COMBINE 2021

B Breakout C Contributed talk D Discussion I Invited talk L Lightning talk P Presentation session

T Tutorial

OCTOBER 11 · MONDAY

01:00 – 01:30 L **Lightning Talks A** Presentations Room

Speakers: Woosub Shin, Yuda Munarko, John Gennari

1. **Woosub Shin** (University of Auckland / University of Washington): *SBMate: A Framework for Evaluating Quality of Annotations in Systems Biology Models*
2. **Yuda Munarko** (Auckland Bioengineering Institute, University of Auckland): *The Searching of Entities in Annotated Biosimulation Models from The PMR and BioModels Using NLIMED*
3. **John Gennari** (University of Washington): *OMEX metadata specification, v 1.2*
4. **Discussion**

- Woosub Shin (University of Auckland / University of Washington); Hellerstein, Joseph (University of Washington); Munarko, Yuda (University of Auckland); Neal, Maxwell (Seattle Children's Research Institute); Nickerson, David (University of Auckland); Rampadarath, Anand (University of Auckland); Sauro, Herbert (University of Washington); Gennari, John (University of Washington)

SBMate: A Framework for Evaluating Quality of Annotations in Systems Biology Models

The interests in repurposing and reusing existing systems biology models have been growing in recent years. Semantic annotations play an important role for this, by providing crucial information on the meanings and functions of models. However, we have not found existing tools that test whether the annotations exist or if they are of high-quality.

In this lightning talk, we introduce SBMate, a python package that would serve as a framework for evaluating the quality of annotations of systems biology models. Three default metrics are provided: coverage, consistency, and specificity. Coverage checks whether annotations exist in a model. Consistency quantifies if the annotations are appropriate for the given model entity. Finally, specificity indicates how detailed the annotations are. We discuss the three metrics and the challenges with evaluating annotations, using the models in the BioModels repository. Additional metrics would be easily added to extend the current version of SBMate.

- Yuda Munarko (Auckland Bioengineering Institute, University of Auckland); Sarwar, Dewan (Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand); Rampadarath, Anand (Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand); Atalag, Koray (Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand); Gennari, John (Biomedical & Health Informatics, University of Washington, Seattle, USA); Neal Maxwell(Seattle Children's Research Institute, Seattle, USA); Nickerson David (Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand)

The Searching of Entities in Annotated Biosimulation Models from The PMR and BioModels Using NLIMED

The semantic annotation of the biosimulation model with standard ontologies allows one to quickly and fully understand the model. Currently, hundreds of models with thousands of entities in the PMR and Biomodels are annotated using RDF. Therefore, it is possible to find those entities in different levels of granularity, such as variable, component, and reaction, using SPARQL. Since composing SPARQL complicated, we developed NLIMED, a natural language query (NLQ) based searching tool, automatically convert NLQ into SPARQL. It works by identifying topics in NLQ using Named Entity Recognition (NER) and information retrieval techniques. Topics associated with ontology classes and their accompanying predicates are then organised into SPARQL with AND logic. Since a topic can relate to more than one ontology class, NLIMED can generate ranked SPARQLs based on its relation probability. Now, NLIMED is available online for the PMR and Biomodels; hence, it may provide more effortless searching of entities. With the NLQ based searching, we can potentially discover similar entities annotated with a different ontology in both repositories. Finally, we are confident that different repositories in the biosimulation model or other domains can implement NLIMED's approach for their entity searching.

- John Gennari (University of Washington); Nickerson, David; Misirli, Goksel; Konig, Matthias; Neal, Maxwell; Waltemath, Dagmar

OMEX metadata specification, v 1.2

Although the COMBINE community has reached consensus on the general approach toward semantic annotation of biosimulation models, details about best practices for annotation have been lacking. The OMEX metadata specification provides these details, and version 1.2 has recently been ratified and published in the Journal of Integrative Bioinformatics. In addition, we have created software libraries (for C/C++: libOmexMeta; for python: pyomexmeta) that allow users to create OMEX archives and RDF annotation files that conform to v1.2 of the OMEX metadata specification.

The 1.2 specification organizes annotations into model-level annotations, simple singular annotations, and composite annotations. For each of these, we provide examples, and recommendations for knowledge sources (e.g., ontologies) to use to achieve consistency in metadata across our community. We also provide a base namespace, “omex-library.org”, so that these archives and their annotations can be unified into a single RDF graph or knowledge base. The resulting knowledge base can be queried for any type of annotation described by the specification, providing improved compliance with FAIR principles.

01:00 – 04:00 P **Session A - presentation room** Presentations

Moderators: David Nickerson

02:00 – 02:20 C **Mathematical model notation - proposal for SBGN** Presentations Room

Speakers: Ilya Kiselev

Kiselev Ilya (Sirius University of Science and Technology, Sochi, Russia); Kolpakov Fedor (Sirius University of Science and Technology, Sochi, Russia), Akberdin Ilya (Sirius University of Science and Technology, Sochi, Russia), Kutumova Elena (Sirius University of Science and Technology, Sochi, Russia)

In BioUML Platform from the start we utilized a visual approach to modeling. It implies that the model is created and modified as a visual diagram, each element of which is associated with a mathematical object (equation, variable, event etc.). Users may edit properties of those elements (e.g. formula, initial value, trigger, etc.) On the basis of visual representation along with specified properties corresponding executable code is generated and used for numerical simulation.

As a mathematical basis for models we use the SBML standard, while for visual representation the main standard is SBGN. Thanks to SBML annotations we are able to include necessary information about layout and other visual specifics into SBML documents. However not any SBML element has an SBGN counterpart and vice-versa. Particularly those elements are equations, functions, events and constraints. On the other hand, many mathematical models do not involve reactions or similar processes. Instead they may be formulated explicitly with algebraic and differential equations and discrete events. Such models still benefit from visual representation and thus require appropriate notation.

We propose to extend SBGN notation for such models. Notation should include elements for abstract variables, equations and other mathematical elements. Additionally arcs representing relations between variables indicate their interactions (e.g. whether increase of one variable’s value also increases value of another variable).

Other interesting cases arise when we consider modular models. The package “comp” for hierarchical SBML models provides a powerful tool for modular models creation, on the other hand, while SBGN supports hierarchical models, unfortunately it lacks convenient instruments for creation of complex multilevel models.

At the same time, the approach to modularity, which we use in BioUML is different from the SBML approach - we consider modules to be connected through mathematical variables - inputs, outputs and shared variables, while SBML allows replacements of arbitrary elements in modules. Thus our aim was to extend visual notation in order to facilitate creating complex modular models while keeping it compatible with SBML.

02:20 – 02:40 C **WikiNetworks: translating manually created biological pathways for topological analysis**
Speakers: Mukta G. Palshikar Presentations Room
Palshikar, Mukta G. (Biophysics, Structural, and Computational Biology Program, University of Rochester School of Medicine and Dentistry); Hilchey, Shannon P. (University of Rochester School of Medicine and Dentistry); Zand. Martin S. (University of Rochester School of Medicine and Dentistry); Thakar, Juilee (University of Rochester School of Medicine and Dentistry);

Summary: WikiPathways is a database of 2979 biological pathways across 31 species created manually using the graphical editor PathVisio. These pathway representations can be exported in GPML (Graphical Pathway Markup Language) and in Systems Biology Graphical Notation (SBGN) format. Currently available tools do not translate pathways into networks that are useful for topological analysis. The WikiPathways app for Cytoscape allows the import and visualization of WikiPathways entries. However, this app is unable to correct drawing errors introduced by hand-curation methods, such as unconnected interactions. In addition, the output of the Cytoscape app has superfluous nodes, duplicate nodes, node insertion at hyperedges, and does not process edges to the components of node groups (e.g., protein complexes).

We developed the WikiNetworks package to standardize and construct directed networks by combining geometric information and manual annotations from WikiPathways. WikiNetworks uses the positions of pathway elements in Cartesian space to resolve ambiguous edges based on physical proximity of pathway elements and eliminates nodes that do not correspond to biological entities. We compared the output of WikiNetworks to 15 manually curated pathways from WikiPathways and found that WikiNetworks performs significantly better than the existing tool. The output of WikiNetworks is thus truly ready for input into network analysis software, and for use with dynamical modeling and simulation software.

Availability and Implementation: WikiNetworks is written in Python3 and is available on github.com/Thakar-Lab/wikinetworks and on PyPI.

02:40 – 03:00 C **The Simulation Experiment Description Markup Language (SED-ML) Level 1 Version 4**
Speakers: Lucian Smith, Matthias König Presentations Room
Smith, Lucian (University of Washington Seattle); König, Matthias (Humboldt University Berlin)

Computational simulation experiments increasingly inform modern biological research, and bring with them the need to provide ways to annotate, archive, share and reproduce the experiments performed. These simulations increasingly require extensive collaboration among modelers, experimentalists, and engineers. The Minimum Information About a Simulation Experiment (MIASE) guidelines outline the information needed to share simulation experiments. SED-ML is a computer-readable format for the information outlined by MIASE, created as a community project and supported by many investigators and software tools.

The first versions of SED-ML focused on deterministic and stochastic simulations of models. Level 1 Version 4 of SED-ML substantially expands these capabilities to cover additional types of models, model languages, parameter estimations, simulations and analyses of models, and analyses and visualizations of simulation results. To facilitate consistent practices across the community, Level 1 Version 4 also more clearly describes the use of SED-ML constructs, and includes numerous concrete validation rules.

SED-ML is supported by a growing ecosystem of investigators, model languages, and software tools, including eight languages for constraint-based, kinetic, qualitative, rule-based, and spatial models, over 20 simulation tools, visual editors, model repositories, and validators.

Additional information about SED-ML is available at https://sed-ml.org/.

Within this talk an overview of SED-ML Level 1 Version 4 is provided.

03:00 – 03:20 C **Modular approach for modeling of the human individual response to antihypertensive therapy with generation of virtual populations** Presentations Room

Speakers: Elena Kutumova

Kutumova Elena (Sirius University of Science and Technology, Sochi, Russia); Kiselev Ilya (Sirius University of Science and Technology, Sochi, Russia); Sharipov Ruslan (Sirius University of Science and Technology, Sochi, Russia); Lifshits Galina (Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia); Fedor Kolpakov (Sirius University of Science and Technology, Sochi, Russia)

Hypertension is the most common risk factor for the development of cardiovascular, cerebrovascular and renal diseases. Therefore, mathematical modeling of circulatory regulation taking into account the response to antihypertensive therapies with different mechanisms of action has become increasingly important. We divide the modelling process into several stages.

Stage 1. Development of a mathematical model of cardiovascular and renal physiology. Our approach involves creation of a set of modules for basic physiological processes and construction of a model using these modules for a given disease.

Stage 2. Pharmacokinetic/pharmacodynamic modelling of antihypertensive medications including angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium-channel blockers, thiazide diuretics and β-Adrenoreceptor blockers. This stage is based on the detection of target points in the model for each of these antihypertensive drug classes and validation of unknown dynamics using available clinical trials.

Stage 3. The model personalization. We consider a stationary parameterization of the model within physiological ranges as a virtual patient. Some values in parametrization we take from the laboratory analyses of a real person. For other difficult to measure values, we consider physiologically acceptable variations, thus getting a virtual population for the person.

Stage 4. Treatment simulation of the virtual population. This stage allows to get groups of constructed patients with a similar response to the medication. Therefore, we can analyze which characteristics of the person result in the effectiveness (or ineffectiveness) of the antihypertensive therapy.

The sequential implementation of these stages can provide a working scheme to optimize the choice of drug therapy for the treatment of arterial hypertension in particular patients.

Our model of cardiovascular and renal systems is contained in the web-edition of the BioUML software (https://sirius-web.org/bioumlweb/) and is available on Git repository (https://gitlab.sirius-web.org/virtual-patient/blood-pressure-regulation) including several representative use cases for working with the model in Jupyter notebook.

03:20 – 04:00 D **Open discussion** Presentations Room

06:00 – 06:30 L **Lightning Talks B** Presentations Room

Speakers: Dilan Pathirana, Rahuman Sheriff, Woosub Shin, Hugh Sorby

1. **Dilan Pathirana** (University of Bonn): *PEtab Select: add model selection to a PEtab-based calibration workflow*
2. **Rahuman Sheriff** (European Bioinformatics Institute (EMBL-EBI)): *Model Curation in BioModels and FROG*
3. **Woosub Shin** (University of Auckland): *Automating Model Annotations: Developing a Recommender System for Annotating Systems Biology Models*
4. **Hugh Sorby** (University of Auckland): *libCellML: almost at version 1*
5. **Discussion**

- Dilan Pathirana (University of Bonn); Hasenauer, Jan (University of Bonn); the PEtab Select contributors

PEtab Select: add model selection to a PEtab-based calibration workflow

PEtab is a file format for the specification of parameter estimation problems, and has been adopted by several systems biology modelling tools. PEtab facilitates use of the different state-of-the-art methods, which are unique to specific calibration tools, on the same PEtab problem. This includes methods related to optimization and uncertainty analysis. The current specification allows for the semi- or full-automation of several tasks in the model calibration process, on the level of a single model. The work presented here extends PEtab, and compatible tools, to support efficient calibration of, and selection from, multiple models. The extension supports specification of the: model space and its constraints; selection algorithms and criteria; parameter estimation problems via PEtab; and initial models. PEtab Select also provides a Python library and command-line interface for the management of the model space and selection problems such that tools, which already support model calibration with PEtab, can easily implement model selection.

- Rahuman Sheriff (European Bioinformatics Institute (EMBL-EBI)); BioModels Team (EMBL-EBI); FBC curation standards working group;

Model Curation in BioModels and FROG

BioModels (https://www.ebi.ac.uk/biomodels) is one of the largest repositories of curated mathematical models of biological and pharmacological processes. Models submitted to BioModels repository are manually curated. Curation process involves encoding models in the standard SBML format, ensuring reproducibility of simulation results in the reference manuscript and semantic enrichment of the model with controlled vocabularies. A new curation milestone was recently achieved by BioModels; over 1000 models are now fully curated. Our recent model curation target areas include tumor immune interaction, cell cycle and COVID-19. Over 99% of the curated models are ordinary differential equation models. BioModels Parameters resource facilitates easy search and retrieval of model parameters from these kinetic models. Furthermore, BioModels has started implementation of FROG analysis (https://www.ebi.ac.uk/biomodels/curation/fbc), a community standard to foster reproducibility and curation of constraint-based models including Genome Scale Metabolic models (GEMs). FROG analysis generates FROG report, a reference dataset that consists of Objective function value, Flux Variability Analysis (FVA), Gene and Reaction deletion fluxes. The FROG report can be then used to assess the reproducibility and curate the model. We invite the community to submit GEMs to BioModels along with the FROG report to facilitate curation and contribute to the community manuscript.

- Woosub Shin (University of Auckland); Hellerstein, Joseph (University of Washington); Sauro, Herbert (University of Washington); Gennari, John (University of Washington)

Automating Model Annotations: Developing a Recommender System for Annotating Systems Biology Models

The utility and usefulness of a biomedical model is greatly increased by providing semantically meaningful annotations of model elements, such as chemical species and reactions. At present, annotations are applied manually, a process that is time-consuming and skill-intensive.

We are developing a recommender system that suggests annotations for biomedical models. Our current focus is

to annotate reactions using available information from annotations of chemical species. Our initial focus is on models in SBML format. In this lightning talk, we describe the current state of annotations in the BiGG and BioModels repositories as well as discuss our progress and challenges with building the recommender system.

- Hugh Sorby (University of Auckland); Garny, Alan (University of Auckland);Nickerson, David (University of Auckland)

libCellML: almost at version 1

We will summarise the current status and future plans of libCellML as we edge ever close to releasing version 1 of the libCellML library.

06:00 – 09:00 P **Session B - presentation room** Presentations
Moderators: Dagmar Waltemath

07:20 – 07:40 C **PEtab-MS – a format for specifying parameter estimation problems for multiscale models**
Speakers: Emad Alamoudi Presentations Room
Alamoudi, Emad (University of Bonn); Starruß, Jörn, Center of Information Services and High Performance Computing (ZIH), Technische Universität Dresden, Dresden, Germany; Hasenauer, Jan, Faculty of Mathematics and Natural Sciences, University of Bonn, 53113 Bonn, Germany, Helmholtz Zentrum München - German Research Center for Environmental Health, Institute of Computational Biology, 85764 Neuherberg, Germany, Technische Universität München, Center for Mathematics, Chair of Mathematical Modeling of Biological Systems, 85748 Garching, Germany; FitMultiCell Consortium, Faculty of Mathematics and Natural Sciences, University of Bonn, 53113 Bonn, Germany, Center of Information Services and High Performance Computing (ZIH), Technische Universität Dresden, Dresden, Germany, Helmholtz Zentrum München - German Research Center for Environmental Health, Institute of Computational Biology, 85764 Neuherberg, Germany, Technische Universität München, Center for Mathematics, Chair of Mathematical Modeling of Biological Systems, 85748 Garching, Germany, Institute for Medical Informatics and Biometry, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany, Centre for Modelling and Simulation in the Biosciences, BioQuant, Heidelberg University, 69120, Heidelberg, Germany

Many biological tissues are highly organized and dynamic. Their Spatio-temporal patterns are relevant in many biological processes, including tissue homeostasis, viral infection, or tumor development. Multiscale modeling that describes the construction of a system model from several sub-models on different length scales became an important tool to simulate the dynamics of these biological tissues. Yet, reproducibility and reusability of the results is still an issue.

Here, we introduce a PEtab extension called PEtab-MS, which facilitates the specification of parameter estimation problems for multiscale models. PEtab-MS encodes the estimation problem, including unknown parameters, parameter bounds, experimental conditions, and measurement data. It builds upon the PEtab format which is well tested and already supported in many simulations and parameter estimation toolboxes with hundreds of users. The new extension is already supported with FitMultiCell pipeline, which is a user-friendly, open-source, and scalable platform for simulating and parameterizing computational models of multi-cellular processes.

07:40 – 08:00 C **Introduction to OpenML** Presentations Room
Speakers: Prabhant Singh
Singh, Prabhant (OpenML)

Machine learning research should be easily accessible and reusable. OpenML is an open platform for sharing datasets, algorithms, and experiments - to learn how to learn better, together. This talk will cover introduction to OpenML and how to use the platform for reproducible science.

08:00 – 08:20 C **Synchronisation of visual and textual formats for mathematical models: Antimony extension in BioUML** Presentations Room

Speakers: Ilya Kiselev

Kiselev Ilya (Sirius University of Science and Technology, Sochi, Russia); Kolpakov Fedor (Sirius University of Science and Technology, Sochi, Russia), Akberdin Ilya (Sirius University of Science and Technology, Sochi, Russia), Kutumova Elena (Sirius University of Science and Technology, Sochi, Russia)

To ensure a model exchange and reproducibility of simulation results a variety standards for model representation have been developed over the past decades (SBML, SBGN, CellML, BioPAX). Depending on a context it can be more suitable to present a model of a biological system either as a graph (for example, as a map of metabolic or signal transduction pathway) or by means of the human-readable text such as Antimony language.

Developing the BioUML Platform we from the start tried to combine visual modeling (creating mathematical models as visual diagrams). As a standard for visual modeling the tool mainly employs SBGN notation, while SBML is used as a foundation for mathematical formalism of the model. In the past few years we also made an effort to support human-readable model representation by adopting Antimony language. We chose Antimony because it is very close to SBML and SBGN standards which facilitates their combining in the frames of one tool.

However, there are some gaps between different standards resulting in the interoperability issues:
- some SBML elements (for example, algebraic equations in unresolved form) can not be described in Antimony syntax;
- Antimony does not provide enough information to generate SBGN diagrams.

In this talk we would like to discuss possible extensions of the Antimony language in order to cover those problems. There are two possible ways of extension:
- Direct addition of new terms into Antimony to make it more compliant with SBML and SBGN.
- Addition of annotation mechanism into Antimony language similar to the scheme of the SBML enabling using custom terms to annotate Antimony objects.

Next are some possible suggestions for Antimony extension. Some of them may be added directly to the language, while others should use custom annotations.
- algebraic equations in unresolved form;
- ports direction (input, output, contact);
- annotation of complexes and entities inner structure (including subunits, units of information and state variables). Possible annotation form is “(A{p}:B):(D)3”;
- SBGN entity types (macromolecule, simple chemical etc.);
- process types (association, dissociation, etc.);
- custom molecule structure (e.g. glycans);
- layout information (node location, borders, edge path);

08:20 – 09:00 D **Open discussion** Presentations Room

PINNED
10:00 – 10:45 I **Invited talk: Anna Niaraki: Do you speak Systems Biology? Shaping a common language for community work** Presentations Room

Speakers: Anna Niarakis

In this talk, I will present recent developments in Disease maps and CoLoMoTo consortia, and discuss how the use of standards and best practices help bring Systems Biology communities closer.

10:00 – 13:00 P **Session C - presentation room** Presentations

Moderators: Matthias König

11:00 – 11:20 C **A physiological-based model of indocyanine green for the evaluation of liver function based on COMBINE standards** Presentations Room

Speakers: Matthias König

König, Matthias (Humboldt-University Berlin); Köller, Adrian (Humboldt-University Berlin); Grzegorzewski, Jan (Humboldt-University Berlin)

11:20 – 11:40 C **Fides: Reliable Trust-Region Optimization for Parameter Estimation of Ordinary Differential Equation Models** Presentations Room

Speakers: Fabian Froehlich

Froehlich, Fabian (Harvard Medical School)

Motivation: Because they effectively represent mass action kinetics, ordinary differential equation models are widely used to describe biochemical processes. Optimization-based calibration of these models on experimental data can be challenging, even for low-dimensional problems. However, reliable model calibration is a prerequisite for many subsequent analysis steps, including uncertainty analysis, model selection and biological interpretation. Although multiple hypothesis have been advanced to explain why optimization based calibration of biochemical models is challenging, there are few comprehensive studies that test these hypothesis and tools for performing such studies are also lacking.

Results: We implemented an established trust-region method as a modular python framework (fides) to enable structured comparison of different approaches to ODE model calibration involving Hessian approximation schemes and trust-region subproblem solvers. We evaluate fides on a set of PEtab benchmark problems that include experimental data. We find a high variability in optimizer performance among different implementations of the same algorithm, with fides performing more reliably that other implementations investigated. Our investigation of possible sources of poor optimizer performance identify shortcomings in the widely used Gauss-Newton approximation. We address these shortcomings by proposing a novel hybrid Hessian approximation scheme that enhances optimizer performance.

11:40 – 12:00 C **Interfacing SBML and CellML with the scientific machine learning and constrained-based modeling ecosystems in Julia** Presentations Room

Speakers: Paul F. Lang

Lang, Paul F (University of Oxford); Iravanian, Shahriar; Kratochvil Miroslav; Jain, Anand; Giordano, Mose; Rackauckas, Chris

Julia is a general purpose programming language that was designed for simplifying and accelerating numerical analysis and computational science. Building on the computer algebra system Symbolics.jl, many high-performing solvers for constrained systems and differential equations have been developed in Julia. In particular the Scientific Machine Learning (SciML) ecosystem of Julia packages includes ModelingToolkit.jl, a modeling framework for high-performance symbolic-numeric computation in scientific computing and scientific machine learning. It allows for users to give a high-level description of a model for symbolic preprocessing to analyze and enhance the model. ModelingToolkit can automatically generate fast functions for model components, along with automatically sparsifying and parallelizing the computations. This enabled highly performing solvers of differential equations, parameter optimisation algorithms and methodologies for automated model discovery. To give the systems biology community easy access to SciML, we developed a set of three tools that import biochemical models specified in the SBML or CellML community standards into the SciML ecosystem: SBML.jl, SBMLToolkit.jl and CellMLToolkit.jl SBML.jl is a Julia wrapper for the main functionality of the libSBML library via its C language bindings. It also enables constrained-based modelling by supporting the Flux Balance Constraints SBML package. SBMLToolkit.jl and CellMLToolkit.jl import dynamic SBML and CellML models, respectively, into the SciML ecosystem. We show that SBML.jl alone is sufficient for backing up a facile reconstruction of large-scale constraint-based metabolic models and that combined community models that consist of tens of millions individual reactions may be constructed interactively on off-the-shelf hardware, without requiring any special programming techniques or other support libraries. We further demonstrate how the SciML tools for dynamic systems can be used to accelerate model simulation and fitting of kinetic parameters. We will also provide an example of filling mechanistic knowledge gaps in epidemiological SBML and CellML toy models using neural networks and sparse identification of nonlinear dynamics. By providing the computational systems biologists with easy access to the open-source Julia ecosystem we hope to catalyse the development of further Julia tools in this domain, and the growth of the Julia bioscience community.

12:00 – 12:20 C **Connecting computational biology models and COVID-19 resources** Presentations Room

Speakers: Lea Gütebier

Gütebier, Lea (Medical Informatics Laboratory, University Medicine Greifswald); Henkel, Ron (University Medicine Greifswald); Bleimehl, Tim (German Center for Diabetes Research); Preusse, Martin (Kaiser & Preusse); Waltemath, Dagmar (University Medicine Greifswald)

Within a short time, a lot of data on COVID-19 has been published from different research groups. Integrating this data is essential for the better understanding of the disease.

In 2020 HealthECCO (https://healthecco.org/) initiated the CovidGraph project [1] to support the exploration of COVID-19 data. The CovidGraph, a Neo4j database, integrates various data sets regarding COVID-19 and the corona virus family in a single knowledge graph. It includes data from different domains such as publications, patents, biological entities, clinical studies and more.
Another graph-based project aiming at better integration of biomedical research data is MaSyMoS [2]. MaSyMoS stores computational models, simulation descriptions and associated meta-data, including reference publications and terms from bio-ontologies in a Neo4j graph database. BioModels serves as one of the data sources, including a highly curated collection of COVID-19 studies.

We integrated MaSyMoS as an additional data provider into the CovidGraph. Therefore, we matched relevant publications from MaSyMoS and CovidGraph and, in addition, terms from biomedical ontologies.
More specifically, our workflow matches the corresponding IDs (e.g., PubMedID) of publications from the COVID-19 Open Research Dataset contained in CovidGraph, and publications from PubMed contained in MaSyMoS. Furthermore, terms from ontologies relevant to COVID-19, and information about biological entities are stored in the biomedical data domain of the CovidGraph. We mapped overlapping ontology terms from both, MaSyMoS annotations and the CovidGraph, by their corresponding IDs. The overlap includes the Gene Ontology, ChEBI, UniProt and Disease Ontology.
The CovidGraph represents a network of integrated data providers and allows for data exploration across different domains, particularly medical research and systems biology.

In our talk we present the CovidGraph and explain how we integrated the MaSyMoS data set. We outline the developed integration concept to show that further data resources can easily be integrated to extend the CovidGraph.

The COVID-19 collection on BioModels was developed as part of Grant: European Commission: EOSCsecretariat.eu - EOSCsecretariat.eu (831644).

[1] Gütebier et al. “COVIDGraph: Connecting biomedical COVID-19 resources and computational biology models.” SEA Data 2021
[2] Henkel et al. “Combining computational models, semantic annotations and simulation experiments in a graph database.” Database

12:20 – 13:00 D **Open discussion** Presentations Room

14:00 – 14:30 L **Lightning Talks D** Presentations Room

Speakers: Matthias König, Krishna Tiwari, Ciaran Welsh, Tung Nguyen

1. **Matthias König** (Humboldt-University Berlin): *SBML4Humans - Interactive SBML report for Humans*
2. **Krishna Tiwari** (EMBL-EBI): *Model reproducibility: implementing reproducibility scorecard and using scorecard for curation prioritization*
3. **Ciaran Welsh** (University of Washington): *Roadrunner: A high performance SBML simulation engine*
4. **Tung Nguyen** (EMBL-EBI): *Technical Updates on BioModels*
5. **Discussion**

- Matthias König (Humboldt-University Berlin); Das, Sankha (BITS Pilani, Rajasthan, India)

SBML4Humans - Interactive SBML report for Humans

The Systems Biology Markup Language (SBML) is the de facto standard for representation and exchange of mathematical models of biological systems. SBML can represent many different classes of phenomena in biology, including metabolic networks, signaling pathways, or regulatory networks, and supports models of arbitrary complexity from single processes to multi-scale models. SBML is a standard format for the representation and exchange of biological models between computers. SBML is difficult to read, comprehend and interpret by humans directly, and tools are required to provide an abstraction layer to interact with SBML objects and the relationships between them.
The objective of the project SBML4Humans was to address this issue by providing an interactive and reactive report for SBML models which will allow Humans (experts as well as novices) to easily comprehend the content of a model. SBML4Humans provides

- interactive SBML report with navigation between SBML objects
- search and filter functionality
- web application with REST API
- support for hierarchical models (comp extension)
- support for distributions and uncertainties (distrib extension)
- support for COMBINE archives

The SBML4Humans reports are available in the form of both, (i) as part of the python package sbmlutils (https://github.com/matthiaskoenig/sbmlutils) and (ii) via a web application at https://sbml4humans.de (without the need for installation). SBML4Humans provides an interactive and reactive abstraction layer around SBML as an entry point to the models without having to know SBML. Such a report for Humans is an important asset to communicate and disseminate computational models and allow one to easily understand computational models created by others. SBML4Humans was developed as part of the Google summer of Code 2021.

- Krishna Tiwari (EMBL-EBI)

Model reproducibility: implementing reproducibility scorecard and using scorecard for curation prioritization

Reproducibility in science is a key factor determining the reliability of the experiments done. In our work (Tiwari et. al, 2020), we have analysed 455 models, and found that nearly half of models could not be reproduced using the information provided in the manuscript. Learning from the reasons of non-reproducibility, we developed a simple reproducibility scorecard to improve the chances of reproducing systems biology models. This scorecard consists of 8 questions with a unit score for each answer as “Yes”. All eight questions may not be always applicable, thus, on the scale of 8, we recommend a score of 4 at the least. We showed that increase in total reproducibility score, increased the chance of model reproducibility.

For the ease of usage, we created a simple one-page reproducibility scorecard document which can be used by journal editors, model curators/developers and reviewers. Document can be downloaded from https://www.ebi.ac.uk/biomodels/reproducibility. This scorecard has been implemented in BioModels’ curation pipeline and used to prioritize curation of models with higher scores.

- Ciaran Welsh (University of Washington); Lucian Smith, Jin Xu, Herbert Sauro

Roadrunner: A high performance SBML simulation engine

Systems biologists frequently model the dynamics of biological systems using mathematical frameworks such as ordinary differential equations or Gillespie simulations. Roadrunner is a high performance, multi-language, multi-platform library written in C++ for reading, writing, creating, simulating and analysing mathematical models stored as SBML. Roadrunner compiles SBML to machine code 'on-the-fly' using LLVM as a just-in-time (JIT) compiler. This results in models with extremely fast execution times. Roadrunner supports the integration of SBML models deterministically, with or without sensitivities, using Backwards-Differentiation and Adams-Moulton for stiff and non-stiff models respectively. Roadrunner supports stochastic simulations, steady state computations and metabolic control analysis. Roadrunner language bindings make the library available from Python, Julia and C in addition to its native language, C++. Lastly, numerous "auxiliary" improvements have been made at the project level, including a new build system, a new suite of tests and a cross-platform continuous integration system for project durability and robustness.

- Tung Nguyen (EMBL-EBI)

Technical Updates on BioModels

BioModels is a leading repository of mathematical models of biological and biomedical systems. Over 21,000 unique IPs access BioModels every month. To provide seamless access to our wide user base, BioModels' web infrastructure has been migrated from virtual machines to cloud based deployments. This migration to cloud architecture across data centres allows BioModels to implement a horizontal scaling approach as well as improve its resilience to adapt for bulk and concurrent submissions. Furthermore, we have developed user centric features, improved model submission and update flow in BioModels.

14:00 – 17:00	P	**Session D - presentation room** *Moderators: Thomas Gorochowski*	Presentations

15:00 – 15:20 C **Motivation and progress towards building a common protocol representation** Presentations Room

Speakers: Jacob Beal

Jacob Beal (Raytheon BBN Technologies); Bartley, Bryan (Raytheon BBN Technologies); Beal, Jacob (Raytheon BBN Technologies); Biggers, Vanessa (Strateos); Bryce, Daniel (SIFT); Goldman, Robert P. (SIFT); Keller, Benjamin (University of Washington); Lee, Peter (Ginkgo Bioworks); Nowak, Joshua (Strateos); Rogers, Miles (Raytheon BBN Technologies); Weston, Mark (Netrias)

Laboratory protocols are critical to biological research and development, but can be difficult to communicate or reproduce due to the differences in context, skills, and resources between different projects, investigators, and organizations. Addressing this problem will require a data model for description of laboratory protocols that is unambiguous enough for precise interpretation and automation, yet simultaneously abstract enough to support reuse and adaptation.

Specifically, an effective protocol representation must be able to support at least the following goals regarding biological experimentation:
1. Execution of a protocol by either humans or machines
2. Maintaining execution records and associated metadata markup
3. Mapping protocols from one laboratory environment to another
4. Recording modifications of protocols and the relationship between different versions
5. Maintaining and improving protocols with multiple contributors over time
6. Verification and validation of protocol completeness and coherence
7. Planning, scheduling, and allocation of laboratory resources

While a number of machine-readable representations for protocols already exist, including protocols.io, Autoprotocol, Aquarium, and various hardware automation libraries, each of these serves only a small fraction of the needs above. To address this gap, we have been working together towards a common protocol representation that is general enough to address all of these needs, building on foundations from RDF, SBOL, UML, and Autoprotocol.

In this talk, we will introduce the layered architecture we have developed, addressing activities, outcomes, and interfaces, as well as prototype implementations in the form of the Protocol Activity Markup Language (PAML) and Open Protocol Interface Language (OPIL). We will discuss our use of PAML in the SD2 program for the representation, execution, and interchange of several protocols. The next stage of development is widening the pool of contributing stakeholders, completing the initial implementation of PAML, and using this as a basis for building initial tools, workflows, and interchange between organizations.

15:20 – 15:40 C **Synthetic Biology Knowledge System: Integration of near real-time curation** Presentations Room

Speakers: Jeanet Mante

Mante, Jeanet (CU Boulder); Hao, Yikai; Jett, Jacob; Joshi, Udan; Keating, Kevin; Lu, Xiang; Nakum, Gaurav; Rodriguez, Nicholas E.; Tang, Jiawei; Terry, Logan; Wu, Xuanyu; Yu, Eric; Downie, Stephen; McInnes, Bridget T.; Nguyen, Mai H.; Sepulvado, Brandon; Young, Eric M.; Myers, Chris J.

The Synthetic Biology Knowledge System (SBKS) is an instance of the SynBioHub repository that includes text and data information that has been mined from papers published in ACS Synthetic Biology. It was built to address the issue of genetic design information being spread over disparate databases and in literature without standard formatting for re-use. Whilst post-hoc data mining is still needed, SBKS is now focusing on near real-time curation. In particular, there is a need for tools to support curation of knowledge at the time of publication, effective means for knowledge enrichment to link data from a variety of sources, and the development of efficient, knowledge-enabled mechanisms to search for information during the genetic circuit design process. Addressing these needs will result in an integrated knowledge sharing resource that will accelerate biological design. This will advance computer-aided design of biological circuits, facilitate dissemination of the genetic tools we use to understand the rules of life, and support the engineering of non-conventional organisms across the tree of life.

15:40 – 16:00 C **SynBioHub3: A more intuitive and maintainable genetic design repository** Presentations Room

Speakers: Benjamin Hatch

Benjamin Hatch (University of Utah); Eric Yu, Jeanet Mante, Chris J. Myers

Synthetic biology has the potential to lead to new or more efficient production of medicines, fuels, and other important compounds. Crucial to the success of synthetic biology is effective standards for the storage and sharing of genetic design knowledge at the abstract and sequence level. In order to achieve this goal, SynBioHub, a web-based genetic design repository, was developed in order to facilitate standardized storing, searching, and sharing of genetic designs. Although SynBioHub is already utilized by many synthetic biologists and organizations, further development is necessary to meet the needs of large-scale synthetic biology projects. This development has proved to be challenging, as SynBioHub's code base is difficult to maintain due to lack of software modularity and outdated technologies. To solve this issue, SynBioHub's back and front-end architecture is being redesigned. This redesigned version of SynBioHub, which will be labeled as SynBIoHub3, will provide a more maintainable and intuitive genetic design repository that will be resilient to new functionality development and future SBOL version releases.

16:00 – 16:20 C **SED-ML Validator: tool for debugging simulation experiments** Presentations Room

Speakers: Bilal Shaikh

Shaikh, Bilal (Ichan School of Medicine at Mount Sinai); Shaikh, Bilal (Icahn School of Medicine at Mount Sinai); Freiburger, Andrew Philip (University of Victoria); Konig, Matthias (Humboldt University); Bergmann, Frank T. (Heidelberg University, California Institute of Technology); Nickerson, David P. (University of Auckland); Sauro, Herbert M. (University of Washington); Blinov, Michael L. (University of Connecticut School of Medicine); Smith, Lucian P. (University of Washington); Moraru, Ion I. (University of Connecticut School of Medicine); Karr, Jonathan R. (Icahn School of Medicine at Mount Sinai);

More sophisticated models are needed to address problems in bioscience, synthetic biology, and precision medicine. To help facilitate the collaboration needed for such models, the community developed the Simulation Experiment Description Markup Language (SED-ML), a common format for describing simulations. However, the utility of SED-ML has been hampered by limited support for SED-ML among modeling software tools and by different interpretations of SED-ML among the tools that support the format. To help modelers debug their simulations and to push the community to use SED-ML consistently, we developed a tool for validating SED-ML files. We have used the validator to correct the official SED-ML example files. We plan to use the validator to correct the files in the BioModels database so that they can be simulated. We anticipate that the validator will be a valuable tool for developing more predictive simulations and that the validator will help increase the adoption and interoperability of SED-ML.

Availability:
The validator is freely available as a webform, HTTP API, command-line program, and Python package at https://run.biosimulations.org/utils/validate and https://pypi.org/project/biosimulators-utils. The validator is also embedded into interfaces to 19 BioSimulators simulation tools. The source code is openly available at https://github.com/biosimulators/biosimulators_utils.

16:20 – 16:40 C **Reaction Rules for Whole-Cell Models** Presentations Room

Speakers: John Sekar

Sekar, John (Mt. Sinai School of Medicine); Goldberg, Arthur; Faeder, James (University of Pittsburgh); Karr, Jonathan (Mt. Sinai School of Medicine)

Whole-cell reaction models can potentially integrate experimental data that capture different aspects of a cell into a unified predictive understanding. However, representing even a fraction of cellular biochemistry using reactions poses challenges for curating, composing and simulating models. One promising approach is rule-based modeling, in which groups of similar reactions are concisely encoded as reaction rules, producing compact executable models. Current rule-based approaches either focus on site-based protein-protein interactions such as complexation and phosphorylation (e.g., BioNetGen, Kappa), or on polymerization reactions such as transcription and translation (e.g., BioCRNPyler, Pinetree). However, each biochemical domain (sites or sequences) requires a custom grammar to represent rules compactly and custom algorithms to simulate efficiently, which makes rule-based implementations difficult to extend, modify or integrate with each other. For example, it is currently infeasible to represent and simulate a gene-regulation model with both site-based and sequence-based rules.

In ongoing work, we are building an open-source modeling and simulation platform that generalizes rule-based principles to a wider range of biochemical mechanisms. To this end, we first developed attributed graphs that concisely encode molecular structures and properties using a generic schema that is extensible to current rule-based formalisms as well as novel design cases. Then, we enabled composing rules using hierarchical verb-like actions (e.g., 'translate' can be composed from 'read transcript', 'convert sequence', and 'create peptide'), which makes rules semantically transparent and readable. Next, we developed a reaction simulator that compiles rule actions to graph action primitives ('create entity', 'set attribute', etc.) and executes them. This lets us improve the simulator independent of model schemas with generic graph algorithms for canonical labeling, partitioning, connectivity and dynamic matching. To simplify the management of large models, we enabled designing models from templates, organizing model parameters into structured namespaces, and seamlessly linking models with structured data. Together, we anticipate these advances will facilitate the collaborative development of expressive biochemical models that capture multiple cellular subsystems, and ultimately enable whole-cell models.

16:40 – 17:00 D **Open discussion** Presentations Room

PINNED
18:00 – 18:15 I **COMBINE Forum Opening Session** Presentations Room

Speakers: Chris Meyers

18:00 – 21:00 P **Community Session - Monday** Presentations

Moderators: Chris Meyers

PINNED
18:15 – 19:00 I **Invited talk: Sylvia Thun: Enabling interoperability of information and processes across health domains** Presentations Room

Speakers: Sylvia Thun

PINNED
19:15 – 20:00 I **Invited talk: Stephen Larson: OpenWorm's first decade: accomplishments and the roadmap ahead** Presentations Room

Speakers: Stephen Larson

OpenWorm's mission to simulate a complete organism began in 2011 and the project has travelled a unique path scientifically and organizationally since then. This talk will focus on what has been achieved and what lies ahead for the project.

PINNED
20:15 – 21:00 I **Invited talk: Pedro Mendes: Strategies for modeling biological systems at multiple scales illustrated with iron physiology** Presentations Room

Speakers: Pedro Mendes

21:00 – 21:30 L **Lightning Talks E** Presentations Room

Speakers: Matthias König, Emilia Chen, Bryan Bartley, John Gennari, Adel Heydarabadipour

1. **Matthias König** (Humboldt-University Berlin): *Developing computational models with SBML - sbmlutils, sbmlsim, cy3sbml*
2. **Emilia Chen** (University of Cambridge/EMBL-EBI): *Immunotherapy Model Curation in BioModels*
3. **Bryan Bartley** (Raytheon BBN Technologies): *Take Your Terms from Ontologies: A Python Tool Enabling Better Annotation Practices for COMBINE Data Standards*
4. **John Gennari** (University of Washington): *Annotation-based model search across SBML and CellML*
5. **Adel Heydarabadipour** (Sharif University of Technology): *An update on Systems Biology Network Editor (SBNE): an API to render and edit the graphical representation of SBML models using their Layout and Render extensions*
6. **Discussion**

- Matthias König (Humboldt-University Berlin)

Developing computational models with SBML - sbmlutils, sbmlsim, cy3sbml

The Systems Biology Markup Language (SBML) is the de facto standard for representation and exchange of mathematical models of biological systems. SBML can represent many different classes of phenomena in biology, including metabolic networks, signaling pathways, or regulatory networks, and supports models of arbitrary complexity from single processes to multi-scale models.
A key challenge in computational modeling, especially for people new in the field, is how to encode or develop ordinary differential equation models in SBML. Here we present the python packages sbmlutils and sbmlsim and the Cytoscape app cy3sbml which address this issue by providing solutions for the programmatic creation, simulation and visualization of SBML models. A hands-on introduction will be provided in the corresponding tutorial.

[1] sbmlutils - Python utilities for SBML, https://github.com/matthiaskoenig/sbmlutils
[2] sbmlsim - SBML model simulation made easy, https://github.com/matthiaskoenig/sbmlsim
[3] cy3sbml - Visualization of SBML models in Cytoscape, https://github.com/matthiaskoenig/cy3sbml

- Emilia Chen (University of Cambridge/EMBL-EBI); Tiwari, Krishna (EMBL-EBI); Nguyen, Tung (EMBL-EBI); Sheriff, Rahuman (EMBL-EBI); Hermjakob, Henning (EMBL-EBI)

Immunotherapy Model Curation in BioModels

Immunotherapy involves exploiting the immune system to target cancers and is the focus of emerging, promising treatments against cancer. However in order to harness its potential, the understanding of underlying tumour-immune dynamics is necessary. Mathematical modelling presents an opportunity to characterise the interactions between tumour, normal and immune cells under different immunotherapy treatment conditions. This includes examples of chimeric antigen receptor T-cell therapy (e.g. Leon-Traina2021), immune checkpoint inhibition (e.g. Creemers2021) and virotherapy (e.g. Bunimovich- Mendrazitsky2007) models. Coupled with experimental work, immunotherapy mathematical models are being employed to mechanistically understand obstacles to immunotherapy. For instance, the Creemers2021 model was used to explain observations of immune checkpoint inhibition-induced dichotomous clinical outcomes. These applications of immunotherapy models makes them key weapons in our arsenal in the fight against cancer. To support immunotherapy research, I performed targeted curation of immuno-oncology models in BioModels. In this lightning talk, I will provide a quick overview of the immuno-oncology model collection.

- Bryan Bartley (Raytheon BBN Technologies)

Take Your Terms from Ontologies: A Python Tool Enabling Better Annotation Practices for COMBINE Data Standards

Adoption of annotation practices that leverage ontologies is difficult for many research groups, given the set of

specialized database and query language skills needed to interface with ontologies. Take Your Terms from Ontologies (Tyto) is a lightweight and easy-to-use Python tool that facilitates use of controlled vocabularies in everyday scripting practice. First introduced at the HARMONY 2021 session, here we present notable updates including high-level abstraction methods for performing ontological inference, a more resilient and generalized distributed systems architecture, and standardization of URIs on identifiers.org. While Tyto was originally developed for synthetic biology applications, its utility may extend to users working with ontologies in other COMBINE data standards.

- John Gennari (University of Washington); Nickerson, David; Neal, Maxwell

Annotation-based model search across SBML and CellML

With the development of the OMEX metadata specification, we have the capability to support highly detailed searches across model repositories. Since the metadata specification is language-independent, we have begun working with both the BioModels repository (using SBML) and the Physiome Model Repository (the PMR, which uses CellML) to annotate these large collections of models. To date, we have auto-generated annotation files for all 1000+ curated models in the BioModels repository, as well as hand-built annotation files for about 200 models in the PMR collection.

Based on the metadata specification, we can query for (a) specific entities, such as ChEBI or Uniprot participants, (b) names of specific processes, such as specified by GO terms, and (c) participation of entities as sources or sinks or mediators in processes. The last capability is unique to composite annotations, where multiple resources are used to provide semantics about a particular process, such as a biochemical reaction or a fluid flow. We demonstrate these capabilities across a combined resource that include the annotation files for both CellML models from the PMR, and SBML model from BioModels.

- Adel Heydarabadipour (Sharif University of Technology); Sauro, Herbert (University of Washington)

An update on Systems Biology Network Editor (SBNE): an API to render and edit the graphical representation of SBML models using their Layout and Render extensions

Systems Biology Markup Language (SBML), the de facto software-independent standard format for storing and exchanging biological models, handles the ongoing diversification of modeling approaches and community requirements with its modular extensions. The information on the graphical representation of a model is stored in its Layout extension, where the position and dimensions of each element are specified, and is further extended by the Render extension, containing additional details on the way those elements are rendered in the network representing a model. To save developers the arduous and burdensome task of manually encoding Layout and Render information about an SBML model, we have developed a higher-level API which enables them to read, manually create, or automatically generate; modify; and write straightforwardly the layout and render features to an SBML model. The original work was presented in COMBINE 2020, and here we provide an update on it where its auto-layout and auto-render features are significantly improved and also supplemented with Graphviz graph drawing algorithms to render a clear illustration of more complicated biological networks. In addition, its GUI is now equipped with an “Item features menu” side panel allowing a user to change any information on the graphical representation of a model visually. It is now much easier to get/set the Layout and Render extension features of an SBML model through the API portable library as well. The portable library is originally written in C/C++ and currently provides language bindings for Python using the wrapper generator SWIG. The source code and its binary installers for Microsoft Windows, macOS, and Linux as well as its complete documentation are available online on GitHub (https://github.com/adelhpour/SBNE) and readthedocs (https://sbne.readthedocs.io /en/latest), respectively.

...

21:00 – 00:00	P	**Session E - presentation room**	Presentations

22:00 – 22:20 C **KG-Microbe: a reference knowledge-graph and platform for harmonized microbial information**
Speakers: marcin pawel joachimiak Presentations Room
marcin pawel joachimiak (Lawrence Berkeley National Laboratory); Harshad Hegde, William D. Duncan, Luca Cappelletti, Justin T. Reese, Anne E. Thessen, Christopher J. Mungall

Microorganisms (microbes) are incredibly diverse, spanning all major divisions of life, and represent the greatest fraction of known species. A vast amount of knowledge about microbes is available in the literature, across experimental datasets, and in established data resources. While the genomic and biochemical pathway data about microbes is well-structured and annotated using standard ontologies, broader information about microbes and their ecological traits is not. We created the KG-Microbe (https://github.com/Knowledge-Graph-Hub/kg-microbe) resource in order to extract and integrate diverse knowledge about microbes from a variety of structured and unstructured sources. Initially, we are harmonizing and linking prokaryotic data for phenotypic traits, taxonomy, functions, chemicals, and environment descriptors, to construct a knowledge graph with over 266,000 entities linked by 432,000 relations. The effort is supported by a knowledge graph construction platform (KG-Hub) for rapid development of knowledge graphs using available data, knowledge modeling principles, and software tools. As part of this framework, we rely on the LinkML linked data modeling language and the Biolink Model, a high level data model for biological entities and relations. KG-Microbe is a microbe-centric Knowledge Graph (KG) to support tasks such as querying and graph link prediction in many use cases including microbiology, biomedicine, and the environment. KG-Microbe fulfills a need for standardized and linked microbial data, allowing the broader community to contribute, query, and enrich analyses and algorithms.

22:20 – 22:40 C **Vivarium: an interface and engine for integrative multi-scale simulation** Presentations Room
Speakers: Eran Agmon
Agmon, Eran (Stanford University); Covert, Markus (Stanford University)

Vivarium is a Pythonic software tool for composing multi-scale simulations from multiple input simulators and models. It provides an interface that makes individual models into modules that can be wired together in large composite models, parallelized across multiple CPUs, and run with Vivarium's discrete-event simulation engine. Vivarium's utility is demonstrated by building composite models that combine several modeling frameworks: agent based models, ordinary differential equations, stochastic reaction systems, constraint-based models, solid-body physics, and spatial diffusion. Some immediate applications for the COMBINE community include composing standardized simulators, such as those provided by BioSimulators.

22:40 – 23:00 C **COYOTE: An Extensible Python-Based Reaction Layout Tool** Presentations Room

Speakers: Jin Xu

Xu, Jin (University of Washington); Geng, Gary; Nguyen, Nhan; Perena-Cortes, Carmen; Samuels, Claire; Sauro, Herbert

We have recently developed an open-source (MIT license) Python-based cross-platform reaction viewer and editor called Coyote. The tool uses wxPython to implement the GUI and support the drawing canvas. It supports the visualization of compartments, species, reactions, and modifiers. We include many options to stylize each of these components. In addition, species can be made from more primitive shapes to create composite nodes. Other features include network zooming, as well as an interactive bird-eye view of the network to allow easy navigation on large networks.

A unique feature of the tool is the extensive plugin API (fully documented) where third-party developers can include new functionality. Example plugins include an alignment tool, autolayout of networks, circularization of nodes, and random network generator to name but a few. Plugins are stored at a GitHub repository and an included plugin manager that can retrieve and install new plugins from the repository on demand. Plugins have version metadata associated with them to make it easy to install new versions. Of particular interest are the SBML import and export plugins that support the SMBL layout and render standard. We will illustrate the exchange of the layout/render format with COPASI. More recently Claire Samuels has developed an experimental plugin that allows simulations to be carried out from within Coyote. More importantly, plugins have access to the canvas by registering with the core OnPaint event. This allows plugins to add further visualization features. For example, the simulator plugin provides an example of visualizing concentrations in real-time on the drawing canvas. This will be demonstrated at the workshop. Availability: https://github.com/evilnose/PyRKViewer.

23:00 – 23:20 C **BioSimulations: Sharing and reusing biomodels, simulations, and visualizations** Presentations Room

Speakers: Bilal Shaikh

Shaikh, Bilal (Icahn School of Medicine at Mount Sinai); Marupilla, Gnaneswara (UConn Health); Wilson, Mike (UConn Health); Blinov, Michael L. (UConn Health); Moraru, Ion I (UConn Health); Karr, Jonathan R (Icahn School of Medicine at Mount Sinai)

More comprehensive and more predictive models of cells could enable the advancement of biology, medicine, and bioengineering. Such models will require collaboration between many investigators and domains. In turn, models must be shareable, executable, and accessible. While many community resources exist, such as BiGG, BNGL, BioModels, CellML, KiSAO, NeuroML, OMEX, SBML, and SED-ML, significant barriers to collaboration remain. Different types of models remain siloed across multiple repositories, modeling frameworks, model formats, simulation algorithms, and simulation tools. For example, flux balance models are often obtained from BiGG and simulated with COBRApy, while deterministic kinetic models are often obtained from BioModels and simulated with tellurium. Future, more comprehensive models will likely require capturing biology at multiple scales, requiring the ability to combine frameworks, formats, algorithms, and tools. Additionally, installation and usage of tools cam be difficult. To address these challenges, we have developed BioSimulations (https://run.biosimulations.org), a cloud platform for reproducing and reusing models with ease. BioSimulations will provide users a unified platform to share, discover, run, and visualize biological models across various modeling scales, formats, frameworks, and algorithms.

Recent advances

Recent improvements to the platform have significantly increased the range of supported simulations to 10 model languages, 20 simulation tools, 70 algorithms, and several modeling frameworks including continuous and discrete kinetic, logical, flux balance, spatial, particle-based, and hybrid modeling. To help investigators navigate these tools, the platform can now automatically inspect models and recommend specific algorithms and tools for executing them. Large-scale simulation results up to 5 TB per project can now be saved and queried using the Highly Scalable Data Service (HSDS). Users can also use Vega to visualize simulation results with advanced, interactive diagrams, such as activity flow diagrams, process description maps, and reaction flux maps which can be created with a variety of tools such as Escher, GINSim, and Newt. Furthermore, BioSimulations makes it easy for investigators to reuse such visualizations across multiple projects, such as to compare multiple models or simulations.

Future directions

We are developing a repository of entire simulation projects, including interactive visualizations of results.

23:20 – 00:00 D **Open discussion** Presentations Room

B Breakout C Contributed talk D Discussion I Invited talk L Lightning talk P Presentation session T Tutorial

OCTOBER 12 • TUESDAY

PINNED
01:00 – 01:45 I **Replay: Invited talk: Anna Niaraki: Do you speak Systems Biology? Shaping a common language for community work** Presentations Room
Speakers: Anna Niarakis

01:00 – 04:00 P **Replay Session C - presentation room** Presentations
Moderators: David Nickerson

02:00 – 02:20 C **Replay: A physiological-based model of indocyanine green for the evaluation of liver function based on COMBINE standards** Presentations Room
Speakers: Matthias König
König, Matthias (Humboldt-University Berlin); Köller, Adrian (Humboldt-University Berlin); Grzegorzewski, Jan (Humboldt-University Berlin)

02:20 – 02:40 C **Replay: Fides: Reliable Trust-Region Optimization for Parameter Estimation of Ordinary Differential Equation Models** Presentations Room
Speakers: Fabian Froehlich
Froehlich, Fabian (Harvard Medical School)

Motivation: Because they effectively represent mass action kinetics, ordinary differential equation models are widely used to describe biochemical processes. Optimization-based calibration of these models on experimental data can be challenging, even for low-dimensional problems. However, reliable model calibration is a prerequisite for many subsequent analysis steps, including uncertainty analysis, model selection and biological interpretation. Although multiple hypothesis have been advanced to explain why optimization based calibration of biochemical models is challenging, there are few comprehensive studies that test these hypothesis and tools for performing such studies are also lacking.

Results: We implemented an established trust-region method as a modular python framework (fides) to enable structured comparison of different approaches to ODE model calibration involving Hessian approximation schemes and trust-region subproblem solvers. We evaluate fides on a set of PEtab benchmark problems that include experimental data. We find a high variability in optimizer performance among different implementations of the same algorithm, with fides performing more reliably that other implementations investigated. Our investigation of possible sources of poor optimizer performance identify shortcomings in the widely used Gauss-Newton approximation. We address these shortcomings by proposing a novel hybrid Hessian approximation scheme that enhances optimizer performance.

02:40 – 03:00 C **Replay: Interfacing SBML and CellML with the scientific machine learning and constrained-based modeling ecosystems in Julia** Presentations Room

Speakers: Paul F. Lang

Lang, Paul F (University of Oxford); Iravanian, Shahriar; Kratochvil Miroslav; Jain, Anand; Giordano, Mose; Rackauckas, Chris

Julia is a general purpose programming language that was designed for simplifying and accelerating numerical analysis and computational science. Building on the computer algebra system Symbolics.jl, many high-performing solvers for constrained systems and differential equations have been developed in Julia. In particular the Scientific Machine Learning (SciML) ecosystem of Julia packages includes ModelingToolkit.jl, a modeling framework for high-performance symbolic-numeric computation in scientific computing and scientific machine learning. It allows for users to give a high-level description of a model for symbolic preprocessing to analyze and enhance the model. ModelingToolkit can automatically generate fast functions for model components, along with automatically sparsifying and parallelizing the computations. This enabled highly performing solvers of differential equations, parameter optimisation algorithms and methodologies for automated model discovery. To give the systems biology community easy access to SciML, we developed a set of three tools that import biochemical models specified in the SBML or CellML community standards into the SciML ecosystem: SBML.jl, SBMLToolkit.jl and CellMLToolkit.jl SBML.jl is a Julia wrapper for the main functionality of the libSBML library via its C language bindings. It also enables constrained-based modelling by supporting the Flux Balance Constraints SBML package. SBMLToolkit.jl and CellMLToolkit.jl import dynamic SBML and CellML models, respectively, into the SciML ecosystem. We show that SBML.jl alone is sufficient for backing up a facile reconstruction of large-scale constraint-based metabolic models and that combined community models that consist of tens of millions individual reactions may be constructed interactively on off-the-shelf hardware, without requiring any special programming techniques or other support libraries. We further demonstrate how the SciML tools for dynamic systems can be used to accelerate model simulation and fitting of kinetic parameters. We will also provide an example of filling mechanistic knowledge gaps in epidemiological SBML and CellML toy models using neural networks and sparse identification of nonlinear dynamics. By providing the computational systems biologists with easy access to the open-source Julia ecosystem we hope to catalyse the development of further Julia tools in this domain, and the growth of the Julia bioscience community.

03:00 – 03:20 C **Replay: Connecting computational biology models and COVID-19 resources** Presentations Room

Speakers: Lea Gütebier

Gütebier, Lea (Medical Informatics Laboratory, University Medicine Greifswald); Henkel, Ron (University Medicine Greifswald); Bleimehl, Tim (German Center for Diabetes Research); Preusse, Martin (Kaiser & Preusse); Waltemath, Dagmar (University Medicine Greifswald)

Within a short time, a lot of data on COVID-19 has been published from different research groups. Integrating this data is essential for the better understanding of the disease.

In 2020 HealthECCO (https://healthecco.org/) initiated the CovidGraph project [1] to support the exploration of COVID-19 data. The CovidGraph, a Neo4j database, integrates various data sets regarding COVID-19 and the corona virus family in a single knowledge graph. It includes data from different domains such as publications, patents, biological entities, clinical studies and more.
Another graph-based project aiming at better integration of biomedical research data is MaSyMoS [2]. MaSyMoS stores computational models, simulation descriptions and associated meta-data, including reference publications and terms from bio-ontologies in a Neo4j graph database. BioModels serves as one of the data sources, including a highly curated collection of COVID-19 studies.

We integrated MaSyMoS as an additional data provider into the CovidGraph. Therefore, we matched relevant publications from MaSyMoS and CovidGraph and, in addition, terms from biomedical ontologies.
More specifically, our workflow matches the corresponding IDs (e.g., PubMedID) of publications from the COVID-19 Open Research Dataset contained in CovidGraph, and publications from PubMed contained in MaSyMoS. Furthermore, terms from ontologies relevant to COVID-19, and information about biological entities are stored in the biomedical data domain of the CovidGraph. We mapped overlapping ontology terms from both, MaSyMoS annotations and the CovidGraph, by their corresponding IDs. The overlap includes the Gene Ontology, ChEBI, UniProt and Disease Ontology.
The CovidGraph represents a network of integrated data providers and allows for data exploration across different domains, particularly medical research and systems biology.

In our talk we present the CovidGraph and explain how we integrated the MaSyMoS data set. We outline the developed integration concept to show that further data resources can easily be integrated to extend the CovidGraph.

The COVID-19 collection on BioModels was developed as part of Grant: European Commission: EOSCsecretariat.eu - EOSCsecretariat.eu (831644).

[1] Gütebier et al. “COVIDGraph: Connecting biomedical COVID-19 resources and computational biology models.” SEA Data 2021
[2] Henkel et al. “Combining computational models, semantic annotations and simulation experiments in a graph database.” Database

03:20 – 04:00 D **Open discussion** Presentations Room

06:00 – 06:30 L **Replay: Lightning Talks D** Presentations Room

Speakers: Matthias König, Krishna Tiwari, Ciaran Welsh, Tung Nguyen

1. **Matthias König** (Humboldt-University Berlin): *SBML4Humans - Interactive SBML report for Humans*
2. **Krishna Tiwari** (EMBL-EBI): *Model reproducibility: implementing reproducibility scorecard and using scorecard for curation prioritization*
3. **Ciaran Welsh** (University of Washington): *Roadrunner: A high performance SBML simulation engine*
4. **Tung Nguyen** (EMBL-EBI): *Technical Updates on BioModels*
5. **Discussion**

- Matthias König (Humboldt-University Berlin); Das, Sankha (BITS Pilani, Rajasthan, India)

SBML4Humans - Interactive SBML report for Humans

The Systems Biology Markup Language (SBML) is the de facto standard for representation and exchange of mathematical models of biological systems. SBML can represent many different classes of phenomena in biology, including metabolic networks, signaling pathways, or regulatory networks, and supports models of arbitrary complexity from single processes to multi-scale models. SBML is a standard format for the representation and exchange of biological models between computers. SBML is difficult to read, comprehend and interpret by humans directly, and tools are required to provide an abstraction layer to interact with SBML objects and the relationships between them.
The objective of the project SBML4Humans was to address this issue by providing an interactive and reactive report for SBML models which will allow Humans (experts as well as novices) to easily comprehend the content of a model. SBML4Humans provides

- interactive SBML report with navigation between SBML objects
- search and filter functionality
- web application with REST API
- support for hierarchical models (comp extension)
- support for distributions and uncertainties (distrib extension)
- support for COMBINE archives

The SBML4Humans reports are available in the form of both, (i) as part of the python package sbmlutils (https://github.com/matthiaskoenig/sbmlutils) and (ii) via a web application at https://sbml4humans.de (without the need for installation). SBML4Humans provides an interactive and reactive abstraction layer around SBML as an entry point to the models without having to know SBML. Such a report for Humans is an important asset to communicate and disseminate computational models and allow one to easily understand computational models created by others. SBML4Humans was developed as part of the Google summer of Code 2021.

- Krishna Tiwari (EMBL-EBI)

Model reproducibility: implementing reproducibility scorecard and using scorecard for curation prioritization

Reproducibility in science is a key factor determining the reliability of the experiments done. In our work (Tiwari et. al, 2020), we have analysed 455 models, and found that nearly half of models could not be reproduced using the information provided in the manuscript. Learning from the reasons of non-reproducibility, we developed a simple reproducibility scorecard to improve the chances of reproducing systems biology models. This scorecard consists of 8 questions with a unit score for each answer as “Yes”. All eight questions may not be always applicable, thus, on the scale of 8, we recommend a score of 4 at the least. We showed that increase in total reproducibility score, increased the chance of model reproducibility.

For the ease of usage, we created a simple one-page reproducibility scorecard document which can be used by journal editors, model curators/developers and reviewers. Document can be downloaded from https://www.ebi.ac.uk/biomodels/reproducibility. This scorecard has been implemented in BioModels’ curation pipeline and used to prioritize curation of models with higher scores.

- Ciaran Welsh (University of Washington); Lucian Smith, Jin Xu, Herbert Sauro

Roadrunner: A high performance SBML simulation engine

Systems biologists frequently model the dynamics of biological systems using mathematical frameworks such as ordinary differential equations or Gillespie simulations. Roadrunner is a high performance, multi-language, multi-platform library written in C++ for reading, writing, creating, simulating and analysing mathematical models stored as SBML. Roadrunner compiles SBML to machine code 'on-the-fly' using LLVM as a just-in-time (JIT) compiler. This results in models with extremely fast execution times. Roadrunner supports the integration of SBML models deterministically, with or without sensitivities, using Backwards-Differentiation and Adams-Moulton for stiff and non-stiff models respectively. Roadrunner supports stochastic simulations, steady state computations and metabolic control analysis. Roadrunner language bindings make the library available from Python, Julia and C in addition to its native language, C++. Lastly, numerous "auxiliary" improvements have been made at the project level, including a new build system, a new suite of tests and a cross-platform continuous integration system for project durability and robustness.

- Tung Nguyen (EMBL-EBI)

Technical Updates on BioModels

BioModels is a leading repository of mathematical models of biological and biomedical systems. Over 21,000 unique IPs access BioModels every month. To provide seamless access to our wide user base, BioModels' web infrastructure has been migrated from virtual machines to cloud based deployments. This migration to cloud architecture across data centres allows BioModels to implement a horizontal scaling approach as well as improve its resilience to adapt for bulk and concurrent submissions. Furthermore, we have developed user centric features, improved model submission and update flow in BioModels.

06:00 – 09:00 P **Replay Session D - presentation room** Presentations
Moderators: Thomas Gorochowski

07:00 – 07:20 C **Replay: Motivation and progress towards building a common protocol representation**

Speakers: Jacob Beal Presentations Room

Jacob Beal (Raytheon BBN Technologies); Bartley, Bryan (Raytheon BBN Technologies); Beal, Jacob (Raytheon BBN Technologies); Biggers, Vanessa (Strateos); Bryce, Daniel (SIFT); Goldman, Robert P. (SIFT); Keller, Benjamin (University of Washington); Lee, Peter (Ginkgo Bioworks); Nowak, Joshua (Strateos); Rogers, Miles (Raytheon BBN Technologies); Weston, Mark (Netrias)

Laboratory protocols are critical to biological research and development, but can be difficult to communicate or reproduce due to the differences in context, skills, and resources between different projects, investigators, and organizations. Addressing this problem will require a data model for description of laboratory protocols that is unambiguous enough for precise interpretation and automation, yet simultaneously abstract enough to support reuse and adaptation.

Specifically, an effective protocol representation must be able to support at least the following goals regarding biological experimentation:
1. Execution of a protocol by either humans or machines
2. Maintaining execution records and associated metadata markup
3. Mapping protocols from one laboratory environment to another
4. Recording modifications of protocols and the relationship between different versions
5. Maintaining and improving protocols with multiple contributors over time
6. Verification and validation of protocol completeness and coherence
7. Planning, scheduling, and allocation of laboratory resources

While a number of machine-readable representations for protocols already exist, including protocols.io, Autoprotocol, Aquarium, and various hardware automation libraries, each of these serves only a small fraction of the needs above. To address this gap, we have been working together towards a common protocol representation that is general enough to address all of these needs, building on foundations from RDF, SBOL, UML, and Autoprotocol.

In this talk, we will introduce the layered architecture we have developed, addressing activities, outcomes, and interfaces, as well as prototype implementations in the form of the Protocol Activity Markup Language (PAML) and Open Protocol Interface Language (OPIL). We will discuss our use of PAML in the SD2 program for the representation, execution, and interchange of several protocols. The next stage of development is widening the pool of contributing stakeholders, completing the initial implementation of PAML, and using this as a basis for building initial tools, workflows, and interchange between organizations.

07:20 – 07:40 C **Replay: Synthetic Biology Knowledge System: Integration of near real-time curation**

Speakers: Jeanet Mante Presentations Room

Mante, Jeanet (CU Boulder); Hao, Yikai; Jett, Jacob; Joshi, Udan; Keating, Kevin; Lu, Xiang; Nakum, Gaurav; Rodriguez, Nicholas E.; Tang, Jiawei; Terry, Logan; Wu, Xuanyu; Yu, Eric; Downie, Stephen; McInnes, Bridget T.; Nguyen, Mai H.; Sepulvado, Brandon; Young, Eric M.; Myers, Chris J.

The Synthetic Biology Knowledge System (SBKS) is an instance of the SynBioHub repository that includes text and data information that has been mined from papers published in ACS Synthetic Biology. It was built to address the issue of genetic design information being spread over disparate databases and in literature without standard formatting for re-use. Whilst post-hoc data mining is still needed, SBKS is now focusing on near real-time curation. In particular, there is a need for tools to support curation of knowledge at the time of publication, effective means for knowledge enrichment to link data from a variety of sources, and the development of efficient, knowledge-enabled mechanisms to search for information during the genetic circuit design process. Addressing these needs will result in an integrated knowledge sharing resource that will accelerate biological design. This will advance computer-aided design of biological circuits, facilitate dissemination of the genetic tools we use to understand the rules of life, and support the engineering of non-conventional organisms across the tree of life.

07:40 – 08:00 C **Replay: SynBioHub3: A more intuitive and maintainable genetic design repository**
Speakers: Benjamin Hatch Presentations Room
Benjamin Hatch (University of Utah); Eric Yu, Jeanet Mante, Chris J. Myers

Synthetic biology has the potential to lead to new or more efficient production of medicines, fuels, and other important compounds. Crucial to the success of synthetic biology is effective standards for the storage and sharing of genetic design knowledge at the abstract and sequence level. In order to achieve this goal, SynBioHub, a web-based genetic design repository, was developed in order to facilitate standardized storing, searching, and sharing of genetic designs. Although SynBioHub is already utilized by many synthetic biologists and organizations, further development is necessary to meet the needs of large-scale synthetic biology projects. This development has proved to be challenging, as SynBioHub's code base is difficult to maintain due to lack of software modularity and outdated technologies. To solve this issue, SynBioHub's back and front-end architecture is being redesigned. This redesigned version of SynBioHub, which will be labeled as SynBIoHub3, will provide a more maintainable and intuitive genetic design repository that will be resilient to new functionality development and future SBOL version releases.

08:00 – 08:20 C **Replay: SED-ML Validator: tool for debugging simulation experiments** Presentations Room
Speakers: Bilal Shaikh
Shaikh, Bilal (Ichan School of Medicine at Mount Sinai); Shaikh, Bilal (Icahn School of Medicine at Mount Sinai); Freiburger, Andrew Philip (University of Victoria); Konig, Matthias (Humboldt University); Bergmann, Frank T. (Heidelberg University, California Institute of Technology); Nickerson, David P. (University of Auckland); Sauro, Herbert M. (University of Washington); Blinov, Michael L. (University of Connecticut School of Medicine); Smith, Lucian P. (University of Washington); Moraru, Ion I. (University of Connecticut School of Medicine); Karr, Jonathan R. (Icahn School of Medicine at Mount Sinai);

More sophisticated models are needed to address problems in bioscience, synthetic biology, and precision medicine. To help facilitate the collaboration needed for such models, the community developed the Simulation Experiment Description Markup Language (SED-ML), a common format for describing simulations. However, the utility of SED-ML has been hampered by limited support for SED-ML among modeling software tools and by different interpretations of SED-ML among the tools that support the format. To help modelers debug their simulations and to push the community to use SED-ML consistently, we developed a tool for validating SED-ML files. We have used the validator to correct the official SED-ML example files. We plan to use the validator to correct the files in the BioModels database so that they can be simulated. We anticipate that the validator will be a valuable tool for developing more predictive simulations and that the validator will help increase the adoption and interoperability of SED-ML.

Availability:
The validator is freely available as a webform, HTTP API, command-line program, and Python package at https://run.biosimulations.org/utils/validate and https://pypi.org/project/biosimulators-utils. The validator is also embedded into interfaces to 19 BioSimulators simulation tools. The source code is openly available at https://github.com/biosimulators/biosimulators_utils.

08:20 – 08:40 C **Replay: Reaction Rules for Whole-Cell Models** Presentations Room

Speakers: John Sekar

Sekar, John (Mt. Sinai School of Medicine); Goldberg, Arthur; Faeder, James (University of Pittsburgh); Karr, Jonathan (Mt. Sinai School of Medicine)

Whole-cell reaction models can potentially integrate experimental data that capture different aspects of a cell into a unified predictive understanding. However, representing even a fraction of cellular biochemistry using reactions poses challenges for curating, composing and simulating models. One promising approach is rule-based modeling, in which groups of similar reactions are concisely encoded as reaction rules, producing compact executable models. Current rule-based approaches either focus on site-based protein-protein interactions such as complexation and phosphorylation (e.g., BioNetGen, Kappa), or on polymerization reactions such as transcription and translation (e.g., BioCRNPyler, Pinetree). However, each biochemical domain (sites or sequences) requires a custom grammar to represent rules compactly and custom algorithms to simulate efficiently, which makes rule-based implementations difficult to extend, modify or integrate with each other. For example, it is currently infeasible to represent and simulate a gene-regulation model with both site-based and sequence-based rules.

In ongoing work, we are building an open-source modeling and simulation platform that generalizes rule-based principles to a wider range of biochemical mechanisms. To this end, we first developed attributed graphs that concisely encode molecular structures and properties using a generic schema that is extensible to current rule-based formalisms as well as novel design cases. Then, we enabled composing rules using hierarchical verb-like actions (e.g., ‘translate’ can be composed from ‘read transcript’, ‘convert sequence’, and ‘create peptide’), which makes rules semantically transparent and readable. Next, we developed a reaction simulator that compiles rule actions to graph action primitives (‘create entity’, ‘set attribute’, etc.) and executes them. This lets us improve the simulator independent of model schemas with generic graph algorithms for canonical labeling, partitioning, connectivity and dynamic matching. To simplify the management of large models, we enabled designing models from templates, organizing model parameters into structured namespaces, and seamlessly linking models with structured data. Together, we anticipate these advances will facilitate the collaborative development of expressive biochemical models that capture multiple cellular subsystems, and ultimately enable whole-cell models.

08:40 – 09:00 D **Open discussion** Presentations Room

10:00 – 10:30 L **Replay: Lightning Talks E** Presentations Room

Speakers: Matthias König, Emilia Chen, John Gennari, Bryan Bartley

1. **Matthias König** (Humboldt-University Berlin): *Developing computational models with SBML - sbmlutils, sbmlsim, cy3sbml*
2. **Emilia Chen** (University of Cambridge/EMBL-EBI): *Immunotherapy Model Curation in BioModels*
3. **Bryan Bartley** (Raytheon BBN Technologies): *Take Your Terms from Ontologies: A Python Tool Enabling Better Annotation Practices for COMBINE Data Standards*
4. **John Gennari** (University of Washington): *Annotation-based model search across SBML and CellML*
5. **Discussion**

- Matthias König (Humboldt-University Berlin)

Developing computational models with SBML - sbmlutils, sbmlsim, cy3sbml

The Systems Biology Markup Language (SBML) is the de facto standard for representation and exchange of mathematical models of biological systems. SBML can represent many different classes of phenomena in biology, including metabolic networks, signaling pathways, or regulatory networks, and supports models of arbitrary complexity from single processes to multi-scale models.
A key challenge in computational modeling, especially for people new in the field, is how to encode or develop ordinary differential equation models in SBML. Here we present the python packages sbmlutils and sbmlsim and the Cytoscape app cy3sbml which address this issue by providing solutions for the programmatic creation, simulation and visualization of SBML models. A hands-on introduction will be provided in the corresponding tutorial.

[1] sbmlutils - Python utilities for SBML, https://github.com/matthiaskoenig/sbmlutils
[2] sbmlsim - SBML model simulation made easy, https://github.com/matthiaskoenig/sbmlsim
[3] cy3sbml - Visualization of SBML models in Cytoscape, https://github.com/matthiaskoenig/cy3sbml

- Emilia Chen (University of Cambridge/EMBL-EBI); Tiwari, Krishna (EMBL-EBI); Nguyen, Tung (EMBL-EBI); Sheriff, Rahuman (EMBL-EBI); Hermjakob, Henning (EMBL-EBI)

Immunotherapy Model Curation in BioModels

Immunotherapy involves exploiting the immune system to target cancers and is the focus of emerging, promising treatments against cancer. However in order to harness its potential, the understanding of underlying tumour-immune dynamics is necessary. Mathematical modelling presents an opportunity to characterise the interactions between tumour, normal and immune cells under different immunotherapy treatment conditions. This includes examples of chimeric antigen receptor T-cell therapy (e.g. Leon-Traina2021), immune checkpoint inhibition (e.g. Creemers2021) and virotherapy (e.g. Bunimovich- Mendrazitsky2007) models. Coupled with experimental work, immunotherapy mathematical models are being employed to mechanistically understand obstacles to immunotherapy. For instance, the Creemers2021 model was used to explain observations of immune checkpoint inhibition-induced dichotomous clinical outcomes. These applications of immunotherapy models makes them key weapons in our arsenal in the fight against cancer. To support immunotherapy research, I performed targeted curation of immuno-oncology models in BioModels. In this lightning talk, I will provide a quick overview of the immuno-oncology model collection.

- Bryan Bartley (Raytheon BBN Technologies)

Take Your Terms from Ontologies: A Python Tool Enabling Better Annotation Practices for COMBINE Data Standards

Adoption of annotation practices that leverage ontologies is difficult for many research groups, given the set of specialized database and query language skills needed to interface with ontologies. Take Your Terms from Ontologies (Tyto) is a lightweight and easy-to-use Python tool that facilitates use of controlled vocabularies in

everyday scripting practice. First introduced at the HARMONY 2021 session, here we present notable updates including high-level abstraction methods for performing ontological inference, a more resilient and generalized distributed systems architecture, and standardization of URIs on identifiers.org. While Tyto was originally developed for synthetic biology applications, its utility may extend to users working with ontologies in other COMBINE data standards.

- John Gennari (University of Washington); Nickerson, David; Neal, Maxwell

Annotation-based model search across SBML and CellML

With the development of the OMEX metadata specification, we have the capability to support highly detailed searches across model repositories. Since the metadata specification is language-independent, we have begun working with both the BioModels repository (using SBML) and the Physiome Model Repository (the PMR, which uses CellML) to annotate these large collections of models. To date, we have auto-generated annotation files for all 1000+ curated models in the BioModels repository, as well as hand-built annotation files for about 200 models in the PMR collection.

Based on the metadata specification, we can query for (a) specific entities, such as ChEBI or Uniprot participants, (b) names of specific processes, such as specified by GO terms, and (c) participation of entities as sources or sinks or mediators in processes. The last capability is unique to composite annotations, where multiple resources are used to provide semantics about a particular process, such as a biochemical reaction or a fluid flow. We demonstrate these capabilities across a combined resource that include the annotation files for both CellML models from the PMR, and SBML model from BioModels.

10:00 – 13:00 P **Replay Session E - presentation room** Presentations
Speakers: Padraig Gleeson

11:00 – 11:20 C **Replay: KG-Microbe: a reference knowledge-graph and platform for harmonized microbial information** Presentations Room
Speakers: marcin pawel joachimiak
marcin pawel joachimiak (Lawrence Berkeley National Laboratory); Harshad Hegde, William D. Duncan, Luca Cappelletti, Justin T. Reese, Anne E. Thessen, Christopher J. Mungall

Microorganisms (microbes) are incredibly diverse, spanning all major divisions of life, and represent the greatest fraction of known species. A vast amount of knowledge about microbes is available in the literature, across experimental datasets, and in established data resources. While the genomic and biochemical pathway data about microbes is well-structured and annotated using standard ontologies, broader information about microbes and their ecological traits is not. We created the KG-Microbe (https://github.com/Knowledge-Graph-Hub/kg-microbe) resource in order to extract and integrate diverse knowledge about microbes from a variety of structured and unstructured sources. Initially, we are harmonizing and linking prokaryotic data for phenotypic traits, taxonomy, functions, chemicals, and environment descriptors, to construct a knowledge graph with over 266,000 entities linked by 432,000 relations. The effort is supported by a knowledge graph construction platform (KG-Hub) for rapid development of knowledge graphs using available data, knowledge modeling principles, and software tools. As part of this framework, we rely on the LinkML linked data modeling language and the Biolink Model, a high level data model for biological entities and relations. KG-Microbe is a microbe-centric Knowledge Graph (KG) to support tasks such as querying and graph link prediction in many use cases including microbiology, biomedicine, and the environment. KG-Microbe fulfills a need for standardized and linked microbial data, allowing the broader community to contribute, query, and enrich analyses and algorithms.

11:20 – 11:40 C **Replay: Vivarium: an interface and engine for integrative multi-scale simulation** Presentations Room

Speakers: Eran Agmon

Agmon, Eran (Stanford University); Covert, Markus (Stanford University)

Vivarium is a Pythonic software tool for composing multi-scale simulations from multiple input simulators and models. It provides an interface that makes individual models into modules that can be wired together in large composite models, parallelized across multiple CPUs, and run with Vivarium's discrete-event simulation engine. Vivarium's utility is demonstrated by building composite models that combine several modeling frameworks: agent based models, ordinary differential equations, stochastic reaction systems, constraint-based models, solid-body physics, and spatial diffusion. Some immediate applications for the COMBINE community include composing standardized simulators, such as those provided by BioSimulators.

11:40 – 12:00 C **Replay: COYOTE: An Extensible Python-Based Reaction Layout Tool** Presentations Room

Speakers: Jin Xu

Xu, Jin (University of Washington); Geng, Gary; Nguyen, Nhan; Perena-Cortes, Carmen; Samuels, Claire; Sauro, Herbert

We have recently developed an open-source (MIT license) Python-based cross-platform reaction viewer and editor called Coyote. The tool uses wxPython to implement the GUI and support the drawing canvas. It supports the visualization of compartments, species, reactions, and modifiers. We include many options to stylize each of these components. In addition, species can be made from more primitive shapes to create composite nodes. Other features include network zooming, as well as an interactive bird-eye view of the network to allow easy navigation on large networks.

A unique feature of the tool is the extensive plugin API (fully documented) where third-party developers can include new functionality. Example plugins include an alignment tool, autolayout of networks, circularization of nodes, and random network generator to name but a few. Plugins are stored at a GitHub repository and an included plugin manager that can retrieve and install new plugins from the repository on demand. Plugins have version metadata associated with them to make it easy to install new versions. Of particular interest are the SBML import and export plugins that support the SMBL layout and render standard. We will illustrate the exchange of the layout/render format with COPASI. More recently Claire Samuels has developed an experimental plugin that allows simulations to be carried out from within Coyote. More importantly, plugins have access to the canvas by registering with the core OnPaint event. This allows plugins to add further visualization features. For example, the simulator plugin provides an example of visualizing concentrations in real-time on the drawing canvas. This will be demonstrated at the workshop. Availability: https://github.com/evilnose/PyRKViewer.

12:00 – 12:20 C **Replay: BioSimulations: Sharing and reusing biomodels, simulations, and visualizations**

Speakers: Bilal Shaikh Presentations Room

Shaikh, Bilal (Icahn School of Medicine at Mount Sinai); Marupilla, Gnaneswara (UConn Health); Wilson, Mike (UConn Health); Blinov, Michael L. (UConn Health); Moraru, Ion I (UConn Health); Karr, Jonathan R (Icahn School of Medicine at Mount Sinai)

More comprehensive and more predictive models of cells could enable the advancement of biology, medicine, and bioengineering. Such models will require collaboration between many investigators and domains. In turn, models must be shareable, executable, and accessible. While many community resources exist, such as BiGG, BNGL, BioModels, CellML, KiSAO, NeuroML, OMEX, SBML, and SED-ML, significant barriers to collaboration remain. Different types of models remain siloed across multiple repositories, modeling frameworks, model formats, simulation algorithms, and simulation tools. For example, flux balance models are often obtained from BiGG and simulated with COBRApy, while deterministic kinetic models are often obtained from BioModels and simulated with tellurium. Future, more comprehensive models will likely require capturing biology at multiple scales, requiring the ability to combine frameworks, formats, algorithms, and tools. Additionally, installation and usage of tools cam be difficult. To address these challenges, we have developed BioSimulations (https://run.biosimulations.org), a cloud platform for reproducing and reusing models with ease. BioSimulations will provide users a unified platform to share, discover, run, and visualize biological models across various modeling scales, formats, frameworks, and algorithms.

Recent advances

Recent improvements to the platform have significantly increased the range of supported simulations to 10 model languages, 20 simulation tools, 70 algorithms, and several modeling frameworks including continuous and discrete kinetic, logical, flux balance, spatial, particle-based, and hybrid modeling. To help investigators navigate these tools, the platform can now automatically inspect models and recommend specific algorithms and tools for executing them. Large-scale simulation results up to 5 TB per project can now be saved and queried using the Highly Scalable Data Service (HSDS). Users can also use Vega to visualize simulation results with advanced, interactive diagrams, such as activity flow diagrams, process description maps, and reaction flux maps which can be created with a variety of tools such as Escher, GINSim, and Newt. Furthermore, BioSimulations makes it easy for investigators to reuse such visualizations across multiple projects, such as to compare multiple models or simulations.

Future directions

We are developing a repository of entire simulation projects, including interactive visualizations of results.

12:20 – 13:00 D **Open discussion** Presentations Room

14:00 – 14:30 L **Replay: Lightning Talks A** Presentations Room

Speakers: Adel Heydarabadipour, Yuda Munarko, Woosub Shin, John Gennari

1. **Adel Heydarabadipour** (Sharif University of Technology): *An update on Systems Biology Network Editor (SBNE): an API to render and edit the graphical representation of SBML models using their Layout and Render extensions*
2. **Woosub Shin** (University of Auckland / University of Washington): *SBMate: A Framework for Evaluating Quality of Annotations in Systems Biology Models*
3. **Yuda Munarko** (Auckland Bioengineering Institute, University of Auckland): *The Searching of Entities in Annotated Biosimulation Models from The PMR and BioModels Using NLIMED*
4. **John Gennari** (University of Washington): *OMEX metadata specification, v 1.2*
5. **Discussion**

- Adel Heydarabadipour (Sharif University of Technology); Sauro, Herbert (University of Washington)

An update on Systems Biology Network Editor (SBNE): an API to render and edit the graphical representation of SBML models using their Layout and Render extensions

Systems Biology Markup Language (SBML), the de facto software-independent standard format for storing and exchanging biological models, handles the ongoing diversification of modeling approaches and community requirements with its modular extensions. The information on the graphical representation of a model is stored in its Layout extension, where the position and dimensions of each element are specified, and is further extended by the Render extension, containing additional details on the way those elements are rendered in the network representing a model. To save developers the arduous and burdensome task of manually encoding Layout and Render information about an SBML model, we have developed a higher-level API which enables them to read, manually create, or automatically generate; modify; and write straightforwardly the layout and render features to an SBML model. The original work was presented in COMBINE 2020, and here we provide an update on it where its auto-layout and auto-render features are significantly improved and also supplemented with Graphviz graph drawing algorithms to render a clear illustration of more complicated biological networks. In addition, its GUI is now equipped with an “Item features menu” side panel allowing a user to change any information on the graphical representation of a model visually. It is now much easier to get/set the Layout and Render extension features of an SBML model through the API portable library as well. The portable library is originally written in C/C++ and currently provides language bindings for Python using the wrapper generator SWIG. The source code and its binary installers for Microsoft Windows, macOS, and Linux as well as its complete documentation are available online on GitHub (https://github.com/adelhpour/SBNE) and readthedocs (https://sbne.readthedocs.io /en/latest), respectively.

- Woosub Shin (University of Auckland / University of Washington); Hellerstein, Joseph (University of Washington); Munarko, Yuda (University of Auckland); Neal, Maxwell (Seattle Children’s Research Institute); Nickerson, David (University of Auckland); Rampadarath, Anand (University of Auckland); Sauro, Herbert (University of Washington); Gennari, John (University of Washington)

SBMate: A Framework for Evaluating Quality of Annotations in Systems Biology Models

The interests in repurposing and reusing existing systems biology models have been growing in recent years. Semantic annotations play an important role for this, by providing crucial information on the meanings and functions of models. However, we have not found existing tools that test whether the annotations exist or if they are of high-quality.

In this lightning talk, we introduce SBMate, a python package that would serve as a framework for evaluating the quality of annotations of systems biology models. Three default metrics are provided: coverage, consistency, and specificity. Coverage checks whether annotations exist in a model. Consistency quantifies if the annotations are appropriate for the given model entity. Finally, specificity indicates how detailed the annotations are. We discuss the three metrics and the challenges with evaluating annotations, using the models in the BioModels repository. Additional metrics would be easily added to extend the current version of SBMate.

- Yuda Munarko (Auckland Bioengineering Institute, University of Auckland); Sarwar, Dewan (Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand); Rampadarath, Anand (Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand); Atalag, Koray (Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand); Gennari, John (Biomedical & Health Informatics, University of Washington, Seattle, USA); Neal Maxwell(Seattle Children's Research Institute, Seattle, USA); Nickerson David (Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand)

The Searching of Entities in Annotated Biosimulation Models from The PMR and BioModels Using NLIMED

The semantic annotation of the biosimulation model with standard ontologies allows one to quickly and fully understand the model. Currently, hundreds of models with thousands of entities in the PMR and Biomodels are annotated using RDF. Therefore, it is possible to find those entities in different levels of granularity, such as variable, component, and reaction, using SPARQL. Since composing SPARQL complicated, we developed NLIMED, a natural language query (NLQ) based searching tool, automatically convert NLQ into SPARQL. It works by identifying topics in NLQ using Named Entity Recognition (NER) and information retrieval techniques. Topics associated with ontology classes and their accompanying predicates are then organised into SPARQL with AND logic. Since a topic can relate to more than one ontology class, NLIMED can generate ranked SPARQLs based on its relation probability. Now, NLIMED is available online for the PMR and Biomodels; hence, it may provide more effortless searching of entities. With the NLQ based searching, we can potentially discover similar entities annotated with a different ontology in both repositories. Finally, we are confident that different repositories in the biosimulation model or other domains can implement NLIMED's approach for their entity searching.

- John Gennari (University of Washington); Nickerson, David; Misirli, Goksel; Konig, Matthias; Neal, Maxwell; Waltemath, Dagmar

OMEX metadata specification, v 1.2

Although the COMBINE community has reached consensus on the general approach toward semantic annotation of biosimulation models, details about best practices for annotation have been lacking. The OMEX metadata specification provides these details, and version 1.2 has recently been ratified and published in the Journal of Integrative Bioinformatics. In addition, we have created software libraries (for C/C++: libOmexMeta; for python: pyomexmeta) that allow users to create OMEX archives and RDF annotation files that conform to v1.2 of the OMEX metadata specification.

The 1.2 specification organizes annotations into model-level annotations, simple singular annotations, and composite annotations. For each of these, we provide examples, and recommendations for knowledge sources (e.g., ontologies) to use to achieve consistency in metadata across our community. We also provide a base namespace, “omex-library.org”, so that these archives and their annotations can be unified into a single RDF graph or knowledge base. The resulting knowledge base can be queried for any type of annotation described by the specification, providing improved compliance with FAIR principles.

14:00 – 17:00 P **Replay Session A - presentation room** Presentations
Moderators: Dagmar Waltemath

15:00 – 15:20 C **Replay: Mathematical model notation - proposal for SBGN** Presentations Room

Speakers: Ilya Kiselev

Kiselev Ilya (Sirius University of Science and Technology, Sochi, Russia); Kolpakov Fedor (Sirius University of Science and Technology, Sochi, Russia), Akberdin Ilya (Sirius University of Science and Technology, Sochi, Russia), Kutumova Elena (Sirius University of Science and Technology, Sochi, Russia)

In BioUML Platform from the start we utilized a visual approach to modeling. It implies that the model is created and modified as a visual diagram, each element of which is associated with a mathematical object (equation, variable, event etc.). Users may edit properties of those elements (e.g. formula, initial value, trigger, etc.) On the basis of visual representation along with specified properties corresponding executable code is generated and used for numerical simulation.

As a mathematical basis for models we use the SBML standard, while for visual representation the main standard is SBGN. Thanks to SBML annotations we are able to include necessary information about layout and other visual specifics into SBML documents. However not any SBML element has an SBGN counterpart and vice-versa. Particularly those elements are equations, functions, events and constraints. On the other hand, many mathematical models do not involve reactions or similar processes. Instead they may be formulated explicitly with algebraic and differential equations and discrete events. Such models still benefit from visual representation and thus require appropriate notation.

We propose to extend SBGN notation for such models. Notation should include elements for abstract variables, equations and other mathematical elements. Additionally arcs representing relations between variables indicate their interactions (e.g. whether increase of one variable's value also increases value of another variable).

Other interesting cases arise when we consider modular models. The package "comp" for hierarchical SBML models provides a powerful tool for modular models creation, on the other hand, while SBGN supports hierarchical models, unfortunately it lacks convenient instruments for creation of complex multilevel models.

At the same time, the approach to modularity, which we use in BioUML is different from the SBML approach - we consider modules to be connected through mathematical variables - inputs, outputs and shared variables, while SBML allows replacements of arbitrary elements in modules. Thus our aim was to extend visual notation in order to facilitate creating complex modular models while keeping it compatible with SBML.

15:20 – 15:40 C **Replay: WikiNetworks: translating manually created biological pathways for topological analysis**
Speakers: Mukta G. Palshikar Presentations Room
Palshikar, Mukta G. (Biophysics, Structural, and Computational Biology Program, University of Rochester School of Medicine and Dentistry); Hilchey, Shannon P. (University of Rochester School of Medicine and Dentistry); Zand. Martin S. (University of Rochester School of Medicine and Dentistry); Thakar, Juilee (University of Rochester School of Medicine and Dentistry);

Summary: WikiPathways is a database of 2979 biological pathways across 31 species created manually using the graphical editor PathVisio. These pathway representations can be exported in GPML (Graphical Pathway Markup Language) and in Systems Biology Graphical Notation (SBGN) format. Currently available tools do not translate pathways into networks that are useful for topological analysis. The WikiPathways app for Cytoscape allows the import and visualization of WikiPathways entries. However, this app is unable to correct drawing errors introduced by hand-curation methods, such as unconnected interactions. In addition, the output of the Cytoscape app has superfluous nodes, duplicate nodes, node insertion at hyperedges, and does not process edges to the components of node groups (e.g., protein complexes).

We developed the WikiNetworks package to standardize and construct directed networks by combining geometric information and manual annotations from WikiPathways. WikiNetworks uses the positions of pathway elements in Cartesian space to resolve ambiguous edges based on physical proximity of pathway elements and eliminates nodes that do not correspond to biological entities. We compared the output of WikiNetworks to 15 manually curated pathways from WikiPathways and found that WikiNetworks performs significantly better than the existing tool. The output of WikiNetworks is thus truly ready for input into network analysis software, and for use with dynamical modeling and simulation software.

Availability and Implementation: WikiNetworks is written in Python3 and is available on github.com/Thakar-Lab/wikinetworks and on PyPI.

15:40 – 16:00 C **Replay: The Simulation Experiment Description Markup Language (SED-ML) Level 1 Version 4**
Speakers: Matthias König, Lucian Smith Presentations Room
Smith, Lucian (University of Washington Seattle); König, Matthias (Humboldt University Berlin)

Computational simulation experiments increasingly inform modern biological research, and bring with them the need to provide ways to annotate, archive, share and reproduce the experiments performed. These simulations increasingly require extensive collaboration among modelers, experimentalists, and engineers. The Minimum Information About a Simulation Experiment (MIASE) guidelines outline the information needed to share simulation experiments. SED-ML is a computer-readable format for the information outlined by MIASE, created as a community project and supported by many investigators and software tools.

The first versions of SED-ML focused on deterministic and stochastic simulations of models. Level 1 Version 4 of SED-ML substantially expands these capabilities to cover additional types of models, model languages, parameter estimations, simulations and analyses of models, and analyses and visualizations of simulation results. To facilitate consistent practices across the community, Level 1 Version 4 also more clearly describes the use of SED-ML constructs, and includes numerous concrete validation rules.

SED-ML is supported by a growing ecosystem of investigators, model languages, and software tools, including eight languages for constraint-based, kinetic, qualitative, rule-based, and spatial models, over 20 simulation tools, visual editors, model repositories, and validators.

Additional information about SED-ML is available at https://sed-ml.org/.

Within this talk an overview of SED-ML Level 1 Version 4 is provided.

16:00 – 16:20 C **Replay: Modular approach for modeling of the human individual response to antihypertensive therapy with generation of virtual populations** Presentations Room

Speakers: Elena Kutumova

Kutumova Elena (Sirius University of Science and Technology, Sochi, Russia); Kiselev Ilya (Sirius University of Science and Technology, Sochi, Russia); Sharipov Ruslan (Sirius University of Science and Technology, Sochi, Russia); Lifshits Galina (Institute of Chemical Biology and Fundamental Medicine SB RAS, Novosibirsk, Russia); Fedor Kolpakov (Sirius University of Science and Technology, Sochi, Russia)

Hypertension is the most common risk factor for the development of cardiovascular, cerebrovascular and renal diseases. Therefore, mathematical modeling of circulatory regulation taking into account the response to antihypertensive therapies with different mechanisms of action has become increasingly important. We divide the modelling process into several stages.

Stage 1. Development of a mathematical model of cardiovascular and renal physiology. Our approach involves creation of a set of modules for basic physiological processes and construction of a model using these modules for a given disease.

Stage 2. Pharmacokinetic/pharmacodynamic modelling of antihypertensive medications including angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, calcium-channel blockers, thiazide diuretics and β-Adrenoreceptor blockers. This stage is based on the detection of target points in the model for each of these antihypertensive drug classes and validation of unknown dynamics using available clinical trials.

Stage 3. The model personalization. We consider a stationary parameterization of the model within physiological ranges as a virtual patient. Some values in parametrization we take from the laboratory analyses of a real person. For other difficult to measure values, we consider physiologically acceptable variations, thus getting a virtual population for the person.

Stage 4. Treatment simulation of the virtual population. This stage allows to get groups of constructed patients with a similar response to the medication. Therefore, we can analyze which characteristics of the person result in the effectiveness (or ineffectiveness) of the antihypertensive therapy.

The sequential implementation of these stages can provide a working scheme to optimize the choice of drug therapy for the treatment of arterial hypertension in particular patients.

Our model of cardiovascular and renal systems is contained in the web-edition of the BioUML software (https://sirius-web.org/bioumlweb/) and is available on Git repository (https://gitlab.sirius-web.org/virtual-patient/blood-pressure-regulation) including several representative use cases for working with the model in Jupyter notebook.

16:20 – 17:00 D **Open discussion** Presentations Room

PINNED
18:00 – 18:15 I **Opening remarks** Presentations Room

Speakers: Chris Meyers

18:00 – 21:00 P **Community Session - Tuesday** Presentations

Moderators: Chris Meyers

PINNED
18:15 – 19:00 I **Invited talk: Gary Mirams: Cardiac models – now we can share them, but how do they behave and where did they come from?** Presentations Room

Speakers: Gary Mirams

The Physiome Model Repository is a great resource that has allowed the cardiac modelling community to share models in CellML format for many years now. I'll discuss our efforts exploring how to build on this by characterising what these models do in different experimental situations using our Cardiac Electrophysiology Web Lab. I'll talk about some of the challenges in providing enough information for reproduction of the model building process, and communicating the resulting uncertainty in model parameters and structures.

PINNED 19:15 – 20:00	|	**Invited talk: David Ross / Elizabeth Strychalski: Toward More Quantitative Engineering of Biological Function** *Speakers: David Ross, Elizabeth Strychalski*	Presentations Room
PINNED 20:15 – 21:00	|	**Invited talk: Angel Goni-Moreno: Characterising contextual information for circuit design** *Speakers: Angel Goni-Moreno*	Presentations Room

21:00 – 21:30 L **Replay: Lightning Talks B** Presentations Room

Speakers: Dilan Pathirana, Rahuman Sheriff, Woosub Shin, Hugh Sorby

1. **Dilan Pathirana** (University of Bonn): *PEtab Select: add model selection to a PEtab-based calibration workflow*
2. **Rahuman Sheriff** (European Bioinformatics Institute (EMBL-EBI)): *Model Curation in BioModels and FROG*
3. **Woosub Shin** (University of Auckland): *Automating Model Annotations: Developing a Recommender System for Annotating Systems Biology Models*
4. **Hugh Sorby** (University of Auckland): *libCellML: almost at version 1*
5. **Discussion**

- Dilan Pathirana (University of Bonn); Hasenauer, Jan (University of Bonn); the PEtab Select contributors

PEtab Select: add model selection to a PEtab-based calibration workflow

PEtab is a file format for the specification of parameter estimation problems, and has been adopted by several systems biology modelling tools. PEtab facilitates use of the different state-of-the-art methods, which are unique to specific calibration tools, on the same PEtab problem. This includes methods related to optimization and uncertainty analysis. The current specification allows for the semi- or full-automation of several tasks in the model calibration process, on the level of a single model. The work presented here extends PEtab, and compatible tools, to support efficient calibration of, and selection from, multiple models. The extension supports specification of the: model space and its constraints; selection algorithms and criteria; parameter estimation problems via PEtab; and initial models. PEtab Select also provides a Python library and command-line interface for the management of the model space and selection problems such that tools, which already support model calibration with PEtab, can easily implement model selection.

- Rahuman Sheriff (European Bioinformatics Institute (EMBL-EBI)); BioModels Team (EMBL-EBI); FBC curation standards working group;

Model Curation in BioModels and FROG

BioModels (https://www.ebi.ac.uk/biomodels) is one of the largest repositories of curated mathematical models of biological and pharmacological processes. Models submitted to BioModels repository are manually curated. Curation process involves encoding models in the standard SBML format, ensuring reproducibility of simulation results in the reference manuscript and semantic enrichment of the model with controlled vocabularies. A new curation milestone was recently achieved by BioModels; over 1000 models are now fully curated. Our recent model curation target areas include tumor immune interaction, cell cycle and COVID-19. Over 99% of the curated models are ordinary differential equation models. BioModels Parameters resource facilitates easy search and retrieval of model parameters from these kinetic models. Furthermore, BioModels has started implementation of FROG analysis (https://www.ebi.ac.uk/biomodels/curation/fbc), a community standard to foster reproducibility and curation of constraint-based models including Genome Scale Metabolic models (GEMs). FROG analysis generates FROG report, a reference dataset that consists of Objective function value, Flux Variability Analysis (FVA), Gene and Reaction deletion fluxes. The FROG report can be then used to assess the reproducibility and curate the model. We invite the community to submit GEMs to BioModels along with the FROG report to facilitate curation and contribute to the community manuscript.

- Woosub Shin (University of Auckland); Hellerstein, Joseph (University of Washington); Sauro, Herbert (University of Washington); Gennari, John (University of Washington)

Automating Model Annotations: Developing a Recommender System for Annotating Systems Biology Models

The utility and usefulness of a biomedical model is greatly increased by providing semantically meaningful annotations of model elements, such as chemical species and reactions. At present, annotations are applied manually, a process that is time-consuming and skill-intensive.

We are developing a recommender system that suggests annotations for biomedical models. Our current focus is

to annotate reactions using available information from annotations of chemical species. Our initial focus is on models in SBML format. In this lightning talk, we describe the current state of annotations in the BiGG and BioModels repositories as well as discuss our progress and challenges with building the recommender system.

- Hugh Sorby (University of Auckland); Garny, Alan (University of Auckland);Nickerson, David (University of Auckland)

libCellML: almost at version 1

We will summarise the current status and future plans of libCellML as we edge ever close to releasing version 1 of the libCellML library.

21:00 – 00:00 P **Replay Session B - presentation room** Presentations

22:20 – 22:40 C **Replay: PEtab-MS – a format for specifying parameter estimation problems for multiscale models** Presentations Room

Speakers: Emad Alamoudi

Alamoudi, Emad (University of Bonn); Starruß, Jörn, Center of Information Services and High Performance Computing (ZIH), Technische Universität Dresden, Dresden, Germany; Hasenauer, Jan, Faculty of Mathematics and Natural Sciences, University of Bonn, 53113 Bonn, Germany, Helmholtz Zentrum München - German Research Center for Environmental Health, Institute of Computational Biology, 85764 Neuherberg, Germany, Technische Universität München, Center for Mathematics, Chair of Mathematical Modeling of Biological Systems, 85748 Garching, Germany; FitMultiCell Consortium, Faculty of Mathematics and Natural Sciences, University of Bonn, 53113 Bonn, Germany, Center of Information Services and High Performance Computing (ZIH), Technische Universität Dresden, Dresden, Germany, Helmholtz Zentrum München - German Research Center for Environmental Health, Institute of Computational Biology, 85764 Neuherberg, Germany, Technische Universität München, Center for Mathematics, Chair of Mathematical Modeling of Biological Systems, 85748 Garching, Germany, Institute for Medical Informatics and Biometry, Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany, Centre for Modelling and Simulation in the Biosciences, BioQuant, Heidelberg University, 69120, Heidelberg, Germany

Many biological tissues are highly organized and dynamic. Their Spatio-temporal patterns are relevant in many biological processes, including tissue homeostasis, viral infection, or tumor development. Multiscale modeling that describes the construction of a system model from several sub-models on different length scales became an important tool to simulate the dynamics of these biological tissues. Yet, reproducibility and reusability of the results is still an issue.

Here, we introduce a PEtab extension called PEtab-MS, which facilitates the specification of parameter estimation problems for multiscale models. PEtab-MS encodes the estimation problem, including unknown parameters, parameter bounds, experimental conditions, and measurement data. It builds upon the PEtab format which is well tested and already supported in many simulations and parameter estimation toolboxes with hundreds of users. The new extension is already supported with FitMultiCell pipeline, which is a user-friendly, open-source, and scalable platform for simulating and parameterizing computational models of multi-cellular processes.

22:40 – 23:00 C **Replay: Introduction to OpenML** Presentations Room

Speakers: Prabhant Singh

Singh, Prabhant (OpenML)

Machine learning research should be easily accessible and reusable. OpenML is an open platform for sharing datasets, algorithms, and experiments - to learn how to learn better, together. This talk will cover introduction to OpenML and how to use the platform for reproducible science.

23:00 – 23:20 C **Replay: Synchronisation of visual and textual formats for mathematical models: Antimony extension in BioUML** Presentations Room

Speakers: Ilya Kiselev

Kiselev Ilya (Sirius University of Science and Technology, Sochi, Russia); Kolpakov Fedor (Sirius University of Science and Technology, Sochi, Russia), Akberdin Ilya (Sirius University of Science and Technology, Sochi, Russia), Kutumova Elena (Sirius University of Science and Technology, Sochi, Russia)

To ensure a model exchange and reproducibility of simulation results a variety standards for model representation have been developed over the past decades (SBML, SBGN, CellML, BioPAX). Depending on a context it can be more suitable to present a model of a biological system either as a graph (for example, as a map of metabolic or signal transduction pathway) or by means of the human-readable text such as Antimony language.

Developing the BioUML Platform we from the start tried to combine visual modeling (creating mathematical models as visual diagrams). As a standard for visual modeling the tool mainly employs SBGN notation, while SBML is used as a foundation for mathematical formalism of the model. In the past few years we also made an effort to support human-readable model representation by adopting Antimony language. We chose Antimony because it is very close to SBML and SBGN standards which facilitates their combining in the frames of one tool.

However, there are some gaps between different standards resulting in the interoperability issues:
- some SBML elements (for example, algebraic equations in unresolved form) can not be described in Antimony syntax;
- Antimony does not provide enough information to generate SBGN diagrams.

In this talk we would like to discuss possible extensions of the Antimony language in order to cover those problems. There are two possible ways of extension:
- Direct addition of new terms into Antimony to make it more compliant with SBML and SBGN.
- Addition of annotation mechanism into Antimony language similar to the scheme of the SBML enabling using custom terms to annotate Antimony objects.

Next are some possible suggestions for Antimony extension. Some of them may be added directly to the language, while others should use custom annotations.
- algebraic equations in unresolved form;
- ports direction (input, output, contact);
- annotation of complexes and entities inner structure (including subunits, units of information and state variables). Possible annotation form is “(A{p}:B):(D)3”;
- SBGN entity types (macromolecule, simple chemical etc.);
- process types (association, dissociation, etc.);
- custom molecule structure (e.g. glycans);
- layout information (node location, borders, edge path);

23:20 – 00:00 D **Open discussion** Presentations Room

OCTOBER 13 • WEDNESDAY

06:00 – 09:00 B **FAIR assessment of COMBINE models** Breakout Sessions

In this breakout, we will discuss the FAIRness of computational models, both inside and outside the COMBINE community. The goal is to summarise the state of the art, to propose FAIR metrics for computational models, and to give recommendations for model FAIRification.

Communities: BioPAX, CellML, COMBINE Archive and OMEX, ModelXchange, NeuroML, SBGN, SBML, SBOL, SBOL Visual, SED-ML, everyone is welcome

Chair: Dagmar Waltemath

12:00 – 13:00 T **Developing computational models with SBML - sbmlutils, sbmlsim, cy3sbml** Tutorials

The Systems Biology Markup Language (SBML) is the de facto standard for representation and exchange of mathematical models of biological systems. SBML can represent many different classes of phenomena in biology, including metabolic networks, signaling pathways, or regulatory networks, and supports models of arbitrary complexity from single processes to multi-scale models.

A key challenge in computational modeling, especially for people new in the field, is how to encode or develop ordinary differential equation models in SBML. Within this tutorial we provide a hands-on introduction in the python packages sbmlutils and sbmlsim and the Cytoscape app cy3sbml which address this issue by providing solutions for the programmatic creation, simulation and visualization of SBML models.

- model creation with sbmlutils
- report generation with SBML4Humans
- model visualization with cy3SBML
- model simulation with sbmlsim

[1] sbmlutils - Python utilities for SBML, https://github.com/matthiaskoenig/sbmlutils
[2] sbmlsim - SBML model simulation made easy, https://github.com/matthiaskoenig/sbmlsim
[3] cy3sbml - Visualization of SBML models in Cytoscape, https://github.com/matthiaskoenig/cy3sbml"

Communities: COMBINE Archive and OMEX, SBML

Chair: Matthias König

14:00 – 15:00 B **SBML Level 3 Package for Spatial Processes, next steps** Breakout Sessions

The SBML Level 3 Package for Spatial Processes is essentially completed at this point. Following a similar breakout at HARMONY this year, this session aims to demonstrate exchange of documents in the SBML spatial format, and to discuss what further steps need to be taken before finalizing a first version of the package

Communities: SBML

Chair: Frank Bergmann

14:00 – 16:00 B **Standardized simulation execution with BioSimulators and co-simulation with Vivarium: Overview and demos** Breakout Sessions 2

Whole-cells models have substantial potential to revolutionize bioscience, bioengineering, and medicine. To predict cellular phenotype from genotype, whole-cell models must capture all of the biochemical activity inside cells across multiple scales, leveraging heterogeneous data and theory. One promising way to simulate entire cells is to co-simulate submodels of individual cellular subsystems. This strategy could enable multiple investigators to develop multiple submodels, and enable each investigator to use the most appropriate modeling framework, model language, and simulation method for each submodel. This strategy requires capabilities to identify connections among models; execute a broad range of models involving multiple modeling frameworks, model languages, and simulation methods; and co-simulate heterogeneous models.

To accelerate whole-cell modeling, we are developing several interoperable resources for systematically sharing, discovering, executing, and combining models.
- BioSimulators: Registry of standardized interfaces to simulation tools
- BioSimulations: Central portal for sharing, discovering, and executing models
- Vivarium: Framework for co-simulating multiple models

Wednesday: Overview and demos
We will introduce the key principles of BioSimulators, BioSimulations, and Vivarium and demonstrate how to use them to find, execute, and combine models.
14:00 UTC: BioSimulators and BioSimulations
15:00 UTC: Vivarium

Communities: CellML, COMBINE Archive and OMEX, ModelXchange, SBML, SED-ML, KiSAO

Chair: Jonathan Karr

15:00 – 17:00 B **Defining a Common Protocol Representation** Breakout Sessions 3

Speakers: Jacob Beal
This session is a working session for stakeholders to discuss needs, requirements, and opportunities for a common protocol representation, leveraging the current prototype specification and implementation of the Protocol Activity Markup Language (PAML).

Communities: SBOL, Protocol languages community that we are organizing

Chair: Jacob Beal

15:00 – 18:00 B **NeuroML Development Workshop** Breakout Sessions 4

Speakers: Padraig Gleeson
A workshop will be held organised by the NeuroML Editorial Board and the Scientific Committee to review the work of the project over the past year, plan for future developments and engage with the members of the NeuroML user and developer communities.

Communities: NeuroML

Chair: Padraig Gleeson

15:00 – 18:00 B **SBOL3 Software Support – Libraries** Breakout Sessions

Speakers: Chris Meyers
There will be three breakout sessions on issues related to SBOL3 software support each day of the breakout sessions. This one will focus on SBOL3 libraries.

Communities: SBOL

Chair: Chris Myers

16:00 – 17:00 B **Funding COMBINE standard libraries** Breakout Sessions 2

Supporting and maintaining good quality APIs in a variety of programming languages for the number of standards involved in COMBINE cannot be done with zero funding. In the Research Software Development Group at the Centre for Advanced Research Computing (ARC), UCL we have started the process of creating a support base for these APIs. Initial funding came from the CZI and ARC has pledged to fund very basic maintenance in the absence of other funding, but this is not enough. Securing grants that fund software maintenance activity is not easy.

As a community we need to discuss means of funding ongoing support for our software infrastructure. At this breakout we hope to gather together some ideas of how we might secure this funding whether it be small donations from multiple research grants to crowd-sourcing ! We know from experience good APIs hugely benefit the adoption and use of our standards.

Communities: BioPAX, CellML, COMBINE Archive and OMEX, FROG, ModelXchange, NeuroML, Personalized Medicine, PEtab, SBGN, SBML, SBOL, SBOL Visual, SED-ML

Chair: Sarah Keating

18:00 – 19:00 B **Discussion of the current SED-ML V1.4** Breakout Sessions

The session will discuss the following topics: What's new in V1.4, what is the current compatibility between tools and how this can be improved, standard ways to ask for tool SEDML capabilities, the SEDML test suite, and a minimal set of features all tools should support.

Communities: CellML, COMBINE Archive and OMEX, SBML, SED-ML

Chair: Herbert Sauro

19:00 – 20:00 B **SBML distrib - Multidimensional distributions and covariance matrices** Breakout Sessions

Biological models often contain elements that have inexact numerical values, since they are based on values that are stochastic in nature or data that contains uncertainty. The Systems Biology Markup Language (SBML) Level 3 Core specification does not include an explicit mechanism to include inexact or stochastic values in a model, but it does provide a mechanism for SBML packages to extend the Core specification and add additional syntactic constructs. The SBML Distributions package for SBML Level 3 adds the necessary features to allow models to encode information about the distribution and uncertainty of values underlying a quantity.

SBML distrib version 1 was designed for the description and sampling from one-dimensional distributions. Many biological and computational problems require the description and sampling of multi-dimensional distributions and encoding information in the form of covariance matrices. Examples given Monte-Carlo sampling methods or non-linear mixed effects models.

Within this breakout solutions will be discussed and a first draft of how to encode this information will be developed. In addition requirements and features for SBML distrib version 2 will be collected.

Communities: SBML, SED-ML

Chair: Matthias König

OCTOBER 14 • THURSDAY

08:00 – 11:00 B **FAIR assessment of COMBINE models** Breakout Sessions 2

Speakers: Dagmar Waltemath

In this breakout, we will discuss the FAIRness of computational models, both inside and outside the COMBINE community. The goal is to summarise the state of the art, to propose FAIR metrics for computational models, and to give recommendations for model FAIRification.

Communities: BioPAX, CellML, COMBINE Archive and OMEX, ModelXchange, NeuroML, SBGN, SBML, SBOL, SBOL Visual, SED-ML, everyone is welcome

Chair: Dagmar Waltemath

10:00 – 13:00 B **SBOL Visual** Breakout Sessions

Speakers: Thomas Gorochowski

People who are engineering biological organisms often find it useful to communicate in diagrams, both about the structure of the nucleic acid sequences that they are engineering and about the functional relationships between sequence features and other molecular species. Some typical practices and conventions have begun to emerge for such diagrams. SBOL Visual aims to organize and systematize such conventions in order to produce a coherent language for expressing the structure and function of genetic designs. At the same time, we aim to make this language simple and easy to use, allowing a high degree of flexibility and freedom in how such diagrams are organized, presented, and styled—in particular, it should be readily possible to create diagrams both by hand and with a wide variety of software programs. Finally, means are provided for extending the language with new and custom diagram elements, and for adoption of useful new elements into the language.

This session will be used to work on general enhancements to the SBOL Visual standard, explore options for further automation and streamlining of the GitHub repository that houses the standard, and to generate an SEP focused on the use of parametric glyph definitions.

Communities: SBOL Visual

Chair: Thomas Gorochowski

12:00 – 13:00 B **openTECR - an open database for Thermodynamics of Enzyme-Catalyzed Reactions** Breakout Sessions 3

openTECR (Open database on Thermodynamics of Enzyme-Catalyzed Reactions) is a database and a community.

Bob Goldberg and colleagues have done a wonderful job in collecting data on equilibrium constants of enzyme-catalyzed reactions from the primary literature. Their work was published in a nice, human-readable form, in a series of reviews. It collects extensive evidence for measured data, with over 4500 data points from 1400 publications.

We, however, usually pursue automated, computer-assisted interpretation and usage of this data. We therefore need a reliable, machine-actionable version of this database.

We organize ourselves openly using a Google Groups (https://groups.google.com/g/opentecr). So far, we aligned our small community (~20 members) on a common data model, and created a GitHub organization where we store our data and code (https://github.com/opentecr/); we collected so far a list of >40 publications that were not included in any previous databases. We want to integrate the previous database (which is abandoned) into ours in the near future, and work together on correcting errors in the existent data, curating additional details about the published equilibrium constants, and continuously integrating new data.

This database is expected to serve computational and experimental scientists in the fields of metabolic engineering, genome-scale metabolic modelling, biocatalysis and related fields by providing curated information. It will probably be used by eQuilibrator and COBRA as the shared data basis for their individual calculations.

Communities: BioPAX, CellML, PEtab, SBML, STRENDA

Chair: Robert T. Giessmann

13:00 – 15:00 T **Modelling and analysis with COPASI using python scripting** Tutorials

During this tutorial, attendees will learn basic techniques for modeling and simulation of biochemical networks using the BasiCO (https://basico.readthedocs.io/). We'll see how to create models or use models in SBML format, simulate them, and set up parameter estimation tasks. The hands-on exercises throughout the course will ensure that attendees become familiar with the software tools and with analyzing, creating, editing, importing, and simulating biochemical networks.

Prerequisites: Some knowledge of mathematical modeling will be advantageous as well as some basic knowledge of the python language.

Communities: COMBINE Archive and OMEX, SBML, SED-ML

Chair: Frank Bergmann

14:00 – 17:00 B **Standardized simulation execution with BioSimulators and co-simulation with Vivarium: Hack with the developers** Breakout Sessions

Whole-cells models have substantial potential to revolutionize bioscience, bioengineering, and medicine. To predict cellular phenotype from genotype, whole-cell models must capture all of the biochemical activity inside cells across multiple scales, leveraging heterogeneous data and theory. One promising way to simulate entire cells is to co-simulate submodels of individual cellular subsystems. This strategy could enable multiple investigators to develop multiple submodels, and enable each investigator to use the most appropriate modeling framework, model language, and simulation method for each submodel. This strategy requires capabilities to identify connections among models; execute a broad range of models involving multiple modeling frameworks, model languages, and simulation methods; and co-simulate heterogeneous models.

To accelerate whole-cell modeling, we are developing several interoperable resources for systematically sharing, discovering, executing, and combining models.
- BioSimulators: Registry of standardized interfaces to simulation tools
- BioSimulations: Central portal for sharing, discovering, and executing models
- Vivarium: Framework for co-simulating multiple models

Thursday: Open hacking with the developers
We invite the community to explore BioSimulators, BioSimulations, and Vivarium. The BioSimulators, BioSimulations, and Vivarium developers will be available for one-on-one discussion. Potential topics include, but are not limited to
Using BioSimulations to execute simulations;
Creating a standardized interface to a simulation tool;
Publishing a modeling project with BioSimulations; and
Implementing a co-simulation of two models with Vivarium.

Communities: CellML, COMBINE Archive and OMEX, ModelXchange, SBML, SED-ML

Chair: Jonathan Karr

15:00 – 17:00 B **Implementation and Roadmap for Common Protocol Representation** Breakout Sessions 3

Speakers: Jacob Beal

This session is a working session that will use the outcomes of the prior “Defining a Common Protocol Representation” session to work on the specifics of implementation of a common protocol representation, leveraging the current prototype specification and implementation of the Protocol Activity Markup Language (PAML), and to determine priorities and next steps.

Communities: SBOL, Protocol languages community that we are organizing

Chair: Jake Beal

15:00 – 17:00 B **PEtab – Interoperable Specification of Parameter Estimation Problems in Systems Biology**
Breakout Sessions 4

Reproducibility and reusability of the results of data-based modeling studies are essential. Yet, there has been – so far – no broadly supported format for the specification of parameter estimation problems in systems biology. Therefore, we developed PEtab, a format which facilitates the specification of parameter estimation problems using Systems Biology Markup Language (SBML) models and a set of tab-separated value files describing the observation model and experimental data as well as parameters to be estimated.

We already implemented PEtab support into eight well-established model simulation and parameter estimation toolboxes with hundreds of users in total. We provide a Python library for validation and modification of a PEtab problem and as well as example parameter estimation problems based on recent studies.

In this breakout session, we will give a brief introduction on the current state of PEtab and how PEtab development is organized. We will discuss current shortcomings and new features to be include in the next version of PEtab. This includes, for example, support for model definitions in other formats than SBML, specification of model selection problems, and the specification of more complex experimental settings than currently possible. Furthermore, we will be happy to discuss any further PEtab-related topics brought up by the breakout participants.

Communities: PEtab, SBML

Chair: Daniel Weindl

15:00 – 18:00 B **SBOL3 Software Support – Test Cases**
Breakout Sessions 2

Speakers: Chris Meyers

There will be three breakout sessions on issues related to SBOL3 software support each day of the breakout sessions. This one will focus on SBOL3 test cases.

Communities: SBOL

Chair: Chris Myers

18:00 – 19:00 B **The future of SED-ML**
Breakout Sessions

Session to discuss the future of SEDML.

Communities: CellML, SBML, SED-ML

Chair: Herbert Sauro

21:00 – 22:00 B **Automating Model Annotations: Considerations for a Recommender System for Annotating Biomedical Models** Breakout Sessions

Annotations are meta data in model files (or in separate annotation files linked with model files) that provide semantics and other information for model elements such as chemical species and reactions. Such meta data are essential for model users to understand the model's scope and mechanisms (e.g., is ATP modeled explicitly?) as well as for the comparative analysis of models.

Unfortunately, many modelers do not annotate their models or they annotate at a very high level (e.g., annotating a reactant as a "chemical species"). This motivates an interest in assistive tools for annotation. Many systems already provide some assistance via GUIs with dropdowns to choose annotations. We are exploring a step beyond that makes use of machine learning techniques to recommend annotations. For example, a recommender system might infer annotations based on the names of chemical species and/or the combinations of chemical species used as reactants and/or products.

We do not expect that the recommender systems will provide a perfect assignment of meaningful annotations to model entities (although it will provide a confidence along with each recommendation). Rather, we expect there will be interactions between the user and the recommender system. A simple interaction would be to choose from an ordered list (e.g., from most likely to least likely) of possible annotations. A more complex interaction might be that having chosen an annotation for a chemical species in one reaction, several other annotations are tentatively inferred. A further complexity occurs if the user decides to re-do some annotations but not others.

This breakout will explore the following:
Interest in and the desirability of automated annotations, especially who (e.g., modeler, model user) would use it and insights into their use cases.
The kinds of GUI interactions that are most convenient.
Software environments in which potential users would like to exploit these capabilities. We have been particularly interested in VSCode because of its extension capabilities.

Communities: COMBINE Archive and OMEX, ModelXchange, SBML

Chairs: Joseph Hellerstein, Herbert Sauro, John Gennari, Woosub Shin

OCTOBER 15 • FRIDAY

10:00 – 13:00 B **SBOL Visual** Breakout Sessions

Speakers: Thomas Gorochowski

People who are engineering biological organisms often find it useful to communicate in diagrams, both about the structure of the nucleic acid sequences that they are engineering and about the functional relationships between sequence features and other molecular species. Some typical practices and conventions have begun to emerge for such diagrams. SBOL Visual aims to organize and systematize such conventions in order to produce a coherent language for expressing the structure and function of genetic designs. At the same time, we aim to make this language simple and easy to use, allowing a high degree of flexibility and freedom in how such diagrams are organized, presented, and styled—in particular, it should be readily possible to create diagrams both by hand and with a wide variety of software programs. Finally, means are provided for extending the language with new and custom diagram elements, and for adoption of useful new elements into the language.

This session will be used to work on general enhancements to the SBOL Visual standard, explore options for further automation and streamlining of the GitHub repository that houses the standard, and to generate an SEP focused on the use of parametric glyph definitions.

Communities: SBOL Visual

Chair: Thomas Gorochowski

11:00 – 13:00 B **Discussion and hands-on training on FROG Analysis - a community standard to foster reproducibility and curation of constraint-based models** Breakout Sessions 2

Speakers: Rahuman Sheriff

Community standards for consistent model reconstruction and curation of constraint-based models such as genome-scale metabolic models (GEMs) are crucial to ensure reproducibility, reliability and FAIR sharing. We initiated a community effort for standardised assessment of model reproducibility and model curation of GEMs, which is currently lacking. Following discussions at dedicated breakout sessions at HARMONY2020, COMBINE2020 and HARMONY 2021 we have now developed FROG analysis, a community standard to foster reproducibility and curation of constraint-based models (https://www.ebi.ac.uk/biomodels/curation/fbc). FROG analysis is currently being used to curate GEMs in BioModels. Reproducibility assessment is an essential part of curation of a model in BioModels. FROG analysis generates FROG report, a reference dataset that consists of Objective function value, Flux Variability Analysis (FVA), Gene and Reaction deletion fluxes. The FROG report is used to assess the reproducibility of the numerical values and curate the model. In this session, we plan to provide hands-on training to use FROG tools. We invite the community to submit GEMs to BioModels along with the FROG report to facilitate curation and contribute to the community manuscript which is currently under preparation.

Communities: COMBINE Archive and OMEX, FROG, SBML, SED-ML

Chair: Rahuman Sheriff

14:00 – 16:00 B **SBML and SED-ML support for Multi approach Modelling (MAM)** Breakout Sessions

Speakers: Rahuman Sheriff

Multi-approach modelling (MAM) involves modelling of biological systems using interconnected multiple frameworks. A multi-approach model will include a combination of ODE, PDE, logical, constraint-based and agent-based models, linked with each other. During Harmony 2021, in a dedicated breakout session we discussed the current state-of-the art technologies and standards (SBML and SED-ML) available to support multi-approach modelling and also identified potential challenges and gaps. In COMBINE 2021, we would like to use this breakout session to continue the discussion, gather the community to initiate development of new SBML package and complementary SED-ML enhancement to support multi-approach modelling.

Communities: COMBINE Archive and OMEX, SBML, SED-ML

Chair: Rahuman Sheriff

15:00 – 18:00 B **SBOL3 Software Support – Conversion Routes** Breakout Sessions 2

Speakers: Chris Meyers

There will be three breakout sessions on issues related to SBOL3 software support each day of the breakout sessions. This one will focus on conversion routines to/from SBOL3.

Communities: SBOL

Chair: Chris Myers

16:00 – 17:00 B **Standardized simulation execution with BioSimulators and co-simulation with Vivarium: Community input and feedback** Breakout Sessions

Whole-cells models have substantial potential to revolutionize bioscience, bioengineering, and medicine. To predict cellular phenotype from genotype, whole-cell models must capture all of the biochemical activity inside cells across multiple scales, leveraging heterogeneous data and theory. One promising way to simulate entire cells is to co-simulate submodels of individual cellular subsystems. This strategy could enable multiple investigators to develop multiple submodels, and enable each investigator to use the most appropriate modeling framework, model language, and simulation method for each submodel. This strategy requires capabilities to identify connections among models; execute a broad range of models involving multiple modeling frameworks, model languages, and simulation methods; and co-simulate heterogeneous models.

To accelerate whole-cell modeling, we are developing several interoperable resources for systematically sharing, discovering, executing, and combining models.
- BioSimulators: Registry of standardized interfaces to simulation tools
- BioSimulations: Central portal for sharing, discovering, and executing models
- Vivarium: Framework for co-simulating multiple models

Friday: Community discussion for input and feedback
We invite the community to join a group discussion to help shape the directions of BioSimulators, BioSimulations, and Vivarium. We invite input and feedback, such as needed features.
16:00 UTC: BioSimulators and BioSimulations
16:30 UTC: Vivarium

Communities: CellML, COMBINE Archive and OMEX, ModelXchange, SBML, SED-ML

Chair: Jonathan Karr