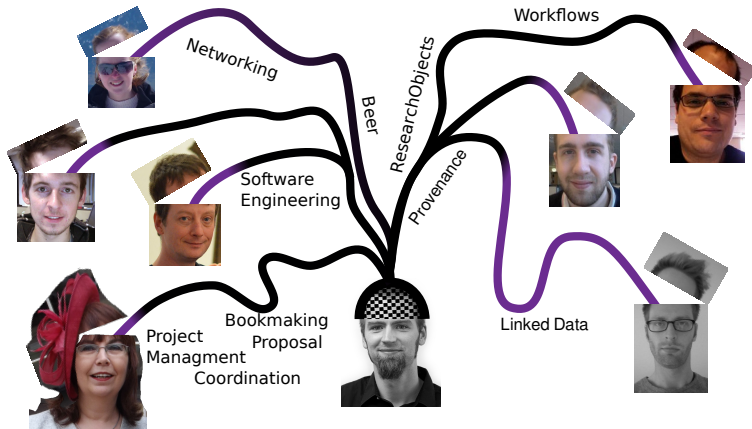


Doing a Case Study
Developing a concept and a schedule



Start writing a proposal
Planning future







Research Object

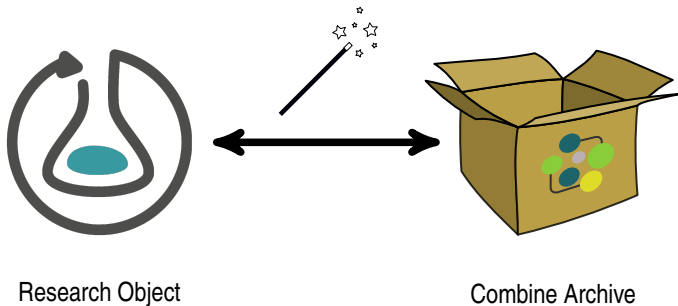
-VS-

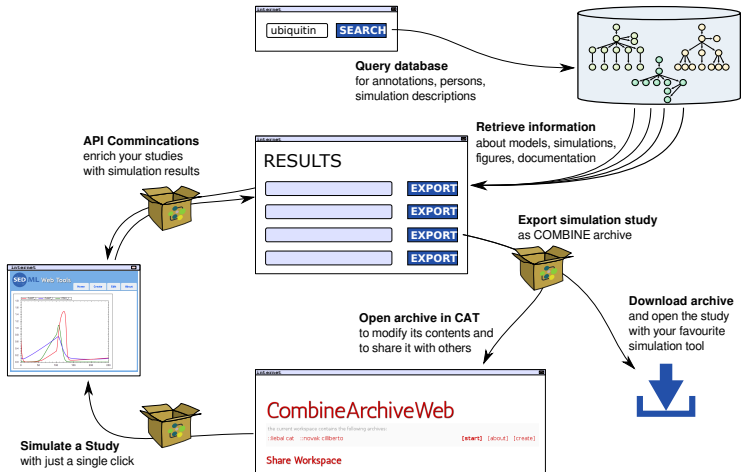


Combine Archive



	ZIP	Docker	VBox	CombineArchive	ResearchObject
Fancy Icon?	✗	✓	✓	✓	✓
Aspect 2	✗	✗	~	~	✓
Aspect 3	✗	~	✗	✓	~
Aspect 4	✗	~	✗	✓	✓
Aspect 5	?	?	?	?	?



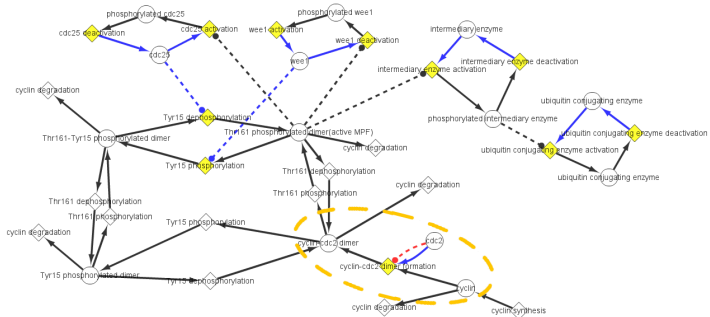


Extracting reproducible simulation studies from model repositories using the CombineArchive Toolkit.
Scharm et. al., DM4LS @ BTW 2015, Hamburg, GER





graphics taken from openclipart.org



✓ What?

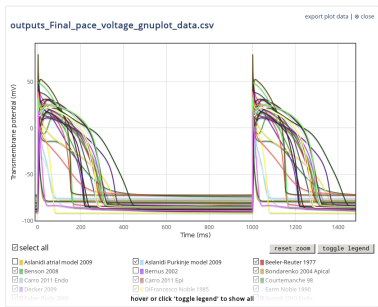
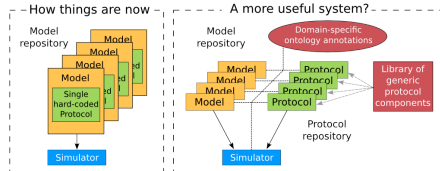
× Who? When? Why? How? ... ??



A call for virtual experiments: Accelerating the scientific process.
Cooper et al., Progress in biophysics and molecular biology (2014).

The Cardiac Electrophysiology Web Lab.

Cooper et al., submitted to Circulation: Arrhythmia and Electrophysiology



How things are now

A more useful system?

Ad sponsored by Jonathan Cooper

Workshop on the Web Lab
10th & 11th September 2015
Department of Computer Science, University of Oxford
<http://s.binfalse.de/fcworkshop>

A call for virtual experiments: Accelerating the scientific process.
Cooper et al., Progress in biophysics and molecular biology (2014).

The Cardiac Electrophysiology Web Lab.
Cooper et al., submitted to Circulation: Arrhythmia and Electrophysiology

Thank you!



Dagmar Waltemath, Ron Henkel, Martin Peters, Olaf Wolkenhauer



@SemsProject

<http://sems.uni-rostock.de>

Ad sponsored by Olaf Wolkenhauer

THE SYSTEMS MEDICINE WEB HUB

POSITIONS RESOURCES PROJECTS

Find jobs and experts Be informed Disseminate your news

PROMOTE YOUR RESEARCH INCREASE YOUR VISIBILITY

EVENTS REPORTS

SYSTEMS MEDICINE.NET

@sysmednet
#sysmednet
The Systems Medicine Web Hub
www.systemsmedicine.net
www.systemsmedicine.net/feed

An initiative of

www.systemsmedicine.net

EUROPEAN UNION
GERMAN MINISTRY OF EDUCATION AND RESEARCH
CASM
PTJ

References

- Sylvia Haus, Sara Jabbari, Thomas Millat, Holger Janssen, Ralf-Jörg Fischer, Hubert Bahl, John King, Olaf Wolkenhauer: *A systems biology approach to investigate the effect of pH-induced gene regulation on solvent production by Clostridium acetobutylicum in continuous culture*. BMC Systems Biology, 2011, 5:10.
- E.T. Papoutsakis: *Equations and calculations for fermentations of butyric acid bacteria*. Biotechnology and bioengineering, 1984, 26:174–187.
- H. Shinto, Y. Tashiro, M. Yamashita, G. Kobayashi, T. Sekiguchi, T. Hanai, Y. Kuriya, M. Okamoto, K. Sonomoto: *Kinetic modeling and sensitivity analysis of acetone–butanol–ethanol production*. Journal of biotechnology, 2007, 131:45–56.
- Thomas Millat, Holger Janssen, Graeme J. Thorn, John R. King, Hubert Bahl, Ralf-Jörg Fischer, Olaf Wolkenhauer: *A shift in the dominant phenotype governs the pH-induced metabolic switch of Clostridium acetobutylicum in phosphate-limited continuous cultures*. Applied Microbiology and Biotechnology, 2013, 97:6451–6466.
- John J Tyson: *Modeling the cell division cycle : cdc2 and cyclin interactions*. Proceedings of the National Academy of Sciences, 1991, 88:7328–7332.

- B Novak, J J Tyson: *Numerical analysis of a comprehensive model of M-phase control in Xenopus oocyte extracts and intact embryos*. Journal of Cell Science, 1993, 106:1153-1168.
- Bela Novak, John J. Tyson: *Quantitative analysis of a molecular model of mitotic control in fission yeast*. Journal of Theoretical Biology, 1995, 173:283-305.
- B Novak, J J Tyson: *Modeling the control of DNA replication in fission yeast*. Proceedings of the National Academy of Sciences of the United States of America, 1997, 94:9147-52.
- M Scharm, O Wolkenhauer, D Waltemath: *Identifying, Interpreting, and Communicating Changes in XML-encoded Models of Biological Systems*. Under review at BIOINFORMATICS.
- M Scharm, D Waltemath: *Extracting reproducible simulation studies from model repositories using the CombineArchive Toolkit*. In proceedings of the Workshop on Data Management for Life Sciences (DMforLS 2015) at BTW 2015, Hamburg, GER.
- J Cooper, M Scharm, G Mirams: *The Cardiac Electrophysiology Web Lab*. Submitted to Circulation: Arrhythmia and Electrophysiology.
- J Cooper, JO Vik, D Waltemath: *A call for virtual experiments: Accelerating the scientific process*. Progress in biophysics and molecular biology, 2014, 117:1, 99-106.



Further Literature

- S Bechhofer, D De Roure, M Gamble, C Goble, I Buchan: *Research objects: Towards exchange and reuse of digital knowledge*. In: The Future of the Web for Collaborative Science (FWCS 2010); 26 Apr 2010-26 Apr 2010; Raleigh, NC, USA.
- S Bechhofer, I Buchan, D De Roure, P Missier, J Ainsworth, J Bhagat, P Couch, D Cruickshank, M Delderfield, I Dunlop, M Gamble, D Michaelides, S Owen, D Newman, S Sufi, C Goble: *Why linked data is not enough for scientists*. Future Generation Computer Systems, 2013, 29(2), 599-611.
- FT Bergmann, R Adams, S Moodie, J Cooper, M Glont, M Golebiewski, M Hucka, C Laibe, AK Miller, DP Nickerson, BG Olivier, N Rodriguez, HM Sauro, M Scharm, S Soiland-Reyes, D Waltemath, F Yvon, NL Novère: *COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project*. BMC bioinformatics, 2014, 15(1), 369.