

Specifications of Standards in Systems and Synthetic Biology

Falk Schreiber^{1*}, Gary D. Bader², Martin Golebiewski³, Michael Hucka⁴, Benjamin Kormeier⁵, Nicolas Le Novère⁶, Chris Myers⁷, David Nickerson⁸, Björn Sommer⁹, Dagmar Waltemath¹⁰ and Stephan Weise¹¹

¹Faculty of IT, Monash University, Clayton, Australia & Martin Luther University Halle-Wittenberg, Germany

²The Donnelly Centre, University of Toronto, Canada

³Heidelberg Institute for Theoretical Studies (HITS), Germany

⁴California Institute of Technology, USA

⁵University of Bielefeld, Germany

⁶Babraham Institute, UK

⁷University of Utah, USA

⁸Auckland Bioengineering Institute, University of Auckland, New Zealand

⁹Faculty of IT, Monash University, Clayton, Australia

¹⁰University of Rostock, Germany

¹¹Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

Standards shape our everyday life. From nuts and bolts to electronic devices and technological processes, standardised products and processes are all around us. Standards have technological and economic benefits, such as making information exchange, production, and services more efficient. However, novel, innovative areas often either lack proper standards, or documents about standards in these areas are not available from a centralised platform or formal body (such as the International Standardisation Organisation).

Systems and synthetic biology is a relatively novel area, and it is only in the last decade that the standardisation of data, information, and models related to systems and synthetic biology has become a community-wide effort. Several open standards have been established and are

*To whom correspondence should be addressed. Email: falk.schreiber@monash.edu

under continuous development as a community initiative. COMBINE, the ‘**C**OMputational **M**odeling in **B**IOlogy’ **N**ETwork [1] has been established as an umbrella initiative to coordinate and promote the development of the various community standards and formats for computational models. There are yearly two meeting, HARMONY (Hackathons on Resources for Modeling in Biology), Hackathon-type meetings with a focus on development of the support for standards, and COMBINE forums, workshop-style events with oral presentations, discussion, poster, and breakout sessions for further developing the standards. For more information see <http://co.combine.org/>.

So far the different standards were published and made accessible through the standards’ web-pages or preprint services. The aim of this special issue is to provide a single, easily accessible and citable platform for the publication of standards in systems and synthetic biology. This special issue is intended to serve as a central access point to standards and related initiatives in systems and synthetic biology, it will be published annually to provide an opportunity for standard development groups to communicate updated specifications.

COMBINE standards covered in this special issue are:

- **CellML** [2] to store and exchange computer-based mathematical models in a modular and reusable manner. In addition, the specification for the **CellML Metadata Framework** is provided.
- **COMBINE Archive Specification** [3] to support the exchange of information necessary for a modeling and simulation experiment in biology. It is a zip-compressed container that includes a manifest file, an optional metadata file, and the files describing the model.
- **SED-ML** [4] (the Simulation Experiment Description Markup Language) to describe the procedures to analyse and simulate models, including model identification, pre-processing, simulation setup, post-processing of simulation result and presentation thereof.
- **SBGN** [5] (the Systems Biology Graphical Notation) to graphically represent processes and networks studied in systems biology. The specifications of the three SBGN languages **SBGN Process Description**, **SBGN Entity Relationship**, and **SBGN Activity Flow** are provided, and these languages allow the representation of different aspects of biological systems at different levels of detail as graphical maps.
- **SBML** [6] (Systems Biology Markup Language) to represent and exchange computational models in systems biology such as models of metabolism, signal transduction and gene regulation. In addition to the specifications for SBML Level 2 and the core of Level 3, this issue includes the specifications for the following extensions to Level 3 (known as “packages” in SBML): **Flux Balance Constraints**, **Hierarchical Model Composition**, **Qualitative Models**, and **Layout**.¹
- **SBOL** [7] (Synthetic Biology Open Language) to exchange data about synthetic biology designs including both structural information, such as hierarchically annotated DNA, RNA, and protein sequences for design components, and behavioral information, such as the interactions between these components.

¹SBML “packages” are extensions to SBML Level 3 only.

No doubt, standards are necessary to ensure exchange, interpretation and reproducibility of scientific results. Access to their definitions is equally important. We hope that this special issue will help to increase the adoption and use of standards in systems and synthetic biology, and thereby support the exchange, distribution, and archiving of models.

GB, MG, MH, NLN, CM, DN, FS and DW are COMBINE coordinators; BS, BK, SW and FS compiled the special issue.

Contact the COMBINE coordinators under combine-coord@googlegroups.com

References

- [1] M. Hucka, D. P. Nickerson, G. D. Bader et al. Promoting coordinated development of community-based information standards for modeling in biology: the COMBINE initiative. *Frontiers in Bioengineering and Biotechnology*, 3:19, 2015.
- [2] A. A. Cuellar, C. M. Lloyd, P. F. Nielsen et al. An overview of CellML 1.1, a biological model description language. *Simulation*, 79(12):740–747, 2003.
- [3] F. T. Bergmann, R. Adams, S. Moodie et al. COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. *BMC Bioinformatics*, 15(1):369, 2014.
- [4] D. Waltemath, R. Adams, F. T. Bergmann et al. Reproducible computational biology experiments with SED-ML – the Simulation Experiment Description Markup Language. *BMC Systems Biology*, 5(1):198, 2011.
- [5] N. Le Novère, M. Hucka, H. Mi et al. The Systems Biology Graphical Notation. *Nature Biotechnology*, 27(8):735–741, 2009.
- [6] M. Hucka, A. Finney, H. M. Sauro et al. The Systems Biology Markup Language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, 2003.
- [7] M. Galdzicki, K. P. Clancy, E. Oberortner et al. The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nature Biotechnology*, 32(6):545–550, 2014.