

# INCOME – integrative, collaborative modeling in systems medicine

## Hackathons for a culture of data and model sharing

by Nina Fischer, Wolfgang Müller, Dagmar Waltemath, Olaf Wolkenhauer and Jan Hasenauer

Systems medicine is an interdisciplinary approach in which physicians and clinical researchers collaborate with experts from the fields of biology, biostatistics, informatics and mathematics to improve the diagnosis, prevention and treatment of diseases. For this, large-scale, heterogeneous data sets must be collected experimentally and merged using holistic models. In reality, however, most projects are limited to individual cellular processes and develop customized models for these. In the e:Med project INCOME, funded by the German Federal Ministry of Education and Research (BMBF), we were able to promote collaboration and exchange between research groups through targeted networking activities. In numerous INCOME meetings, researchers and developers have jointly created a culture of data and model sharing, developing methods to improve the technical links and reusability of existing simulation models. In this way, the meetings provided a platform for strengthening the community.

Mathematical models are used in systems medicine to study disease-related biochemical processes behind diagnostic, prognostic and therapeutic decisions. These models can be formulated in open standards such as the Systems Biology Markup Language (SBML) and stored in model databases such as BioModels. Standards such as the Simulation Experiment Description Language (SED-ML) aim to unambiguously link data and

models and describe various experimental conditions. The development and establishment of standards and databases have had a huge impact on improving the availability and reusability of models and the reproducibility of research results. Nevertheless, many research groups and even widely used software tools have not yet adapted these standards. Instead, they use their own, less standardized formats, which often do not allow rigorous annotation. Correspondingly, according to recent surveys, only about 50% of published modeling studies are still reproducible (Tiwari *et al.*, 2021).

In addition to the models themselves, the data sets used for parameterization and validation are often difficult to access. They are not published, sufficiently annotated, or clearly linked to the model. The aforementioned factors prevent the reuse and expansion of existing models and data sets, thus compromising the success of research projects.

The severely limited reusability and the time required to develop high-quality models are two reasons why many mechanistic models represent only single signaling pathways and often ignore interactions with others. It is high time to leverage existing infrastructures, familiarize researchers with available resources, and make software developers explicitly aware of available libraries and tools for standard-compliant modeling.



Figure 1: Impressions from two INCOME events. INCOME 2019 in Berlin and INCOME 2021 as a virtual event. (Photos: Jan Hasenauer)

## Main objectives of the INCOME project

The objectives of the INCOME alliance project were to:

1. improve networking between research groups and to promote the use of standards
2. establish an integrated database for models and data
3. facilitate collaborative development of large-scale models

Work toward implementing these goals was carried out at several locations in Germany. INCOME (2017–2021) was coordinated by Jan Hasenauer (Helmholtz Zentrum München and University of Bonn), Wolfgang Müller (HITS Heidelberg), Olaf Wolkenhauer (University of Rostock) and Dagmar Waltemath (University Medicine Greifswald, formerly University of Rostock) also worked on the project.

## Training, networking and community-building

INCOME held a total of five events: three conferences with integrated hackathons and two separate hackathons (see Figure 1).

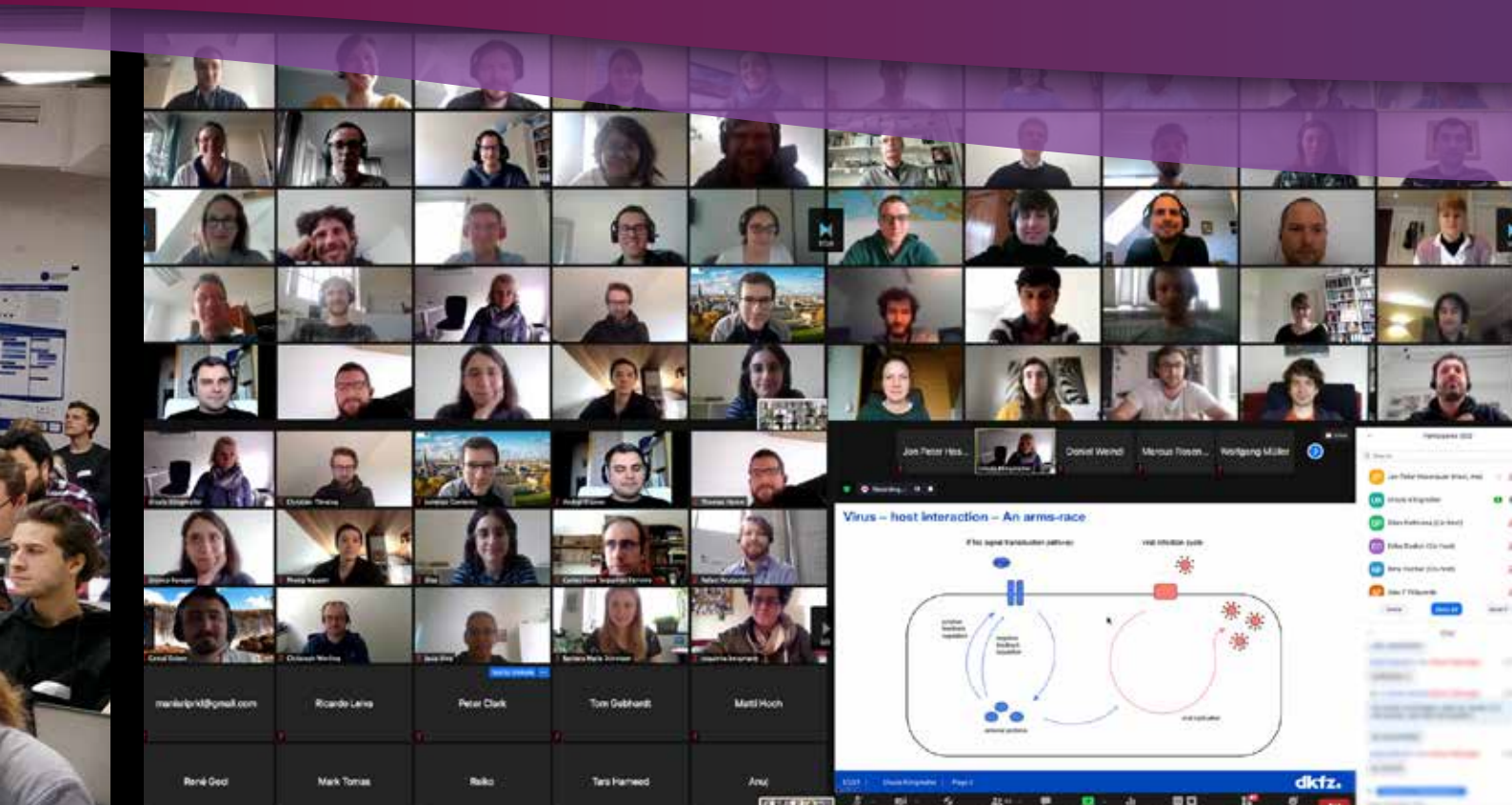
The conferences provided information on recent results in the areas of model building, method development and application. A diverse range of presentations also introduced standards

to a new generation of systems biologists and systems physicians and demonstrated their utility. Panel discussions, as well as group and individual meetings, were used to address open problems and develop solution approaches. The conferences offered extensive networking opportunities, enabling young scientists to exchange ideas with experienced scientists in the community.

Complementing the conferences, the hackathons offered enough space and time for working on joint projects. To significantly simplify the use of standards and improve both reproducibility and reusability, support for standards was provided in many available software tools (e.g., Copasi, Data2Dynamics, Dmod and pyPESTO). In addition, developers took time to demonstrate new software and advise users on any problems. In return, they received direct feedback on their own tools and were able to align further development with user needs.

The events attracted participants from over 20 countries. This internationality and the interdisciplinarity contributed significantly to the development of a strong community at the intersection of modeling, method development and standardization. The increasing number of participants in the events reflects





the growing interest in this area, as well a general interest in collaborating more intensively (see Figure 2). Many aspects of the INCOME events have been directly incorporated into the development of the global standardization network COMBINE ([www.combine.org](http://www.combine.org)).

### Closing gaps

In addition to implementing standards in existing workflows, INCOME also identified and addressed open issues. For example, at the first event in October 2018, it already became clear that there was no suitable solution to formulate parameterization problems. To address this gap, the Parameter Estimation table (PETab) format was developed during the hackathons, but also between meetings (see Figure 3). PETab allows a new level of standardization, improving the reusability of data and models for new studies, such as the development of integrated models.

A large number of research groups have collaborated in the development of PETab, and as a result, the new standard has already been made accessible to thousands of users in eight software tools in a very short time (Schmiester *et al.*, 2021). This would not have been possible without a strong, thematically focused community and the joint hackathons.

### Creating resources

INCOME has helped create community resources for advancing systems biology and systems medicine as research disciplines. For example, a previously established repository for data sets and models (Hass *et al.*, 2019) was transferred to PETab and has been continuously expanded. This not only facilitates reuse, but also enables a realistic benchmarking of methods. A study on the scalability and robustness of existing numerical solution methods has already been completed (Städter *et al.*, 2021), and several more studies on optimization and sample generation methods should be available soon. In addition, as part of a large global community led by Marek Ostaszewski (Luxembourg Centre for Systems Biomedicine), the INCOME consortium has been working on a model for SARS-CoV-2 infections (Ostaszewski *et al.*, 2020). The resulting COVID-19 Disease Map is probably the most comprehensive description of molecular and cellular processes during infection.

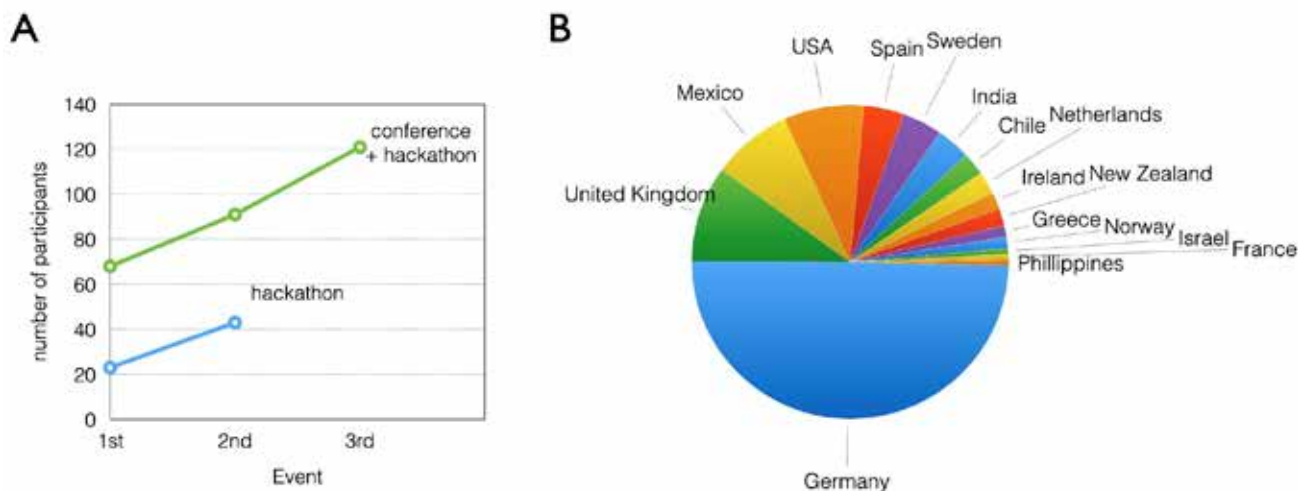


Figure 2: A shows statistics from registration data (number of participants) of the five events held from 2018–2021, while B shows statistics from registration data of the virtual event held in 2021 (Source: Jan Hasenauer).

All resources that have been created are freely available and can be easily used in further projects due to their high degree of standardization. This facilitates knowledge transfer and simplifies a holistic view of diseases.

### But where do we go from here?

INCOME has shown all consortium members how much more can be achieved when they are part of a strong community, and also how important community-building is. The training sessions during the conferences, exchange and collaborative work during the hackathons, and numerous discussions have all helped create a closer connection between model, software and standards development. This has been a process. Individual meetings would not have been sufficient, and there was substantial added value gained from repeated exchange between returning participants, who in turn promoted the meetings in their research circles. We firmly believe that the knowledge that has been and continues to be carried into individual research institutions through INCOME events has significantly improved and will continue to increase the reusability of models. Our goal is to reach many more

researchers in the future. We therefore plan to continue the event series beyond the end of the project (and are working on the feasibility).

**We thank all the participants from the conferences and hackathons who have contributed to the success of INCOME.**

### References:

- Tiwari *et al.* (2021). Reproducibility in systems biology modeling. *Molecular Systems Biology*, 17:e9982.
- Schmiester *et al.* (2021). PETab - Interoperable specification of parameter estimation problems in systems biology. *PLOS Computational Biology*, 17(1):e1008646.
- Hass *et al.* (2019). Benchmark problems for dynamic modeling of intracellular processes. *Bioinformatics*, 35(17):3073-3082.
- Städter *et al.* (2021). Benchmarking of numerical integration methods for ODE models of biological systems. *Scientific Reports*, 11:2696.

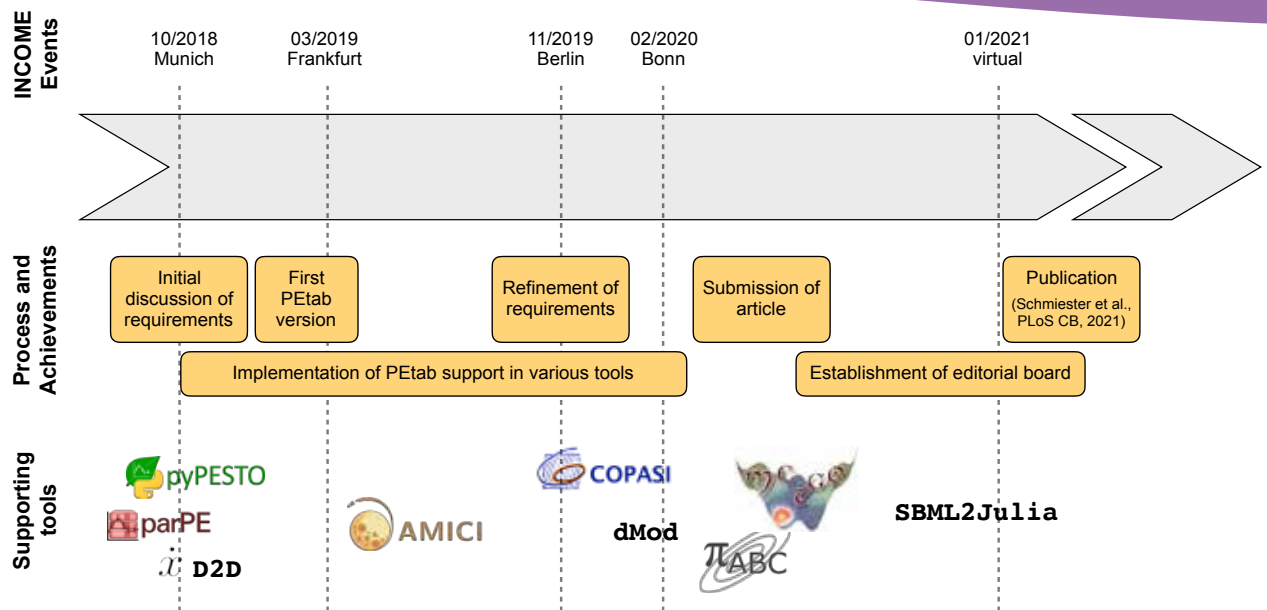


Figure 3: Development of the PETab standard within the framework of the INCOME conferences (Source: Jan Hasenauer).

Ostaszewski *et al.* (2020). COVID-19 Disease Map, building a computational repository of SARS-CoV-2 virus-host interaction mechanisms. *Scientific Data*, 1:136.

**Contact:**



**Prof. Dr. Jan Hasenauer**  
 Life and Medical Sciences (LIMES) Institute  
 University of Bonn  
 jan.hasenauer@uni-bonn.de

[www.mathematics-and-life-sciences.uni-bonn.de](http://www.mathematics-and-life-sciences.uni-bonn.de)



**E:MED NETWORKING FUNDS**

The **projects** supported by the **e:Med Networking Funds** show how valuable resources like software tools, repositories and standards can be developed in hackathons, workshops and summer schools. One success of these projects is specific translational applications in clinical trials on diseases like gastric cancer, breast cancer and inflammatory bowel disorders. Six research alliances link together several e:Med projects in a targeted manner. The scientists contribute their expertise on relevant cross-cutting topics, thereby generating crucial added value through their interdisciplinary collaboration on clinical research, basic research, bioinformatics and modeling. This initiative has been made possible by the BMBF as part of its e:Med funding scheme through resources from the Networking Funds.

[www.sys-med.de/en](http://www.sys-med.de/en)