

Physiome: encouraging reproducibly FAIR computational modelling

David Nickerson, Tommy Yu, Soroush Safaei, Peter Hunter

https://doi.org/10.17608/k6.auckland.7040321





AUCKLAND BIOENGINEERING INSTITUTE

FAIR guiding principles for scientific data

https://doi.org/10.1038/sdata.2016.18

FAIRsharing.org standards, databases, policies

http://fairsharing.org/

Findable

Accessible

Interoperable

Reusable



http://fairdomhub.org/



https://models.physiomeproject.org/



http://fairmetrics.org/











Physiome papers describe the details of mathematical models and computational simulations associated with a 'primary' experimental/modelling paper that has been accepted to a peer-reviewed physiology, bioengineering or biophysics journal.

journal.physiomeproject.org

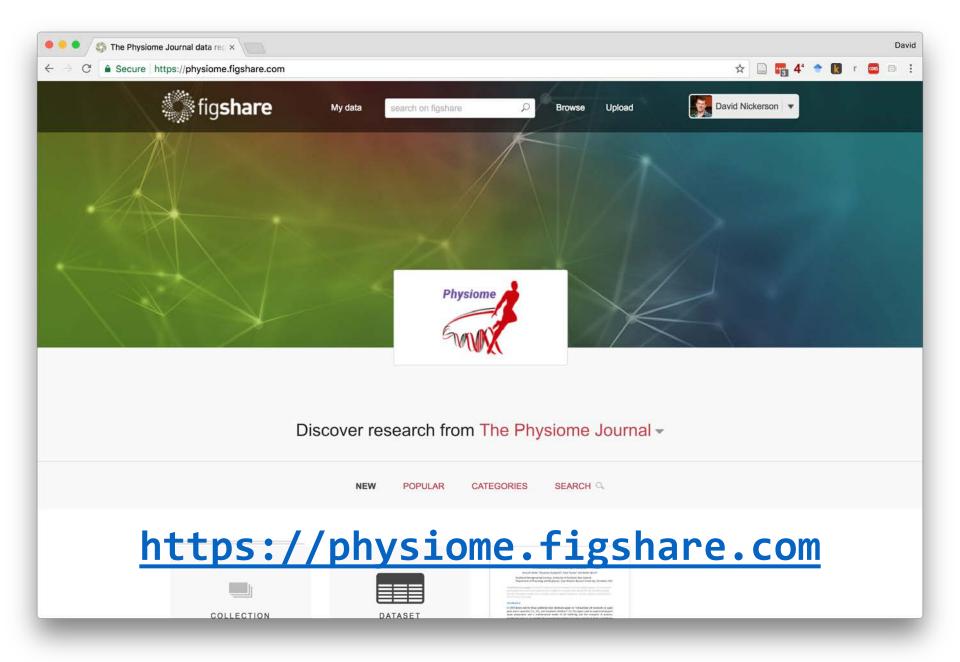
The purpose of *Physiome* is to encourage the reproducibility and reuse of models by providing citation credit for papers that describe and document curated and annotated models. The journal will assist authors in making suitable submissions available in the Physiome encoding standards.

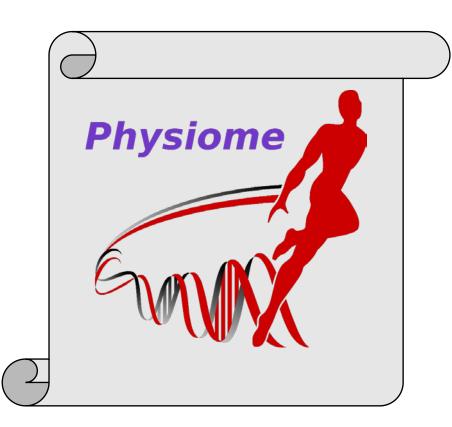
Open access with low cost.

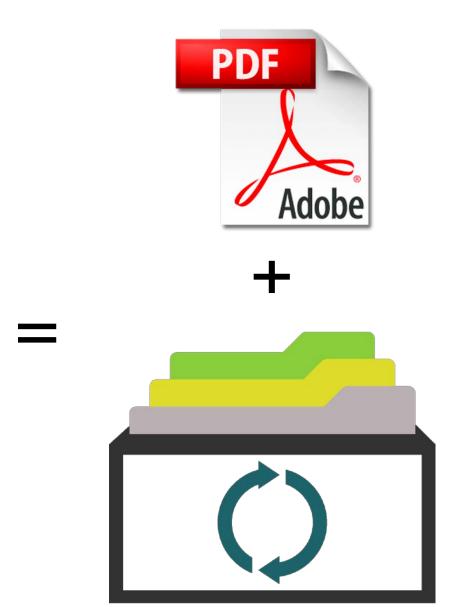






















Core Standards Standards for Knowledge Representation

BioPAX



Standards for Visual Representation





Standards for Models and their Analyses









Associated Standards Used by core standards

Projects



Infrastructure

BioModels.net qualifiers



Controlled Vocabularies



http://co.mbine.org/

Meeting in Boston this October: http://co.mbine.org/events/COMBINE_2018

Encouraging collaboration and reuse

• Track and recognise contributions







Metrics to demonstrate impact

- people get credit for making models FAIR
- e.g. altmetric donut

The Colours of the Donut



http://www.springersource.com/an-introduction-to-altmetric-data-what-can-you-see/







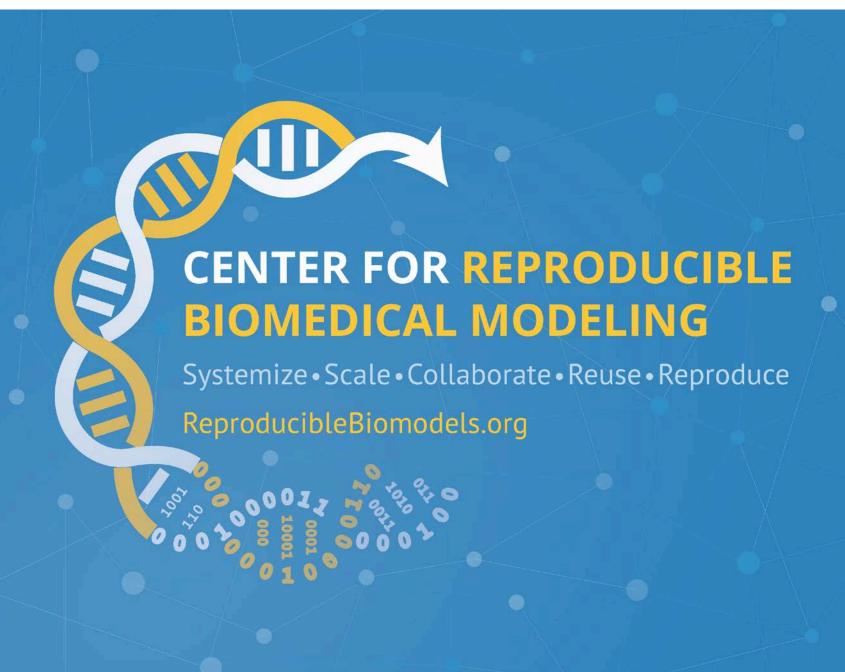
Encouraging collaboration and reuse

- Track and recognise contributions
- Needs to be easy
- Harmonising annotations
- "Rules" for how we construct models
- Reproducibility is key!











Systemize • Scale • Collaborate • Reuse • Reproduce

TECHNOLOGY DEVELOPMENT

Data discovery & integration
Scalable model building
Reproducible simulation
Reproducible visualization
Mining simulation results
Model provenance

AUTHOR & JOURNAL SERVICES

Technical guidance
Technology customization
Model annotation
Pre-publication model review
Simulation results repository
Online simulation & visualization

EDUCATION & OUTREACH

Scalable modeling Reuse & reproduciblity Ontologies & standards

REPRODUCIBLE BIOMODELS.ORG • INFO@REPRODUCIBLE BIOMODELS.ORG U Auckland • UConn • Mount Sinai • U Washington | NIBIB • NIGMS

Current state of *Physiome*

- Soft launch at IUPS World Congress
- https://journal.physiomeproject.org
- https://physiome.figshare.com
- Aiming for "first volume" end of 2018
- Submission site available soon after







Acknowledgements

- Physiome management board: Denis Noble, Walter Boron, Stig Omholt, Andrew McCulloch, Peter Hunter
- Jim Bassingthwaighte, Dan Beard, many others







Aotearoa Foundation













Key enabling technologies

- Unique, persistent, and resolvable identifiers
- Tools to encourage collaboration and reuse
- Metrics to demonstrate impact









Identifiers

Standard formats which provide identifier assignment













Identifiers

- Standard formats which provide identifier assignment
- Repositories adding:
 - persistence
 - versioning
 - resolvability

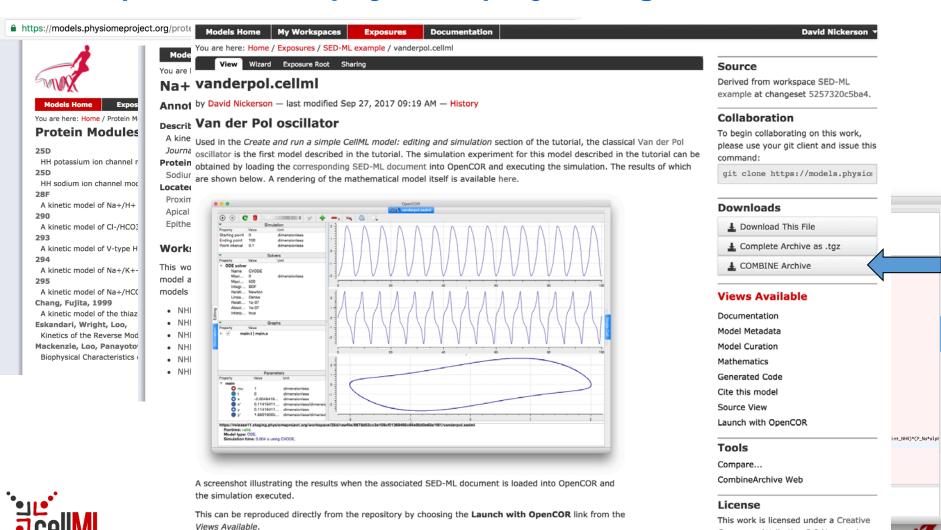






Physiome Model Repository (PMR)

https://models.physiomeproject.org



Commons Attribution 3.0 Unported

License.

Identifiers

- Standard formats which provide identifier assignment
- Repositories adding:
 - persistence
 - versioning
 - resolvability
- Annotations









Systemize • Scale • Collaborate • Reuse • Reproduce

Herbert M Sauro: University of Washington, Bioengineering Michael L Blinov: UConn Health, Cell Analysis and Modeling Dan Cook: University of Washington, Health Informatics John H Gennari: University of Washington, Health Informatics Arthur P Goldberg: Mt Sinai, Genetics and Genomic Sciences Jonathan R Karr: Mount Sinai, Genetics and Genomic Sciences Ion I Moraru: UConn Health, Cell Analysis and Modeling David P Nickerson: Auckland Bioengineering Institute Jim C Schaff: UConn Health, Cell Analysis and Modeling

Introduction

The Center for Reproducible Biomedical Modeling is a new technology development center with funded by the National Institute of Biomedical Imaging and Bioengineering with contributions from the National Insitute of General Medical Sciences. The long-term goal of the center is to enable comprehensive predictive models of biological systems, such as whole-cell models, that could help physicians personalize therapy and help bioengineers rationally design microorganisms. Achieving this goal requires procedures for system-

atically, scalably, and reproducibly building, simulating, verifying, and applying models. **Technology Development** TR&D 1: Data to Models Outreach Jonathan Karr Herbert Sauro MODEL CURATION TR&D 1 will develop tools for reproducibly TR&D3 building models. This will include (1) aggre TR&D1 gating large and heterogeneous data needed to build models, (2) organizing this Models data for model construction, and (3) design-Other domains ing models from this data. These tools will accelerate model construction. Our tools will also make model building transparent and TR&D2 You Tube DREAM \mathcal{H} reproducible by recording the data sources, Informatics assumptions, and design decisions used to build models. Simulation results Journal Curation Workflow David Nickerson Manuscripts received by journals will be curated to make sure that any author supplied code will faithfully reproduce the results presented in the manuscript Curator Data to Models Currently, we cannot comprehensively predict all of the behaviors of entire cells, tissue, organs, or organisms, in part because le with authors to correct emodel in PMR with revi we do not have mechanistic dynamical models that account for each individual biological process and environmental interaction. Consequently, scientists cannot fully explain how genotype influ ences phenotype, bioengineers cannot rationally design microorganisms, and physicia TR&D 2: Informatics genomic data to personalize medicine. Comprehensive models, such as multicellular models, that account for more processes are needed to accurately predict biology and advance bioengineering John Gennari and medicine. Some of the most significant barriers to better dynamical models, including "small" pathway models and "large" whole-cell models, are the onerous manual effort to aggregate model TR&D 2 will develop tools for annotating input data and design models and the opaqueness of the expert-driven dynamical model conthe meaning and provenance of models as struction process which obscures the data and assumptions used to build models well as annotating simulation results model behavior and model validation. This will include developing the schema and Model Provenance and ontologies for describing the prove-Distribution of Collaborative Projects nance, simulation data and validation Semantic Annotation As researchers build larger and more predictive models, it is becoming increasingly difficult to comprehend and interpret models, establish confidence in models, and analyze the results of their simulations. The SBML and CellML standards have made it easier to exchange models and the SED-ML standard has made it easier to exchange simulation experiments. However, it is still difficult to understand how models were constructed, to understand which parts of models were validated TR&D 3: Simulation Many models are still incomprehensible due to lack of annotation. We will develop novel schemas, Selection of Collaborative Projects: ontologies, and software tools for describing the provenance and semantic meaning of models, Ion Moraru simulation experiments, simulation results, and model validations MULTISCALE SIMULATION OF ACETAMINOPHEN-INDUCED LIVER TOXICITY: James Sluka TR&D 3 will develop tools for reproducibly ONLINE INTERACTIVE SIMULATION FOR BIOMODELS: Henning Hermiakols simulating and analyzing models online. This PRIME, MODELING VIRAL IMMUNITY AND ANTAGONISM: Stuart Sealfon will include (1) web-based tools for design-Web-based simulation MODEL ANNOTATION AND CURATION FOR PHYSIOME: Peter Hunter ing simulation experiments and visualizing MINICELL - A MODEL-DRIVEN APPROACH TO MINIMAL CELL ENGINEERING, Luis Serrano and visualization ent 3) simulation results, (2) a universal simulator MODELING CELL REGULATORY NETWORKS: Ravi Ivengar for simulating biomodels and (3) a database Standards such as SBML and SED-ML have improved the for organizing and storing simulation results. comprehensibility, reusability, and SP 1: Whole-cell modeling consulting (TR&D 1) will help researchers build, edit, and simueproducibility of biomodels late whole-cell models. The center is uniquely positioned to disseminate this new area of systems biology. We will provide whole-cell modeling expertise to provide researchers ad However, it remains challenging to hoc whole-cell modeling assistance. Starting in Year 2, SP 1 will begin to help researchers reuse many models and reproduce many simulation results and analyses. This irreproducibility limits the impact of biomodeling by inhibiting researchers from reusing models and simulation re utilize the software tools developed in TR&D 1 to systematically build whole-cell models sults for additional studies or combining models of individual biological processes into metamod-SP 2: Model curation (TR&D 2), SP 2 will curate and annotate models for authors and jourels of entire biological systems. It is hard to reuse and reproduce their results because (a) few modelers report the metadata needed to reproduce simulations, (b) there are many incompatible simu



Simulation results





lators, (c) there is no simulation results repository, (d) there is no standard for analyzing simulation

results, (e) there is no standard for describing visualizations of simulation results, and (f) the tools

for visualizing simulation results are inadequate





tial simulation results.

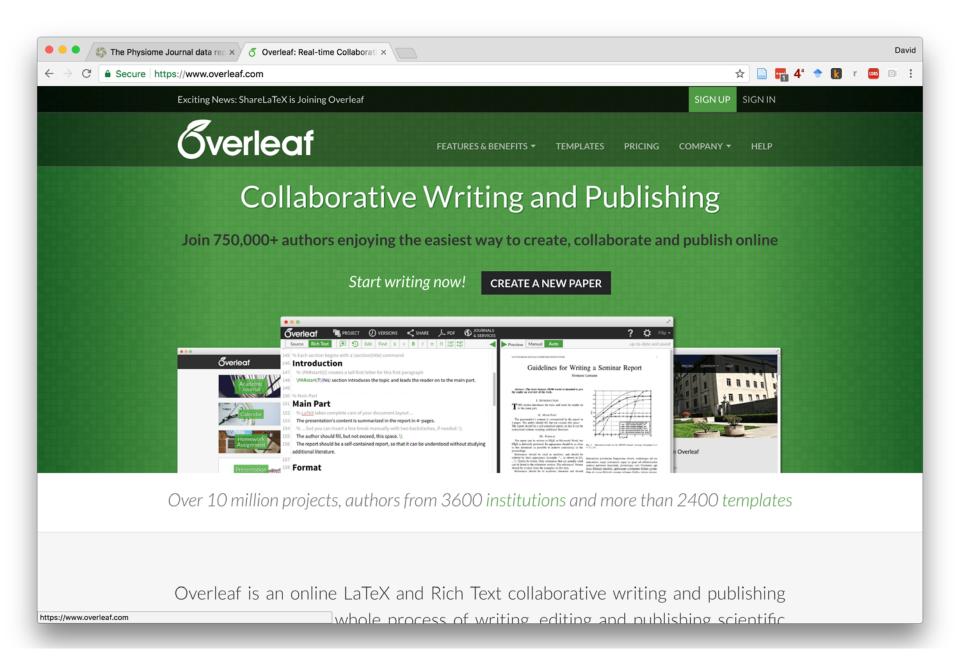


SP 3: Online simulation (TR&D 3) will help authors and journals use the tools developed in

TR&D 3 to simulate and visualize models online, as well as to analyze large datasets of spa-



https://journal.physiomeproject.org



https://www.overleaf.com

