Stuttgart 2024

COMBINE 2024 - Conference of the COmputational Modeling in BIology NEtwork

September 1-5, 2024, University of Stuttgart

Sebastian Höpfl[1], Dagmar Walthemath[2], David Nickerson[3], Nicole Radde[1,*]

[1]*Institute for Stochastics and Applications, University of Stuttgart, Stuttgart, Germany*
[2]*Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany*
[3]*Auckland Bioengineering Institute, University Of Auckland, New Zealand*
[*]nicole.radde@simtech.uni-stuttgart.de (corr. author)

Abstract

The COMBINE consortium is an initiative to coordinate the development of community standards and formats for computer models in biology and medicine. COMBINE members organize COMBINE as an annual conference at varying places. This conference is aimed at scientists at all career levels interested in using and developing these standards. It also provides a platform for exchange, discussion, and interactive experimentation and learning.
From September 1 to 5, the COMBINE 2024 conference took place at the University of Stuttgart, Campus Stuttgart-Vaihingen. The event was held as an in-person workshop-style event with the opportunity for remote participation during break-out sessions to enable broad community participation. COMBINE 2024 has taken place as a satellite event of the Virtual Physiological Human (VPH) 2024 Conference, which also took place in Stuttgart from September 4-6, 2024. The event's co-location was advantageous and promoted exchange and collaboration between the two scientific networks.
This year's COMBINE was co-hosted by the Stuttgart Cluster of Excellence EXC2075 "Data-Integrated Simulation Science (SimTech)". SimTech is an engineering-driven cluster that develops and applies multi-scale computational models and simulation schemes in various fields. Developing standards and workflows to enable and facilitate seamless management, exchange, and reuse of models and data is an important topic in the cluster. Moreover, the development of research software and strategies for long-term maintenance for the community are intensively discussed in the cluster across different communities and application fields. We believe that the long-term maintenance of models and computational tools and, along with this, a broad usage requires community efforts. In this respect, we see the COMBINE consortium as a successful role model in the Systems Biology field. On the other hand, SimTech researchers have also successfully developed software tools, e.g., for model coupling (preCICE) or efficient simulations of engineering-driven models and multi-scale models (Dynamore) and in biocatalysis (EnzymeML). The idea was to enable an exchange between COMBINE and SimTech and lively discussions, which worked out nicely.

Keywords: Reproducibility; Standards; Formats; Systems Biology, COMBINE; Stuttgart

Contents

1 Invited talks

1.1 The EnzymeML framework: improving efficiency and quality of biocatalytic science

Jürgen Pleiss, Jan Range, and Max Häußler, University of Stuttgart, Germany

Biocatalysis is entering a promising era as a data-driven science. High-throughput experimentation generates a rapidly increasing stream of biocatalytic data, which is the raw material for mechanistic and data-driven modeling to design improved biocatalysts and bioprocesses. However, data management has become a bottleneck to progress in biocatalysis. In order to take full advantage of rapid progress in experimental and computational technologies, biocatalytic data should be findable, accessible, interoperable, and reusable (FAIR). The EnzymeML framework (`https://github.com/EnzymeML`) provides reusable and extensible tools and a standardized data exchange format for FAIR and scalable data management in biocatalysis (Range et al., 2022). To enable storage, retrieval, and exchange of enzymatic data, the XML–based markup language EnzymeML has been developed (Pleiss, 2021). An EnzymeML document contains information about reaction conditions and the measured time course of substrate or product concentrations. Kinetic modelling is performed by uploading EnzymeML documents to the modelling platforms COPASI or PySCeS or by using the JAX platform. The rate equation and the estimated kinetic parameters are then added to the EnzymeML document. The EnzymeML document containing the experimental and the modelling results is then uploaded to a Dataverse installation or to the reaction kinetics database SABIO-RK. The workflow of a project is encoded as Jupyter Notebook, which can be re-used, modified, or extended The feasibility and usefulness of the EnzymeML toolbox was demonstrated in six scenarios, where data and metadata of different enzymatic reactions are collected, analysed, and uploaded to public data repositories for future re-use (Lauterbach et al., 2023). FAIRification of data and software and the digitalization of biocatalysis improve the efficiency of research by automation and guarantee the quality of biocatalytic science by reproducibility4. Most of all, they foster reasoning and creating hypotheses by enabling the reanalysis of previously published data, and thus promote disruptive research and innovation.

1.2 Reproducible digital twins for personalized liver function assessment

Matthias König, Humboldt-University Berlin, Germany

Essential prerequisites for the practical application and translation of computational models include: (i) reproducibility of results; (ii) model reusability and extensibility; (iii) data availability; and (iv) strategies for model stratification and individualization. Here, we present a modeling workflow built around these foundational prerequisites, with a focus on liver function tests. Despite the paramount significance of liver function assessment in hepatology, reliable quantification remains a clinical challenge. Dynamic liver function tests offer a promising method for non-invasive in vivo assessment of liver function and metabolic phenotyping. By leveraging whole-body physiologically-based pharmacokinetic (PBPK) models, we're simulating these tests and positioning PBPK models as digital twins for metabolic phenotyping and liver function assessment. To develop and validate our models, we established the open pharmacokinetics database, PK-DB, containing curated data from 600+ clinical studies (Grzegorzewski et al., 2021b; Grzegorzewski et al., 2021a). Our models are individualizable and stratifiable, enabling simulation of lifestyle factors and co-administration effects on drug metabolism. Our models have been instrumental in clinical scenarios: from predicting individual outcomes post-hepatectomy (Köller et al., 2021b; Köller et al., 2021a) to discerning the impact of CYP2D6 gene variants on liver function tests (Grzegorzewski et al., 2022). These models are constructed hierarchically, describing metabolic and other biological processes in organs like the liver and kidneys, seamlessly integrated with whole-body physiology. Notably, all models and data are readily available and reproducible for reuse, encoded in the Systems Biology Markup Language (SBML) (Keating et al., 2020). We will provide an overview of these PBPK models and demonstrate how SBML and FAIR principles can facilitate model development, coupling, and reuse.

1.3 preCICE – A General-Purpose Simulation Coupling Library

Benjamin Uekermann, University of Stuttgart, Germany

preCICE (`https://precice.org/`) is an open-source coupling software for partitioned multi-physics and multi-scale simulations including PDE-PDE and PDE-ODE coupling. Thanks to the software's library approach (the simulations call the coupling) and its high-level API, only minimally-invasive changes are required to prepare an existing (legacy) simulation software for coupling. Moreover, ready-to-use adapters for many popular simulation software packages are available, e.g. for OpenFOAM, SU2, CalculiX, FEniCS, and deal.II. For the actual coupling, preCICE offers methods for fixed-point acceleration (quasi-Newton acceleration), fully parallel communication (MPI or TCP/IP), data mapping (radial-basis function interpolation), and time interpolation (waveform relaxation). Today, although being an academic software project at heart, preCICE is used by more than 100 research groups in both academia and industry. In this presentation, I introduce the basic concepts of preCICE and discuss existing and potential applications in biology.

1.4 Improving Curation: Biomodels and Annotation

Lucian Smith, Herbert Sauro, John Gennari, David Nickerson, S. Malik-Sheriff Rahuman, V. N. Nguyen Tung, University of Washington, USA

The BioModels Database has over 1000 curated models from published papers. Curators at the EBI ensure that the model can be used to reproduce at least one figure from the paper, and extensively annotate the model as well. However, until the advent of SED-ML, it was impossible to store what the curator did to reproduce the model in a standard format, and until more widespread use of SED-ML, it was impossible to reliably validate any SED-ML that was produced. The Center for Reproducible Biomedical Modeling has produced new SED-ML interpreters and validators that have bridged this gap, and we have partnered with the EBI to 'retro-curate', as far as possible, the curated branch of BioModels, to include validated SED-ML, which we have then tested using the SED-ML interpreters on multiple simulation engines. In addition, we have extended the Antimony modeling language, and present the Antimony Web Editor, with particular features useful for adding curation of species, reactions, and parameters.

1.5 CompuTiX: A library for agent based modeling (not only) at a tissue-scale

Jiří Pešek, Jules Dichamp, Peter Kottman, Boulitrop Charles, Dirk Drasdo, INRIA, team SIMBIOTX, France

In recent years, many studies have shown that the tissue microarchitecture along with the mechanical environment has a crucial yet poorly understood impact on the biological processes inside living tissues This have a significant impact on progression of any potential disease or treatment. The limitations of in-vivo imaging techniques together with the small scale and isolated nature of many in-vitro experiments, makes these systems a suitable candidate for in-silico approach, where initial in-vitro experiments can be used to formulate and tune the underlying models and in-vivo imaging is then used to generate a patient specific setup. In particular, an agent based models, where the global effect is achieved by interaction between many, relatively simple, entities, are suitable to capture the spatial and behavioral heterogeneity and complexity of living tissues. In this talk we will present a new open-source computational library, CompuTiX, suitable for agent based simulations of tissues, organoids and more. We will split the talk into two parts. In the first part we will briefly introduce basic bio-physical models starting from simple center based models to more complex models like deformable cell model. In the second, more technical, part we will discuss the architecture of the library, design choices, trade-offs and challenges in our goal to provide a versatile and extensible platform for agent based simulations.

1.6 MeDaX – two years towards bioMedical Data eXploration

Judith AH Wodke, University Medicine Greifswald

Research based on clinical care data is gaining attention across the world. However, the quality of clinical care data is generally not maximised for research purposes. Instead, according to economic principles, medical staff and time costs are commonly minimised, rendering the enrichment with sufficient metadata for easy data reuse at least challenging. In addition, a heterogeneous landscape of laws concerning medical data reuse on national, state, and county levels make (international) interoperability an ambitious aim. The MeDaX project was initiated about

two years ago and its underlying idea presented at COMBINE 2022: connect and semantically enrich highly diverse clinical and other biomedical data in knowledge graphs (KG) to design, implement, and use graph technologies for innovative data exploration. The MeDaX-KG prototype has been designed and implemented building on the BioCypher framework to harmonise biomedical knowledge graphs and using synthetic patient data. The proof of concept pipeline consists of i) a FHIR input adapter, including an optimisation module for the generically generated graph structure, ii) a semi-automatic data schema generation based on the BioLink ontology, and iii) the visualisation of the resulting MeDaX-KG using Neo4j. Currently, the pipeline is improved, a user interface is implemented, and the first pilot in a german university clinic's data integration center is set up while the first stable release is prepared. To particularly tackle the gaps between the different scientific domains in medical informatics, we got involved in or coordinate several community projects. The international MIRAPIE community project aims to propose a provenance standard for biomedicine by defining a MInimal Requirements for Automated Provenance Information Enrichment guideline (`https://codeberg.org/MIRAPIE/MIRAPIE`). Participating in the BioCypher project, we adopted and are currently adapting the Biomedical Resource Ontology (BRO) (`https://github.com/biocypher/biomedical-resource-ontology`) to FAIRify our own software but also to allow the better classification of biomedical data. Within the Medical Informatics Initiative (MII) Germany we are coordinating the FAIRification of the MII core data set (`https://github.com/medizininformatik-initiative`). In summary, we are aligning several interdisciplinary efforts towards exploration of high quality clinical care data for biomedical research.

1.7 Reproducible tools for dealing with highly variable data

Nicole Radde, University of Stuttgart, Germany

In the biomedical context, data is often sparse, and replicates show a high variability. This is because complex procedures, costs, and ethical aspects constrain measurements. Sparsity and high variability pose a challenge for modeling, especially when building models aiming to capture quantitatively dynamic responses. Here, we present two complementary approaches we developed in our group to deal with sparse and variable data. Bayesian Modeling of Time Series Data (BayModTS) uses a Bayesian approach and a simulation model to process sparse and highly variable serial data (Höpfl et al., 2024). BayModTS can be used to quantify uncertainty in the observed process or as a noise filtering approach, as we will demonstrate with selected examples. Second, Eulerian Parameter Inference (EPI) formulates the parameter estimation problem for a simulation model from experimental data as a stochastic inverse problem and infers a parameter distribution that can reproduce the variability of the input data (Wagner et al., 2024). Both approaches are implemented as documented software packages that use standards such as SBML (Keating et al., 2020) or PEtab (Schmiester et al., 2021). In my talk, I will briefly explain our methods and discuss the current challenges regarding reproducibility and FAIR principles from a modeler's perspective.

1.8 The past, present and possible futures

Herbert Sauro, University of Washington, USA

It has almost been 25 years since Hiroaki Kitano initiated the development of SBML as part of the ERATO project. Together with Bolouri, Doyle, Finney, Hucka, myself and a number of key stakeholders (who continue to meet at COMBINE), we published the first draft and software support libraries for the SBML specification. Around the same time we also saw the publication of the specification for CellML that was a more mathematically oriented proposal. What resulted was most unexpected, the emergence of a new vibrant ecosystem which stimulated further development, created a host of new ancillary standards as well as the indispensable BioModels repository. That ecosystem still exists today. In this talk I will review what I feel remains to be done or is incomplete, what new modeling challenges we face, and describe what the center of model reproducibility in the US is doing in terms of software provision. In particular I will describe a number of new client-based web tools and desktop apps. The client-based tools are unusual in that they can be hosted from any free basic server such as a GitHub, Neocities or Cloudflare page. This makes such apps very low maintenance and tend to persist long after funding stops. Examples from our center include a model annotation (AWE) platform, a simple model checking app (ratesb), a high speed BioModels cache, a reproducibility portal, a model verification service, a new SBML/Antimony web utility, a Biosimulators/Biosimulations repository, a new SBML compliant desktop app, a number of new

Python packages for network visualization, a new desktop network editor (Alcuin), new extensions to Antimony (See talk by Lucian Smith), a standard protocol for multi-scale modeling (See talk by Eran Agmon), and the first model credibility hackathon held this summer.

1.9 Computational design of biological receivers using multi-scale models and data standards

Goksel Misirli, Keele University, UK

Engineering genetic regulatory circuits that sense external molecules and respond is essential for developing diverse biological applications. As the complexity of designs increases, a model-driven design process becomes desirable to explore large design spaces that involve different biological parts and parameters. Moreover, the amount of these molecules reaching a receiver is usually assumed to be constant, and the diffusion dynamics and the interference caused by late-arriving molecules and the cellular dynamics are often not integrated. Additionally, each molecule type may represent a single biological signal and be unsuitable for encoding and decoding multiple data bits. Here, we present the virtual parts repository, a computational framework that provides modular, reusable and composable models. The framework facilitates automating the design of predictable applications via simulations. It builds on the Systems Biology Markup Language to model cellular behaviour and the Synthetic Biology Open Language to capture the details of genetic circuits. We then extend this automation approach to design the end-to-end transmission of signalling molecules from a transmitter to cellular receivers for multi-bit data communications. The resulting framework can be used to understand the cellular response for a sequence of custom data bits, each representing a group of molecules released from a transmitter and diffusing over a molecular channel. The framework validates and verifies various communication parameters and identifies the best communication scenarios. We also present a novel algorithm to minimise signal interference by employing equalisation techniques from communication theory. Our data standards-enabled and multi-scale modelling workflow combines engineering genetic circuits and molecular diffusion dynamics to encode and decode data bits, design efficient cellular signals, minimise noise, and develop biologically plausible applications.

1.10 Networks, simple models and model diversity in the description of biological systems

Marc Hütt, Constructor University Bremen, Germany

My talk will address three distinct, but interrelated, topics: (1) networks as structural models to interpret high-throughput data; (2) the distinction between mathematical models and their computer implementations; (3) simple models vs. complicated, parameter-rich models.

Systems biology and systems medicine frequently use network-based strategies for data interpretation and data contextualization. These methods, at times, lack standardization and comparability. Here I briefly discuss, how such methods work, and which implicit hypotheses are associated with them.

The formal representation of a mathematical model is often incomplete, compared to the details required for an implementation of the model to run numerical simulations. Implementation differences can in principle lead to drastically different results. For the case of models of excitable dynamics, I illustrate this point, showing that even the simplest models can display such implementation differences.

Lastly, residing on the topic of simple models, I briefly draw the attention to the co-existence of parameter-rich and simple models of biological systems, outlining a few pros and cons and caveats.

2 Lightning Talks

2.1 Computational Model Development Using SBML: sbmlutils, sbm4humans, cy3sbml

Matthias König, Humboldt-University Berlin, Germany

The Systems Biology Markup Language (SBML) (Keating et al., 2020) is recognized as the standard framework for representing and exchanging complex mathematical models in biological

systems research. SBML facilitates the depiction of a diverse array of biological phenomena, encompassing metabolic networks, signaling pathways, and regulatory networks. It is versatile enough to handle models ranging from simple individual processes to intricate multi-scale representations.

One of the primary challenges faced by newcomers in computational biology is the encoding and development of ordinary differential equation (ODE) models within the SBML framework. Addressing this hurdle, we introduce two innovative Python tools: sbmlutils (`https://github.com/matthiaskoenig/sbmlutils`), sbml4humans (`https://sbml4humans.de`), and the Cytoscape application cy3sbml (`https://github.com/matthiaskoenig/cy3sbml`). These tools collectively streamline the process of SBML model creation, enhancing both the programmatic aspect and the user experience. Specifically, sbmlutils facilitates the programmatic construction of SBML models, while sbml4humans generates user-friendly reports for model interpretation. Furthermore, cy3sbml integrates with Cytoscape to offer advanced visualization capabilities, thereby augmenting the comprehension and analysis of SBML-encoded models.

These advancements significantly contribute to the ease of SBML model development and interpretation, fostering greater accessibility and understanding for those entering the field of computational systems biology.

2.2 Utilizing Nix for rapid BayModTS development

Simon Hauser and Fritz Otlinghaus, Helsinki Systems, Germany

BayModTS, a Python project for FAIR Bayesian Modelling of Time Series workflows, has a complex setup that requires users of that software to install and compile multiple Python packages that have native C dependencies (Höpfl et al., 2024). This is a complex endeavor and currently it is not possible to fully resolve these issues using poetry install. We present a solution that utilizes a general purpose package manager called Nix, that guarantees that a package and all its dependencies can be built reproducibly. This package manager can be used to build all kinds of software packages, including C libraries and Python packages, which we need to realize our solution. We utilize prepackaged system and Python dependencies already made available by the Nix community to build new, complex packages, like libsbml, libroadrunner, tellurium, and others, to realize a portable and reproducible local development environment. In the end, we automate our solution by using a well established GitHub continuous integration solution that builds all packages and makes them available via HTTP using a binary cache. This can be used so that packages no longer have to be build locally and can be downloaded from CI instead, improving the onboarding process for new team members and simplifying collaborations for external researchers.

2.3 Biological and Biophysics Simulation in Tissue Forge

T.J. Sego, University of Florida, USA

Tissue Forge is open-source simulation software for interactive particle-based physics, chemistry and biology modeling and simulation. Tissue Forge allows users to create, simulate and explore models and virtual experiments based on soft condensed matter physics at multiple scales, from the molecular to the multicellular, using a simple interface. While Tissue Forge is designed to simplify solving problems in complex subcellular, cellular and tissue biophysics, it supports applications ranging from classic molecular dynamics to agent-based multicellular systems with dynamic populations. Tissue Forge users can build and interact with models and simulations in real-time and change simulation details during execution, or execute simulations off-screen and/or remotely in high-performance computing environments. Tissue Forge provides a growing library of built-in model components along with support for user-specified models during the development and application of custom, agent-based models. Tissue Forge includes an extensive Python API for model and simulation specification via Python scripts, an IPython console and a Jupyter Notebook, as well as C and C++ APIs for integrated applications with other software tools. Tissue Forge supports installations on Windows, Linux and MacOS systems and is available for local installation via conda. The talk complements a tutorial at COMBINE 2024 that intends to introduce the basic concepts, modeling and simulation features, and some relevant modeling applications of Tissue Forge through guided simulation scripting.

2.4 A COMBINE Standard for Multi-Approach Multi-Scale (MAMS) Modelling

Sheriff Rahuman, EMBL-EBI, UK, and T.J. Sego, University of Florida, USA

Multi-approach Multi-scale (MAMS) modelling represents a cutting-edge method for modelling and analysis of biological systems, leveraging an integrated suite of diverse modelling frameworks. This multi-approach modelling will encompass a combination of diverse modelling formalisms, such as ordinary differential equations (ODE), partial differential equations (PDE), logical, constraint-based, and agent-based models across multiple scales. These models are intricately tied together to facilitate complex simulations. During the dedicated breakout sessions at Harmony 2021, COMBINE 2021, and HARMONY 2024, we delved into the existing state-of-the-art technologies and standards, including SBML and SED-ML, and their support for multi-approach modelling. These discussions also illuminated the current challenges and gaps within the field. For COMBINE 2024, our objective is to further this conversation by identifying published MAMS models and bringing together the community to enable the creation of novel standards or the enhancement of existing COMBINE standards to support MAMS. This effort aims to foster interoperability and support the rapidly evolving paradigm of MAMS modelling. The talk complements a breakout session at COMBINE 2024 that intends to continue the described work and invite new collaborators to join continuing efforts.

2.5 Continuing Work Towards Reproducible Stochastic Biological Simulation

T.J. Sego, University of Florida, USA, and Sheriff Rahuman, EMBL-EBI, UK

Stochastic simulations are commonly used to quantitatively or semi-quantitatively describe the dynamics of biological systems. At various scales and in multiple applications, stochastic simulation better reflects observed biological processes and robustness. Various methods are widely used to incorporate stochasticity into biological simulation, such as the Gillespie stochastic simulation algorithm for systems biology modeling, stochastic Boolean networks for network modeling, and the Cellular Potts model methodology for multicellular modeling. Proving reproducibility of simulation results is critical to establishing the credibility of a model. To this end, BioModels, the largest repository of curated mathematical models, tests and reports the reproducibility of simulation results for all submitted models when possible. A recent study showed that about 50% of the deterministic ordinary differential equation models on BioModels could not be reproduced when applying criteria for reproducibility to the information provided in their associated publication, reflecting a current crisis of reproducibility. Furthermore, there are no well-accepted metrics or standards for reproducing stochastic simulation results, thus perpetuating the crisis of reproducibility for a broad class of biological models. This lightning talk survey recent progress to establish an accepted framework for testing the reproducibility of stochastic simulations in biological modeling. The talk will provide a brief overview of recent progress towards defining quantitative measures to determine whether stochastic simulation results can be reproduced, and when results have been reproduced. The talk complements a breakout session at COMBINE 2024 that intends to continue the described work and invite new collaborators to join continuing efforts.

2.6 Morpheus model repository: Experiences with reproducible multi-cellular models

Lutz Brusch, Jörn Starruß, Diego Jahn, and Robert Müller, Technische Universität Dresden, Germany

Collaborative modeling and simulation become increasingly important for studying self-oganization, patterning, morphogenesis and disease processes from the intracellular to the tissue and organ scales. To support collaborations, we have developed the Morpheus model repository (`https://morpheus.gitlab.io/models`). This model repository is an open access and citable platform for publishing, sharing and archiving multi-scale and multi-cellular models that are encoded in the model description language MorpheusML (`https://doi.org/10.25504/FAIRsharing.78b6a6`). We will explore statistics and examples of the usage of the Morpheus model repository. Different simulators like Artistoo (`https://artistoo.net/converter.html`) and Morpheus (`https://morpheus.gitlab.io`) can process MorpheusML models from the repository. Among them, the model editor and simulator Morpheus is open source and allows to develop multi-scale models in a modular manner and manage the entire workflow through a user-friendly GUI. Moreover, Mor-

pheus is SBML-compliant, supports simulations based on experimental data, e.g. segmented cell configurations, and is integrated with the FitMultiCell toolbox for robust and efficient parameter estimation of stochastic models (`https://gitlab.com/fitmulticell/fit`, `https://doi.org/10.1093/bioinformatics/btad674`). Importantly, the strict separation of the model description in MorpheusML from any solver code allows to readily reproduce model results locally in different (future) versions of the simulation software. Beyond reproduction of published results, MorpheusML models can easily be extended (copy-paste parts between models) and merged among each other, thus vitalizing model reuse and exchange.

2.7 Modeling and simulation using industrial standards Modelica, FMI and web components.

Tomas Kulhanek, Charles University, Prague, Czech Republic

We use industrial standard Modelica to express complex models of human physiology (Mateják et al., 2014; Ježek et al., 2017). Recently we have published enabling technology that allows to export complex models in standard functional mockup interface API (FMI) as a web component to be integrated with other web standards and technologies to create modern web application (Kulhanek et al., 2023) (`https://bodylight.physiome.cz`). Thanks to it the models does not necessarry need to be implemented in Modelica language, but and standard FMI needs to be implemented by other standards to compute model derivatives and do simulation step using a prefered numerical method.

In the exemplar case report of metabolic disorder we will demonstrate the process of creating a component models, export them as web component and integrate with other web components or web standards to create interactive application. Model implementation of cardioavscular system, respiratory and blood gas exchange in Modelica will be used. Co-simulation and enriched with chart and numbers presented in virtual monitor of vital signs will control application flow. The foundation of technologies are published with open source license and thanks to chain of scientific and/or industrial standards and tools. The resulting simulator can be executed on any device without the need to install special software. Platform needs only modern web browser and supported are Windows/Linux/iOS computers, mobile phones, tablets, virtual reality headset, etc. Exemplar application with accompanied learning material to learn pathophysiology of metabolic disorder is available online at `https://egolem.online/dka/`.

2.8 A Standardized Protocol for Integrative, Multiscale Modeling

Ion Moraru, and Eran Agmon, University of Connecticut, USA

We are developing a standardized protocol for multi-algorithmic model composition, based on standardized schemas for process interfaces, composition patterns, and orchestration patterns. This will provide the foundation for robust infrastructure for systems biology models. The BioSimulators project aims to establish this protocol, ensuring reproducibility, tool compatibility, and "plug-and-play" integration of new processes and data. Software tools built around these schemas can include databases, applications, graphical user interfaces, and simulation tools, supported by both local and remote operations, such as containerized and web-based services. By aligning with existing standard formats like SBML and CellML, standard formats for spatial models, and multi-cellular models, the protocol can foster a unified approach that connects these efforts. This approach addresses many challenges by advancing the FAIR (Findable, Accessible, Interoperable, Reusable) principles, allowing researchers to more reliably find simulation modules, understand those models, and connect them reliably into hybrid, multiscale models. We initiated the project with a Verification API, that brings together COPASI, Tellurium, and AMICI–each of them fit with a standardized process interface for uniform timescourses–load them with the same SBML model and simulation instructions, runs them in parallel, and compares results.

2.9 openTECR: community curation of Thermodynamics of Enzyme-Catalyzed Reactions

Robert Giessmann, Institute for Globally Distributed Open Research and Education (IGDORE), Berlin, Germany

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

We create a data collection of apparent equilibrium constants of enzyme-catalyzed reactions, being reliable, open and machine-actionable, with a clear change process to integrate new data and correct errors. We believe that Open Science principles, and specifically Open Data and Open Source are key to achieving our vision.

The openTECR database serves computational and experimental scientists in the fields of metabolic engineering, genome-scale metabolic modelling, biocatalysis and related fields by providing curated information. It is used by eQuilibrator as the data basis for making predictions about any possible reaction.

Recently, we organized an open community curation effort (`https://opentecr.github.io/invitation-to-curate`). We prepared a curation workflow to analyze 278 pages of pages densely packed with tables and textual information. We invited volunteer contributions, and are immensely grateful about 17 volunteers investing almost 100 working hours.

At the COMBINE 2024 meeting, I would like to present our initiative and share our lessons learned about organizing successful community curation. I believe that our example can serve as a blueprint for other databases / project ideas which require a large amount of working hours.

We discovered that key to receiving contributions is to offer very small packages of work and a detailed curation manual. Our smallest task was 3 minutes long and well received.

Our small community (40 members) shares a mailing list (`https://w3id.org/opentecr`) and a GitHub organization where we store our data and code under open licenses (`https://github.com/opentecr/`).

2.10 OpenVT – Developing Framework Description Standards for MultiCellular Agent-Based Virtual Tissue Models

James Glazier, Indiana University, USA

Many simulation frameworks implement multicellular agent-based models using a variety of methodologies (center model, vertex model, Cellular Potts model, finite-element mechanics, ...) and support a variety of biological and mathematical processes it can be often confusing and time consuming to for a researcher to know which simulation framework can fulfill their modeling needs. In our breakout session, we will discuss an approach to defining and categorizing simulation framework capabilities. The session will start with an overview of the various methods employed in multicellular simulations, highlighting their unique features and common challenges. It will present our approach to describing framework descriptions in a standardized way followed by discussion on this approach.

2.11 OpenVT: MultiCellular Agent-Based Virtual Tissue Models: Defining Topics and Priorities for Working Groups and Virtual Workshops

James Glazier, Indiana University, USA

Virtual Tissues (VT), agent based multicellular modeling has become indispensable in understanding complex biological phenomena, from tissue development to disease progression. But the diversity in simulation methods poses challenges in reproducibility, modularity, reusability, and integration for multiscale models, leading to a fragmented ecosystem and hindering growth. The OpenVT Community is trying to address these challenges by bringing siloed research groups together to improve the sharing of VT knowledge. The OpenVT Community supports the expansion of and broader adoption of multicellular modeling beyond academic research labs into greater industry practice. Development of best practices and better reproducibility will ultimately lead to models that more closely follow FAIR (Findable, Accessible, Interoperable, and Reusable) principles, leading to wider use in therapeutic approaches, toxicology, drug discovery and personalization of testing and treatment.

2.12 OpenVT – Developing Reference Models for Multicellular Agent-Based Virtual Tissue Models

James Glazier, and James Osborne (University of Melbourne, Australia), Indiana University, USA

An increasing number of packages implement multicellular agent-based models using a variety of methodologies (center model, vertex model, Cellular Potts model, finite-element mechanics, ...). In principle a set of underlying biological and physical processes should yield the same result independent of the package in which they are implemented. However, at the moment, comparison between methodologies or even between different packages implementing the same methodology, is quite challenging. As a first step to building a shared understanding of modeling capabilities and to improve rigor and reproducibility, we define a minimal set of standard reference models which should be implemented in each framework to illustrate their capabilities and reveal hidden discrepancies of approach. This talk will discuss these efforts and introduce our planned breakout session.

2.13 A functional tissue unit approach to understanding lung function in health and disease

Ruobing Li, Alys Clark, Merryn Tawhai, David Nickerson, Kelly Burrowes, Auckland Bioengineering Institute, University of Auckland, New Zealand

The primary functional tissue unit of the lungs is the acinus. An acinar unit brings together diverse functions, including airflow, blood flow, gas exchange, mechanical deformation and the effect of surfactant on this, and fluid transport from the blood to the lymphatic vessels. Surfactant is important in reducing alveolar surface tension, ensuring stability and preventing alveolar collapse. Ventilation is driven by the dynamic processes of lung tissue expansion and recoil during breathing. Regional tissue recoil pressures also influence pulmonary perfusion, impacting the distribution of blood flow within the lung. The lymphatic system, integral for maintaining fluid balance and optimal immune function, is affected by these mechanical forces too. The interdependence of these factors is vital for maintaining optimal pulmonary function under physiological conditions. Existing models of varying geometric complexity have been developed to simulate lung mechanical behaviours and various fluid transport, currently as separate systems. The ventilation model of Swan et al. [J Theor Biol. 2012; 300:222-31] combines lung airway structure and tissue mechanics with airflow dynamics. A perfusion model by Clark et al. (2010) simulates pulmonary blood flow within the vasculature. A lung lymphatic model developed by Ashworth et al. (2023) estimates the transfer of fluid from the capillary blood vessels into the interstitial space and the lymphatic vessels. The surfactant model based on Otis et al. (1994) work simulates the dynamic adsorption-desorption process at the air-liquid interface and estimates the impact of surfactant on tissue compliance. These different models reflect the diverse functions occurring within each acinus that work together to determine emergent lung function. However, there is no comprehensive model that integrates these aspects to form a complete functional tissue unit (FTU) of the lung. This study addresses this gap by developing a respiratory FTU that integrates these different models to simulate acinar function and link this to represent whole lung function. Model implemented by CellML, Fortran, and Python. By integrating these individual models, we aim to provide a better understanding of the interactions and dependencies within the lungs, essential for simulating lung function in health and disease.

3 Breakouts

3.1 Workshop: refineGEMs and SPECIMEN for automated model reconstruction and annotations

Gwendolyn O. Döbel, Martin Luther University Halle-Wittenberg, Germany

"Metabolic model reconstruction usually relies on several cumbersome steps. Different tools exist, which are only partially automated and need to be connected manually. Our aim is to simplify and reduce the manual workload. Thus, we developed the toolbox refineGEMs and the workflow collection SPECIMEN. A stable release of refineGEMs was already used in practice (Bäuerle et al., 2023). Both tools are currently under active development (enhancement and extension).

This workshop aims to give the attendees a brief introduction to automatic metabolic modelling with the tools refineGEMs (`https://github.com/draeger-lab/refinegems`) and SPECIMEN (`https://github.com/draeger-lab/SPECIMEN`). As part of the workshop, an open discussion will be held about issues arising from automatic energy-generating cycle (EGC) dissolution and gap filling.

3.2 Introduction into Nix for scientific software

Simon Hauser and Fritz Otlinghaus, Helsinki Systems, Germany

Working on software in a team brings all kinds of challenges, especially because everyone has a slightly different development environment. These challenges usually start with onboarding new team members, include complications of moving your local environment to a high performance cluster and end in unreproducible bugs that boil down to ""works on my machine"". Some of these issues can be resolved by providing dependency pinning using poetry or other package managers, but these solutions do not cover the operating system and require additional install documentation that usually contains apt usage. Nix is a general purpose package manager that emerged in the last couple of years that solves these issues, by not just pinning the version of dependencies but also system libraries and tools, like the glibc library, python and also python packages. This session will cover the fundamentals of Nix, including installation, command usage and writing your own custom development environment for a specific software. Participants will learn how to leverage Nix to create reproducible scientific workflows, manage dependencies, and ensure consistent software environments across different systems. Through practical demonstrations and hands-on activities, attendees will gain the skills necessary to integrate Nix into their scientific projects, enhancing both the reliability and portability of their software. Join us to discover how Nix can streamline your scientific software development and deployment processes, fostering greater collaboration and innovation in your research endeavors.

3.3 Biological and Biophysics Simulation in Tissue Forge: Introduction and Guided Simulation Building

T.J. Sego, University of Florida, USA

Tissue Forge is open-source simulation software for interactive particle-based physics, chemistry and biology modeling and simulation. Tissue Forge allows users to create, simulate and explore models and virtual experiments based on soft condensed matter physics at multiple scales, from the molecular to the multicellular, using a simple interface. While Tissue Forge is designed to simplify solving problems in complex subcellular, cellular and tissue biophysics, it supports applications ranging from classic molecular dynamics to agent-based multicellular systems with dynamic populations. Tissue Forge users can build and interact with models and simulations in real-time and change simulation details during execution, or execute simulations off-screen and/or remotely in high-performance computing environments. Tissue Forge provides a growing library of built-in model components along with support for user-specified models during the development and application of custom, agent-based models. Tissue Forge includes an extensive Python API for model and simulation specification via Python scripts, an IPython console and a Jupyter Notebook, as well as C and C++ APIs for integrated applications with other software tools. Tissue Forge supports installations on Windows, Linux and MacOS systems and is available for local installation via conda. This tutorial introduces the basic concepts, modeling and simulation features, and some relevant modeling applications of Tissue Forge through guided simulation scripting. Tutorial concepts will introduce basic Tissue Forge modeling concepts and simulation features through the development of interactive simulations in Python. Attendees are encouraged, but not required, to code along as the tutorial interactively develops and tests simulations in multicellular and biophysics modeling applications.

3.4 A COMBINE Standard for Multi-Approach Multi-Scale (MAMS) Modelling

Sheriff Rahuman, EMBL-EBI, UK

Multi-approach Multi-scale (MAMS) modelling represents a cutting-edge method for modelling and analysis of biological systems, leveraging an integrated suite of diverse modelling frameworks. This multi-approach modelling will encompass a combination of diverse modelling formalisms, such as ordinary differential equations (ODE), partial differential equations (PDE), logical, constraint-based, and agent-based models across multiple scales. These models are intricately tied together to facilitate complex simulations. During the dedicated breakout sessions at Harmony 2021, COMBINE 2021, and HARMONY 2024, we delved into the existing state-of-the-art technologies and standards, including SBML and SED-ML, and their support for multi-approach modelling. These discussions also illuminated the current challenges and gaps within the field. For COMBINE 2024,

our objective is to further this conversation by identifying published MAMS models and bringing together the community to enable the creation of novel standards or the enhancement of existing COMBINE standards to support MAMS. This effort aims to foster interoperability and support the rapidly evolving paradigm of MAMS modelling.

3.5 Continuing Work Towards Reproducible Stochastic Biological Simulation

T.J. Sego, University of Florida, USA

Stochastic simulations are commonly used to quantitatively or semi-quantitatively describe the dynamics of biological systems. At various scales and in multiple applications, stochastic simulation better reflects observed biological processes and robustness. Various methods are widely used to incorporate stochasticity into biological simulation, such as the Gillespie stochastic simulation algorithm for systems biology modeling, stochastic Boolean networks for network modeling, and the Cellular Potts model methodology for multicellular modeling. Proving reproducibility of simulation results is critical to establishing the credibility of a model. To this end, BioModels, the largest repository of curated mathematical models, tests and reports the reproducibility of simulation results for all submitted models when possible. A recent study showed that about 50% of the deterministic ordinary differential equation models on BioModels could not be reproduced when applying criteria for reproducibility to the information provided in their associated publication, reflecting a current crisis of reproducibility. Furthermore, there are no well-accepted metrics or standards for reproducing stochastic simulation results, thus perpetuating the crisis of reproducibility for a broad class of biological models. This breakout session will continue work towards establishing an accepted framework for testing the reproducibility of stochastic simulations in biological modeling. The session will provide a brief overview of recent progress towards defining quantitative measures to determine whether stochastic simulation results can be reproduced, and when results have been reproduced. Attendees will discuss current issues to address towards consensus and broad adoption in relevant modeling communities, as well as future work towards reproducibility of stochastic simulation results using multiscale and complex models.

3.6 Morpheus: A user-friendly simulation framework for multi-cellular systems biology

Lutz Brusch and Jörn Starruß, Technische Universität Dresden, Germany

Multi-cellular modeling and simulation become increasingly important to study tissue morphogenesis and disease processes. This tutorial introduces Morpheus (`https://morpheus.gitlab.io`) in an overview presentation with live demos and hands-on exercises runnable in sync on the presenter's and your own laptop. The focus lies on importing SBML models into Morpheus, extending them in space as reaction-diffusion processes and automatically ""cloning"" them into many individual cells that can dynamically interact. Also, own modeling ideas can be explored with the help of a tutor. Morpheus offers modeling and simulation of multi-cellular dynamics in a Graphical User Interface (GUI) without the need to program code. It uses the domain-specific language MorpheusML to define and simulate multicellular models in 3D space including the most common cell behaviors and tissue mechanics. Morpheus is open-source software and provides readily installable packages for macOS, Windows, Linux (`https://morpheus.gitlab.io/download/latest/`). Please download before the tutorial and have a look around the homepage incl. ¿90 example models.

3.7 Combine spatial multi-cellular modelling with SBML

Jörn Starruß, TU Dresden, Germany

Modularity is key to creating complex multi-cellular models while preserving the accessibility of meaningful submodels. Naturally, composition also encourages reusability and the likes. We want to discuss and establish a common practice how to overlay the spatial dynamics of multi-cellular models with reaction dynamics defined in the SBML standard.

Most obvious features to be represented separately from the spatial cell dynamics are intracellular regulatory systems, inter-cellular communication and spatial reaction-diffusion processes using SBML-spatial (`https://sbml.org/documents/specifications/level-3/version-1/spatial/`).

Further issues arise when inter-connecting identical submodels residing in individual cells and the definition of instantaneous assignments upon entity operations (e.g. cell birth and death).

As an introductory motivation we will present our latest Morpheus (`https://morpheus.gitlab.io`) results in embedding spatial reaction-diffusion submodels within moving cells. Using that experience we will sketch a way how to exploit the HMC package (`https://sbml.org/documents/specifications/level-3/version-1/comp/`) to compose SBML models and attach them in a second step to the individual scopes of our spatial model. We hope for a lively discussion on best practice approaches interconnecting spatial multi-cellular modeling and the SBML standard.

3.8 Training Models using PEtab

Fabian Fröhlich, The Francis Crick Institute, UK

PEtab is a standardized file format for specifying parameter estimation problems (Schmiester et al., 2021). The interoperable format is currently supported by 11 different tools (`https://github.com/PEtab-dev/petab#petab-support-in-systems-biology-tools`), enabling users to benefit from standardized parameter estimation across frameworks based in Python, Julia, R, MATLAB, C++, or GUIs.

Although PEtab was initially developed for parameter estimation, recent efforts have extended the format to improve standardization of various adjacent tasks, including: model selection, multi-scale modeling, PKPD and NLME modeling, optimal control, and visualization.

In this breakout session, based on audience interests, we will present introductions to PEtab and its extensions, then discuss current efforts to improve PEtab. People unfamiliar with PEtab are welcome to attend, and might first like to check out the tutorial (`https://petab.readthedocs.io/en/latest/tutorial.html`).

3.9 SBGN PD: current and future development

Adrien Rougny, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg

Visualization of biological processes plays an essential role in life science research. Over time, diverse forms of diagrammatic representations, akin to circuit diagrams, have evolved without well-defined semantics potentially leading to ambiguous network interpretations and difficult programmatic processing. The Systems Biology Graphical Notation (SBGN) standard aims to reduce ambiguity in the visual representation of biomolecular networks. It provides specific sets of well-defined symbols for various types of biological concepts. SBGN comprises three complementary languages: Process Description (PD), Entity Relationship (ER), and Activity Flow (AF). The XML-based SBGN Markup Language (SBGN-ML) facilitates convenient storage and exchange of SBGN maps. The SBGN languages as well as SBGN-ML are described in detail in specifications (see sbgn.org). This breakout session will focus on the development of SBGN PD. We invite all participants interested in SBGN to join this session, where we will discuss specific issues related to the next version of the PD specification, as well as more open issues related to a future level of SBGN PD.

3.10 Developing a proof of concept for a heap of git-versioned json files as a FAIR alternative to relational databases

Robert Giessmann, Institute for Globally Distributed Open Research and Education (IGDORE), Berlin, Germany

I propose, if there is interest of fellow participants, to think about, and just try out, a heap of json files, versioned in git, as an alternative form to store structured data.

Relational databases are great, but hard for "non-computational people" (read as: the typical experimental, wet-lab person) to create and to change. A heap of json files on GitHub seems still far fledged for some of those persons, but might be the minimal necessary technical barrier they have to cross.

Of course, data inside those json files must, for instance, be "quality controlled", i.e. checked for correctness and sticking to schemata. But that could be implemented as CI/CD actions. Git itself might put practical limits on the general feasibility of that idea – which is to be tested out.

If anyone else is up for it, let's try it out!

3.11 OpenVT – Developing Framework Description Standards for MultiCellular Agent-Based Virtual Tissue Models

James Glazier, Indiana University, USA

Many simulation frameworks implement multicellular agent-based models using a variety of methodologies (center model, vertex model, Cellular Potts model, finite-element mechanics, ...) and support a variety of biological and mathematical processes it can be often confusing and time consuming to for a researcher to know which simulation framework can fulfill their modeling needs. This session will discuss an approach to defining and categorizing simulation framework capabilities. The session will start with an overview of the various methods employed in multicellular simulations, highlighting their unique features and common challenges. It will present our approach to describing framework descriptions in a standardized way followed by discussion on this approach.

3.12 OpenVT: MultiCellular Agent-Based Virtual Tissue Models: Defining Topics and Priorities for Working Groups and Virtual Workshops

James Glazier, Indiana University, USA

Virtual Tissues (VT), agent based multicellular modeling has become indispensable in understanding complex biological phenomena, from tissue development to disease progression. But the diversity in simulation methods poses challenges in reproducibility, modularity, reusability, and integration for multiscale models, leading to a fragmented ecosystem and hindering growth. The OpenVT Community is trying to address these challenges by bringing siloed research groups together to improve the sharing of VT knowledge. The OpenVT Community supports the expansion of and broader adoption of multicellular modeling beyond academic research labs into greater industry practice. Development of best practices and better reproducibility will ultimately lead to models that more closely follow FAIR (Findable, Accessible, Interoperable, and Reusable) principles, leading to wider use in therapeutic approaches, toxicology, drug discovery and personalization of testing and treatment. This session aims to discuss current progress undertaken by the OpenVT community towards a shared ecosystem and look to attendees for insight into what they believe will encourage broader adoption of community guidelines.

3.13 OpenVT – Developing Reference Models for Multicellular Agent-Based Virtual Tissue Models

James Glazier, Indiana University, USA

An increasing number of packages implement multicellular agent-based models using a variety of methodologies (center model, vertex model, Cellular Potts model, finite-element mechanics, ...). In principle a set of underlying biological and physical processes should yield the same result independent of the package in which they are implemented. However, at the moment, comparison between methodologies or even between different packages implementing the same methodology, is quite challenging. As a first step to building a shared understanding of modeling capabilities and to improve rigor and reproducibility, we define a minimal set of standard reference models which should be implemented in each framework to illustrate their capabilities and reveal hidden discrepancies of approach. This session will discuss these efforts and look for feedback from attendees.

4 Poster

4.1 SPECIMEN: Collection of Workflows for Automated and Standardised Reconstruction of Genome-Scale Metabolic Models

Carolin Brune, Gwendolyn O. Döbel, Famke Bäuerle (QBiC), Natia Leonidou (IBMI, DZIF, QBiC), Reihaneh Mostolizadeh (Justus Liebig University Gießen), and Andreas Dräger, Martin Luther University Halle-Wittenberg, Germany

SPECIMEN (`https://github.com/draeger-lab/SPECIMEN`) is an open-source collection of different workflows designed for the automated and standardised curation of genome-scale models. It is Python-based and integrates a variety of tools, including MCC, SBOannotator, refineGEMs, and BOFdat, and more, to concatenate modelling steps like gap filling, annotation, duplicate removal, and biomass normalisation into a single pipeline. SPECIMEN offers different workflows tailored to various modelling approaches and types of input data, facilitating efficient and consistent genome-scale model reconstruction.

4.2 What is an appropriate standard for modeling microbial communities?

Beatrice Ruth, Peter Dittrich (FSU Jena), and Bashar Ibrahim, Friedrich Schiller Universität Jena, Germany

Microbial communities exist almost everywhere on earth and play a major role in environmental processes as well as health and aging. A microbial community consists of a complex network of taxa and their interactions with one another. My task at the moment is to find a way to use measurements to understand the hierarchical structure and to predict interactions within the community. In this context, I noticed that there are no comprehensive standards for modeling microbial communities. Therefore, I would like to ask the question, what would be a suitable standard for modeling microbial communities? To represent the hierarchical structure, the intersections and unions of self-maintaining taxa combinations (organizations) are ordered in a lattice. While the intersections highlight the core microbiome and the general similarities between different organizations, additional unions reveal possible negative interactions. When working with measurements, in addition to the individual taxa composition, other environmental parameters such as at least time and location are important. Based on chemical organization theory, the individual organizations of a given reaction network and thus also their intersection and union sets can be calculated. The reverse path is now examined for interaction prediction. Here I am currently using resource-consumer models including toxins to also take negative interactions into account. Many different models can create the same hierarchical structure. Therefore, the focus is on a model for representation with as few resources as possible, assuming as few interactions as possible. I'm curious how this could be accounted for in a standard, as it would increase clarity and simplify exchange. Does it need more than, for example, SBML or a simple extension?

4.3 Integrate modelling standards with Energy-based System Analysis

Weiwei Ai, Peter Hunter, and David Nickerson, Auckland Bioengineering Institute, The University Of Auckland, New Zealand

Energy is a fundamental concept in physical processes. Physiological systems often involve various physical processes, with energy serving as a universal language across different domains. Energy-based modelling frameworks, such as bond graphs and port-Hamiltonian formulation, have been recently introduced into the computational biology community. The energy-based frameworks adopt a hierarchical and modular approach, which captures traceable energy storage, dissipation, and transduction, offering comprehensive insights into the systems under investigation.

The COMBINE community has established standards for computational biological models, such as CellML and SED-ML, which are used for encoding mathematical models and simulation experiments of physiological processes. This talk will explore the potential integration between these standards and energy-based approaches. We will demonstrate how we leverage these modelling standards to extract information from models for system analysis and invite input and suggestions from the community.

4.4 The preCICE v3 coupling library and the emerging preCICE ecosystem

Gerasimos Chourdakis, Jun Chen; Ishaan Desai, Carme Homs-Pons, Benjamin Rodenberg (Technical University of Munich), David Schneider, Miriam Schulte, Frédéric Simonis, and Benjamin Uekermann, University of Stuttgart, Germany

The coupling library preCICE (precice.org) allows coupling simulation codes at runtime, enabling flexible and efficient partitioned multi-physics simulations, exchanging data over point locations during a time loop. preCICE v1, released in 2016, introduced a high-level API, offering

massively parallel communication and mapping methods, as well as advanced IQN coupling algorithms. preCICE v2, released in 2020, included several integrations to established open-source codes (such as OpenFOAM, SU2, deal.II, FEniCS, and more), allowing users to execute simulations often without having to write any code. This increasing number of components and examples is now a citable preCICE Distribution. The latest preCICE Distribution (v2404) includes 23 components and 31 different example scenarios coupling multiple possible combinations of codes, featuring preCICE v3, which introduces several new features (such as time interpolation, much faster Partition-of-Unity RBF mapping methods, experimental Geometric Multiscale mapping), and a much simplified API. Common applications now extend far beyond fluid-structure interaction, now including applications in biomechanics (via codes such as OpenDiHu or FEBio), porous media (e.g., via DuMux), ice-sheet modelling, and more. This poster will summarize important updates in the preCICE project, and discuss current plans for integrating the community towards a community-driven ecosystem of FAIR (Findable, Accessible, Interoperable, and Reusable) components and simulation cases, opening up to new applications and scientific communities.

4.5 ModelPolisher: Enhancing the Quality and Completeness of Genome-Scale Metabolic Models (GEMs)

Dario Eltzner, Bahaa Ziadah, Thomas J. Zajac, Matthias König, Kaustubh Trivedi and Andreas Dräger, Computational Systems Biology of Infections and Antimicrobial-Resistant Pathogens, Institute for Biomedical Informatics (IBMI), Tübingen, Germany

Background:

Genome-scale metabolic models (GEMs) play a central role in systems biology and enable detailed simulations and predictions of metabolic processes. However, the quality and completeness of these models can vary considerably, which compromises their utility. GEMs often suffer from inconsistencies, incomplete annotations, and structural inaccuracies that can limit their utility. A measure of model quality, the MEMOTE score, often highlights these shortcomings, indicating areas such as missing gene associations, metabolite inconsistencies, and incorrect mass balances. To address this problem we developed ModelPolisher, a tool for standardizing, annotating and refining SBML models.

The results:

ModelPolisher has drastically improved the quality of the annotation of GEMs. By aligning model components with BiGG IDs, the tool enriches models with consistent and detailed metadata, facilitating model sharing and reproducibility. In addition, ModelPolisher's built-in checks for structural correctness, such as mass balance and metabolite connectivity, have proven effective in identifying and correcting errors. Our application of ModelPolisher to a number of models from the BiGG and BioModels databases has resulted in a major improvement in model quality and metadata completeness.

Conclusion:

ModelPolisher is a must-have tool for systems biology that addresses the critical need for high-quality, well-annotated GEMs. Its power to automatically enhance model metadata and ensure structural integrity not only improves model utility, but also promotes collaboration and data sharing. The tool's impact is particularly evident in large modeling projects where consistency and accuracy are paramount.

Availability:

ModelPolisher is open-source on GitHub at https://github.com/draeger-lab/ModelPolisher. It can be used from the command line or integrated into larger workflows and offers a flexible solution for researchers. Extensive documentation and examples make it easy to use, and the community is encouraged to contribute to the ongoing development and improvement of the tool.

4.6 Partitioned simulations using the neuromuscular simulation framework OpenDiHu

Carme Homs-Pons, and Miriam Schulte, University of Stuttgart, Germany

"OpenDiHu is a high-performance computing framework for skeletal muscle simulations. Created and developed at the University of Stuttgart, it is an open-source project written in C++. It uses python scripting and provides CellML support. OpenDiHu offers ready-to-use physics-specific solvers that can be combined by the user to create a tailored muscle solver. The available models include 3D finite element hyperelastic models, subcellular models and a motor neuron pool

model among others. The user does not have to worry about the mapping between models or time sub-cycling, as this is automatically done by the internal coupling tool from OpenDiHu. Besides, OpenDiHu has a preCICE adapter. preCICE is a coupling library for partitioned multi-physics simulations. Using preCICE, we can couple OpenDiHu to other software, e.g., FEBio and deal.ii. Finally, the poster will include use-cases to better showcase what we can do using OpenDiHu and preCICE. In particular, we will present our latest results for a human biceps simulation and our work-in-progress towards a model of the agonist-antagonist myoneural interface."

4.7 SBSCL: A Library of Efficient Java Solvers and Numerical Methods to Analyze Computational Models in Systems Biology

Arthur Neumann (Eberhard Karl University of Tübingen, Germany), Max Hatfield (German Center for Infection Research (DZIF), Quantitative Biology Center (QBiC), Eberhard Karl University of Tübingen), Taichi Araki (Graduate School of Science and Technology, Keio University, Japan), Akira Funahashi (Graduate School of Science and Technology, Keio University, Japan), and Andreas Dräger (Data Analytics and Bioinformatics, Institute of Computer Science Martin Luther University Halle-Wittenberg, German Center for Infection Research (DZIF), Quantitative Biology Center (QBiC), Eberhard Karl University of Tübingen, Germany)

Numerical calculations are at the heart of systems biology. Such computations require interpreting models in a specific framework and performing several preprocessing steps to pass the model to a specialized solver. The Systems Biology Simulation Core Library (SBSCL) is a software library that simulates and analyzes diverse systems biology models, including flux balance constraints, stochastic simulation, and ordinary differential equation systems. It parses SBML models (the Systems Biology Markup Language), a common language used to describe biological processes, and SED-ML files to conduct more involved analyses. The library supports various algorithms for deterministic and stochastic simulations, allowing precise and efficient simulation of even complex biological and biochemical processes. Furthermore, the library supports integration with other tools and frameworks. SBSCL has been well-tested and benchmarked against the entire SBML Test Suite. It supports several extension packages and provides a highly efficient solver package for SBML models that can be incorporated into any program that runs on the Java Virtual Machine (JVM). SBSCL is available free of charge, even for commercial purposes, at `https://github.com/draeger-lab/SBSCL`.

4.8 BayModTS: A Bayesian workflow to process variable and sparse time series data.

Sebastian Höpfl and Nicole Radde, University of Stuttgart, Germany

Biomedical data generation is limited due to cost and ethical aspects. This leads to sparse time series with only a few replicates available. The analysis of this data is further complicated by the inter-individual variability of organisms and the variability within one organism over time. In this context, analyses that consider only the means and ignore the data variability fail due to low signal-to-noise ratios.

Bayesian Modeling of Time-Series Data (BayModTS) processes the data and takes the uncertainty from highly variable time-series data into account. It employs the retarded transient functions of C. Kreutz as a universal simulation model and can be easily adapted to user-specified SBML models. Using an appropriate noise model, a parameter posterior distribution is inferred via Markow-Chain-Monte-Carlo sampling. Posterior predictive distributions transfer parameter samples from the posterior to model predictions, providing continuous predictions with filtered noise. We demonstrate BayModTS' feasibility on rats' in vivo liver perfusion after 60% Portal Vein Ligation. BayModTS acts as a noise filter and transforms MRI perfusion measurements into time-continuous predictions about the perfusion of individual liver lobes equipped with credibility tubes. These can be used as input for liver function models.

In summary, BayModTS is a Findable, Accessible, Interoperable, and Reusable (FAIR) Bayesian workflow to analyse variable and sparse time series data. A user-friendly toolbox can be found on GitHub.

4.9 The role of standards in defining an ecosystem for Virtual Human Twins (VHTs)

Gerhard Mayer, Martin Golebiewski, and Wolfgang Müller, HITS, Heidelberg, Germany

The European EDITH (Ecosystem Digital Twins in Healthcare) project (https://www.edith-csa.eu) funded by the European Commission is paving the road for building a European infrastructure for Virtual Human Twins (VHTs) in healthcare. A Virtual Human Twin (VHT) is data- and/or knowledge-driven multiorgan and multiscale representation of the quantitative human physiology of a single individual or a group of individuals and can be used for complex personalized predictive computational simulations that are applicable for personal health forecasting, for disease and treatment prognosis prediction, for personalized clinical decision support systems to simulate medical treatment options, for the development of personalized medical products, and for the use in biomedical research (e.g. for the data-driven generation of hypotheses in the development of mechanistic models), as well as for many other possible applications in the health domain. Building such an infrastructure for Virtual Human Twins requires interoperability of the manifold data and computational models constructed based on those data, and thus, a high degree of standardization of data and models, as well as applied workflows, modelling approaches and provenance information for traceability. Such standards are defined by initiatives of the scientific community, such as COMBINE, GA4GH, ASME and others, as well as by formal Standard Defining Organizations (SDOs), such as ISO with their technical committees. EDITH develops a proof of concept for a data and model repository and a simulation platform and comprises also ethical, legal, social implications (ELSI) and regulatory compliance aspects, so that in the long run, EDITH will establish a marketplace for digital twins in healthcare.

4.10 Recommendations and requirements for implementing computational models in clinical integrated decision support systems

Martin Golebiewski (HITS gGmbH, Heidelberg, Germany), Pérez Laura López (Universidad Politécnica de Madrid-Life Supporting Technologies Research Group, Madrid, Spain), Elena Martinelli (Unit of Maxillofacial Surgery, Department of Medicine and Surgery, University Hospital of Parma, Parma, Italy), Marc Kirschner (Forschungszentrum Jülich GmbH, Projekt Management Jülich, Jülich, Germany), Sylvia Krobitsch (Forschungszentrum Jülich GmbH, Projekt Management Jülich, Jülich, Germany), Heike Moser (DIN - German Institute for Standardization, Berlin, Germany), Giuseppe Fico (Universidad Politécnica de Madrid-Life Supporting Technologies Research Group, Madrid, Spain), and Tito Poli (Unit of Maxillofacial Surgery, Department of Medicine and Surgery, University Hospital of Parma, Parma, Italy)

Years of progress in biomedical technology have generated a vast number of omics, medical imaging, and health data in multiple formats and described by corresponding metadata in heterogenous ways. Despite its significant promise for clinical use, this big data remains underutilized. The EU-funded EU-STANDS4PM project has established a pan-European expert forum for evaluating existing standards and develop new guidelines for in silico methodologies in personalized medicine. In this context an ad-hoc working group has been created to discuss the practical recommendations and requirements that should be considered for implementing computational models in clinical integrated decision support systems. The outcome of these discussions has resulted in the standard draft ISO/TS 9491-2 "Guidelines for implementing computational models in clinical integrated decision support systems" submitted to and accepted by the ISO committee ISO/TC 276 Biotechnology. Its publication by ISO is anticipated.

This standard draft delivers fundamental requirements for: 1) clinically-driven projects standardization, 2) data handling, 3) assessment of data availability and quality in clinically-driven projects, 4) data modeling and interpretability, 5) validation of existing and development of new models for different populations, 6) uncovering patient-specific and population-related patterns that can improve care, 7) reinforcing a multidisciplinary decision-making process, 8) creating a virtuous cycle of learning, 9) patient involvement and 10) risk management.

We here introduce a guideline for setting up, detailing, annotating, as well as ensuring the interoperability and integration of health data and resulting models, along with their accessibility and origin, in a way that is both understandable and grounded in evidence. It outlines the integration of these guidelines with the conduct of clinical trials through standard operating pro-

cedures. Additionally, it deals with the criteria and advice for the data needed to build or validate these models. These recommendations aim to contribute to the standardization of a framework to regulate the use of data-driven systems for clinical research.

4.11 Shining Light on Single-Cell Dynamics and Heterogeneity: Design and analysis of a hybrid population model for an epigenetic memory system.

Viviane Klingel and Nicole Radde, University of Stuttgart, Germany

Heterogeneity in biological systems can be quantified efficiently by single-cells measurement techniques like flow cytometry. However, many modelling approaches currently cannot capture this behavior, as often only the average cell is covered in commonly used ODE models. Single-cell data can be reduced to summary statistics for use with these models, but this leads to a loss of information in the data and, more importantly, only provides a complete description if the data is close to a normal distribution. Especially in the case of bimodal distributions, which can for example occur in bistable systems, averages are poor descriptions of the data and lack the ability to reproduce important features of the system. The synthetic epigenetic memory system from Graf et al. (2022) is such a particular system. It is characterized by the ability to switch from an OFF- to an ON-state through a transient metabolic trigger. This ON-state is sustained via positive feedback based on DNA methylation. A large part of the cells can remember this state for many days, but eventually, more and more cells switch back to the OFF-state. In the experimental data, this is observable as a transient appearance of two subpopulations, ON- and OFF-cells, with a drift towards the OFF state. We aim to capture this transient bimodality by a tailored model which describes heterogeneous single-cell trajectories. Our hybrid model combines the simulation speed of differential equations with a stochastic process describing cell division, as well as distributed parameters and measurement noise. We trained the model by comparing the simulated population to the data using the Kolmogorov metric, a shape sensitive distance between distributions. The model reproduces the experimental single-cell data as well as bulk methylation measurements well and is able to predict previously unseen data, including experiments of cyclic ON-OFF-switching with an additional input. Our trained model provides insights into the switching behavior and in particular the mechanisms behind the drift towards the OFF-state on the population and on the single-cell scale. Our analysis suggests that the stochastic nature of the cell division plays an important role in the destabilization of the ON-state, but its effect is only observable over long time.

4.12 TFpredict

Dóra Viktória Molnár (German Center for Infection Research (DZIF), Tübingen, Germany, Quantitative Biology Center (QBiC), Eberhard Karl University of Tübingen, Germany), Michael Gaas (German Center for Infection Research (DZIF), Tübingen, Germany Quantitative Biology Center (QBiC), Eberhard Karl University of Tübingen, Germany), Andreas Dräger (German Center for Infection Research (DZIF), Tübingen, Germany, Quantitative Biology Center (QBiC), Eberhard Karl University of Tübingen, Germany, Data Analytics and Bioinformatics, Institute of Computer Science, Martin Luther University Halle-Wittenberg, Germany)

Gene-regulatory and signaling networks have matured into substantial instruments for analyzing biological functions. Constructing such networks highly depends on the availability of detailed information about protein functions, which often needs to be completed, particularly for less well-studied organisms. Transcription factors (TFs) play a crucial role in regulating gene expression and are vital for understanding complex biological networks. TFpredict is a cutting-edge supervised machine-learning tool designed to facilitate network building by accurately predicting TF interactions and regulatory pathways. In combination with SABINE, TFpredict can even predict the nature of interactions and binding domains, providing a comprehensive view of the regulatory mechanisms at play. The methodology involves BLAST score extraction, superclass prediction with different classifier models, and the identification of DNA-binding domains with InterProScan. The application of TFpredict in constructing robust and comprehensive biological networks showcases its potential to revolutionize regulatory network analysis. By automating the prediction process, TFpredict significantly reduces the time and effort required for network construction. Its ability to predict interactions and binding domains offers a detailed understanding of TF dynamics, facilitating the study of complex regulatory pathways. This capability is particularly beneficial for

research on less well-studied organisms, where experimental data may be sparse. TFpredict is a powerful tool in the field of systems biology, enabling researchers to gain deeper insights into TF dynamics and regulatory networks. Its integration into network analysis workflows enhances the accuracy and comprehensiveness of the resulting models, paving the way for discoveries in gene regulation. TFpredict is available as an open-source Java project on GitHub, providing the scientific community access to its functionalities. The tool can be easily integrated into existing bioinformatics pipelines, and comprehensive documentation facilitates its use. Researchers are encouraged to contribute to its ongoing development and application, ensuring its continued evolution and relevance in network modeling.

4.13 A Computational Pipeline for Evaluating Agreement Between Large-Scale Models and Diverse Datasets

Jonah Huggins, Isabel Leal, Benjamin Childs, Marc Birtwistle, Clemson University, USA

Computational models capturing the function of every gene within a cell, known as Whole-Cell Models (WCMs), can predict complex, multi-gene phenotypes while reconciling discrepancies in current understanding (Karr et al., 2012; Macklin et al., 2020). Constructing such models requires the integration of diverse datasets of varying sizes from different labs and assay types. However, aggregating these datasets into a model-readable format to scalably identify model-data mismatch (i.e. knowledge gaps) pose a large challenge for model construction (Szigeti et al., 2018). We are creating a computational pipeline to rapidly evaluate agreement of a large-scale mechanistic model of a human epithelial cell (the SPARCED model (Erdem et al., 2022)) with a compendium of data spanning multiple sources and modalities. Conditions, duration, and results of wet-lab experiments are converted into a machine readable format based in-part on PEtab guidelines (Schmiester et al., 2021). To ensure this pipeline covers a broad range of potential use case scenarios, we constructed 13 benchmarks SPARCED has previously been validated against, comprising various biological conditions, perturbations, and measurement techniques. Initial deployment (i.e. creating new benchmarks) on the LINCS Microenvironment (ME) perturbation dataset (Gross et al., 2022) indicates mixed agreement with Reverse Phase Protein Array (RPPA) data. Further model agreement is being evaluated with RNAseq, ATACseq, and highly multiplexed immunofluorescence perturbation data. This pipeline will provide a means to rapidly evaluate how diverse datasets collectively compare to model variants, thereby improving the accuracy and scalability of SPARCED and contributing to the creation of a human Whole-Cell Model.

4.14 Standard compliant data and model management for systems medicine projects

Olga Krebs, Susan Eckerle, Martin Golebiewski, Anne Elin Heggland, Xiaoming Hu, Maja Rey, Urlike Wittig, and Wolfgang Müller, Heidelberg Institute for Theoretical Studies HITS gGmbH, Germany

Large collaborative projects need to share data during and after, within and beyond the consortium. FAIRDOM-SEEK (`https://fairdomseek.org/`) is an open-source software for storing, cataloguing, sharing and reusing research outcomes designed to support the principles of FAIR (Findable, Accessible, Interoperable, and Reusable) research data management. Originally developed for the needs of systems biology of microorganisms, SEEK is used in numerous projects of systems biology, systems medicine, and related domains. All data types can be handled and the use of files or references to files is possible. Users can change the visibility of files and references, making it a platform for projects and data publication. Its properties make it an interoperability resource for combining different tools for scientific work and subsequent publication of the outcomes. The systems medicine approach to quantification and characterization of large complex systems involves integration of multipledata types (e.g. genomics, proteomics, metabolomics, phenomics, images, patient related data, etc.), stored in several specialized systems used within one project. LiSyM-Cancer for example, uses REDCap (`https://www.project-redcap.org/`) as a clinical data system that manages information about patients and samples; openBIS (`https://openbis.ch/`) as primary system for experimental raw data and its metadata; Nextcloud (`ttps://nextcloud.com/`) for short-term raw data exchange; and OMERO for microscopic images. The harmonisation and integration of (meta)data between these platforms is mandatory to make the data comparable and publishable in open data repositories. Here, we describe our experience in combining multiple open-source data repository systems for the benefit of large collaborative system medicine projects.

4.15 MomaPy: a Python library to work with molecular maps programmatically

Adrien Rougny, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg

Molecular mechanisms of biological systems and pathways may be described and represented graphically under the form of molecular maps. In a molecular map, nodes represent biological entities such as (pools of) bio-molecules and edges relationships between these entities. Molecular maps may be represented and exchanged using standard graphical languages and formats (e.g., SBGN, SBGN-ML, SBML) that are supported by a handful of editors (e.g., SBGN-ED, Newt, CellDesigner) and libraries (e.g., libSBGN, libsSBML, SBMLDiagrams). While these tools allow users to easily build and save maps as images or in standard exchange formats, they are not well suited to work with the content of maps programmatically. Here we introduce MomaPy, a Python library that allows users to perform a wide variety of tasks on maps, including reading, comparing and rendering them efficiently. At its core, MomaPy separates the model of a map (what is represented) from its layout (how it is represented) à la SBML+layout/render, easing the navigation of their biological content. MomaPy currently supports SBGN PD, SBGN AF, and CellDesigner maps, and may be easily extended to support other types of maps and additional tasks to be performed.

4.16 Eulerian Parameter Inference: Modelling of Single-Cell Data

Vincent Wagner and Nicole Radde, University of Stuttgart, Germany

Single-cell measurement techniques and spatial omics call for modelling approaches capable of capturing inherent cell population heterogeneity. In-silico models must be adequately parameterised to reflect the data and enable accurate predictions beyond data reproduction. We here present Eulerian Parameter Inference (EPI), a probabilistic inference method based on a change of variables. The input of EPI is 1) a deterministic simulation model, such as reaction rate equations or other ordinary differential equations, and 2) data exhibiting large variations, for instance, single-cell gene-expression data. EPI translates all information from the data into distributed model parameters. Further, the employed change of variables formulation allows for point-wise evaluation of the inferred parameter density. Each evaluation only requires one forward simulation of the model and its Jacobian. In particular, we do not require an explicit formulation of the inverse mapping from the data to the parameters. We demonstrate EPI's capabilities on diverse models ranging from algebraic equations to ordinary and even partial differential equation systems, thereby proving its practical applicability. The eulerpi Python package is available on the Python Package Index PyPI. It provides all necessary functionalities and only requires a model and a data sample as user input. We hope this easy-to-use package will facilitate EPI's applicability in numerous and diverse research groups.

Figure 1: Participants at this year's COMBINE.

Acknowledgments

We acknowledge the support of the German Research Foundation (DFG) through the international scientific events program (project number 546497022), the Stuttgart Center for Simulation Science (SimTech EXC 2075 - 390740016), and the Stuttgart Research Center Systems Biology (SRCSB).

References

Ashworth, E. T. et al. (2023). "An in silico approach to understanding the interaction between cardiovascular and pulmonary lymphatic dysfunction". In: *American journal of physiology. Heart and circulatory physiology* 324.3, H318–H329. DOI: 10.1152/ajpheart.00591.2022.

Bäuerle, Famke et al. (2023). "Genome-scale metabolic models consistently predict in vitro characteristics of Corynebacterium striatum". In: *Frontiers in bioinformatics* 3, p. 1214074. DOI: 10.3389/fbinf.2023.1214074.

Clark, A. R., K. S. Burrowes, and M. H. Tawhai (2010). "Contribution of serial and parallel microperfusion to spatial variability in pulmonary inter- and intra-acinar blood flow". In: *Journal of applied physiology (Bethesda, Md. : 1985)* 108.5, pp. 1116–1126. DOI: 10.1152/japplphysiol.01177.2009.

Erdem, Cemal et al. (2022). "A scalable, open-source implementation of a large-scale mechanistic model for single cell proliferation and death signaling". In: *Nature communications* 13.1, p. 3555. DOI: 10.1038/s41467-022-31138-1.

Gross, Sean M. et al. (2022). "A multi-omic analysis of MCF10A cells provides a resource for integrative assessment of ligand-mediated molecular and phenotypic responses". In: *Communications biology* 5.1, p. 1066. DOI: 10.1038/s42003-022-03975-9.

Grzegorzewski, Jan, Janosch Brandhorst, and Matthias König (2022). "Physiologically based pharmacokinetic (PBPK) modeling of the role of CYP2D6 polymorphism for metabolic phenotyping with dextromethorphan". In: *Frontiers in pharmacology* 13, p. 1029073. ISSN: 1663-9812. DOI: 10.3389/fphar.2022.1029073.

Grzegorzewski, Jan et al. (2021a). "Pharmacokinetics of Caffeine: A Systematic Analysis of Reported Data for Application in Metabolic Phenotyping and Liver Function Testing". In: *Frontiers in pharmacology* 12, p. 752826. ISSN: 1663-9812. DOI: 10.3389/fphar.2021.752826.

Grzegorzewski, Jan et al. (2021b). "PK-DB: pharmacokinetics database for individualized and stratified computational modeling". In: *Nucleic acids research* 49.D1, pp. D1358–D1364. DOI: 10.1093/nar/gkaa990.

Höpfl, Sebastian et al. (2024). "Bayesian modeling of time series data (BayModTS) - a FAIR workflow to process sparse and highly variable data". In: *Bioinformatics (Oxford, England)*. DOI: 10.1093/bioinformatics/btae312.

Ježek, Filip et al. (2017). "Lumped models of the cardiovascular system of various complexity". In: *Biocybernetics and Biomedical Engineering* 37.4, pp. 666–678. ISSN: 02085216. DOI: 10.1016/j.bbe.2017.08.001.

Karr, Jonathan R. et al. (2012). "A whole-cell computational model predicts phenotype from genotype". In: *Cell* 150.2, pp. 389–401. DOI: 10.1016/j.cell.2012.05.044.

Keating, Sarah M. et al. (2020). "SBML Level 3: an extensible format for the exchange and reuse of biological models". In: *Molecular systems biology* 16.8, e9110. DOI: 10.15252/msb.20199110.

Köller, Adrian, Jan Grzegorzewski, and Matthias König (2021a). "Physiologically Based Modeling of the Effect of Physiological and Anthropometric Variability on Indocyanine Green Based Liver Function Tests". In: *Frontiers in physiology* 12, p. 757293. ISSN: 1664-042X. DOI: 10.3389/fphys.2021.757293.

Köller, Adrian et al. (2021b). "Prediction of Survival After Partial Hepatectomy Using a Physiologically Based Pharmacokinetic Model of Indocyanine Green Liver Function Tests". In: *Frontiers in physiology* 12, p. 730418. ISSN: 1664-042X. DOI: 10.3389/fphys.2021.730418.

Kulhanek, Tomas et al. (2023). "Bodylight.js 2.0 - Web components for FMU simulation, visualisation and animation in standard web browser". In: *Proceedings of the 15th International Modelica Conference 2023, Aachen, October 9-11*. Linköping Electronic Conference Proceedings. Linköping University Electronic Press, pp. 443–452. DOI: 10.3384/ecp204443.

Lauterbach, Simone et al. (2023). "EnzymeML: seamless data flow and modeling of enzymatic data". In: *Nature Methods* 20.3, pp. 400–402. DOI: 10.1038/s41592-022-01763-1.

Macklin, Derek N. et al. (2020). "Simultaneous cross-evaluation of heterogeneous E. coli datasets via mechanistic simulation". In: *Science* 369.6502. ISSN: 0036-8075. DOI: 10.1126/science.aav3751.

Mateják, Marek et al. (2014). "Physiolibrary - Modelica library for Physiology". In: *Proceedings of the 10th International Modelica Conference, March 10-12, 2014, Lund, Sweden*. Linköping Electronic Conference Proceedings. Linköping University Electronic Press, pp. 499–505. DOI: 10.3384/ECP14096499.

Otis, D. R. et al. (1994). "Dynamic surface tension of surfactant TA: experiments and theory". In: *Journal of Applied Physiology* 77.6, pp. 2681–2688. ISSN: 8750-7587. DOI: 10.1152/jappl.1994.77.6.2681.

Pleiss, Jürgen (2021). "Standardized Data, Scalable Documentation, Sustainable Storage – EnzymeML As A Basis For FAIR Data Management In Biocatalysis". In: *ChemCatChem* 13.18, pp. 3909–3913. ISSN: 1867-3880. DOI: 10.1002/cctc.202100822.

Range, Jan et al. (2022). "EnzymeML-a data exchange format for biocatalysis and enzymology". In: *The FEBS journal* 289.19, pp. 5864–5874. DOI: 10.1111/febs.16318.

Schmiester, Leonard et al. (2021). "PEtab-Interoperable specification of parameter estimation problems in systems biology". In: *PLoS computational biology* 17.1, e1008646. DOI: 10.1371/journal.pcbi.1008646.

Szigeti, Balázs et al. (2018). "A blueprint for human whole-cell modeling". In: *Current Opinion in Systems Biology* 7. • Future of systems biology• Genomics and epigenomics, pp. 8–15. ISSN: 2452-3100. DOI: https://doi.org/10.1016/j.coisb.2017.10.005. URL: https://www.sciencedirect.com/science/article/pii/S2452310017301853.

Wagner, Vincent et al. (2024). *Eulerian Parameter Inference: A Probabilistic Change of Variables for Model-Based Inference with High-Variability Data Sets*. DOI: 10.21203/rs.3.rs-4003283/v1.