How Modeling Standards, Software, and Initiatives Support Reproducibility in Systems Biology and Systems Medicine

Dagmar Waltemath* and Olaf Wolkenhauer

Abstract*—*Objective: **Only reproducible results are of significance to science. The lack of suitable standards and appropriate support of standards in software tools has led to numerous publications with irreproducible results. Our objectives are to identify the key challenges of reproducible research and to highlight existing solutions.** ***Results:*** **In this paper, we summarize problems concerning reproducibility in systems biology and systems medicine. We focus on initiatives, standards, and software tools that aim to improve the reproducibility of simulation studies.** ***Conclusions:*** **The long-term success of systems biology and systems medicine depends on trustworthy models and simulations. This requires openness to ensure reusability and transparency to enable reproducibility of results in these fields.**

***Index Terms*—Reproducibility, standards, systems biology, systems medicine.**

I. Introduction

MOST scientific discoveries build upon previous or other findings. The discovery process therefore relies heavily on the reproducibility of scientific results. A lack of transparency and openness led to what many consider a reproducibility crisis [1]–[3].

The failure to reproduce a scientific result has been repeatedly reported over the last years [4]–[13] and led to open discussions of reproducibility issues [14], [15]. In 2011, *Bayer* found that only 43 of 67 findings in cancer studies could be replicated [8], leading to discussions on the reliability of scientific results [9], [15]. In a similar effort, researchers at Amgen could not reproduce 47 of 53 landmark oncological findings for potential drug targets [10]. This study was complemented two years later with tests for reproducibility on 50 other cancer studies [12]. In another investigation, Garcia *et al.* [4] showed that the correctness of *p*-values stated in reports published in 2001 did not correspond to the given test statistics in 38% of articles published in the journal *Nature* and 25% of articles published in the *British Medical Journal*. Further examples of irreproducible studies include microarray gene expression analyses [5], mass spectrometry-based proteomics [6], trials in medical research [7], and many more. The discussions surrounding these cases carry the risk that the public loses trust in biomedical research.

There are four root causes for irreproducibility. First, a lack of standards for data generation leads to problems with the comparability and integration of data sets. Second, a lack of quality and quantity of data often reduces the significance of findings. Third, a lack of openness limits the availability of data and models, as does missing long-term access of resources. Fourth, a lack of transparency arises from missing information on the methods, tools, workflows, and parameter values used in the interpretation of data and models.

In 2013, Garijo *et al.* [11] set out to reproduce a result obtained from Drugome, a method used in tuberculosis research. They spent a total of 280 working hours to fully reproduce the result, including time to analyse the publication, to explore additional materials (data, scripts, and configuration files), to locate and prepare the code, to find appropriate parameter settings, to implement the workflows, to ask questions to the authors, and to validate the workflows. Similarly, Topalidou *et al.* report in detail how they spent three months to reuse a computational model [13]. These examples illustrate how irreproducibility hinders researchers and the scientific community by wasting time and money. Despite the fact that every scientist should have an interest in providing profound, reproducible results, this does not always seem to be the case. Contrarily, the desire to support one's own theory, to be the first person to report a finding, or even the pressure to publish a result may distort—consciously or unconsciously—what we draw from research results [16], [17]. In order to assess the progress and acceptance of scientific findings, Mobley *et al.* [14] asked scientists whether they ever tried to reproduce a finding from a scientific paper. About 54.6% answered that they tried but were not able to do so. Of those, 78.0% contacted the authors of the finding. In 33.3% of the cases the differences were then resolved. About 33.3% published disagreeing results. However, 43.8% of the scientists reported that they faced difficulties in publishing the contradictory results. These disappointing results urge us to reconsider the value of openness and transparency for individuals and the scientific community at large [15], [18].

Manuscript received October 18, 2015; revised February 15, 2016 and March 15, 2016; accepted April 5, 2016. Date of publication June 2, 2016; date of current version September 16, 2016. The work of D. Waltemath was supported by the BMBF e:Bio program under Grant 0316194. The work of O. Wolkenhauer was supported by the BMBF project SysMet under Grant 0316171 and the DFG funding for the project Models, mechanisms and complexity under Grant WO 991/10-1. This work was supported by CaSyM, the EC FP7 coordinating action Coordinating Systems Medicine across Europe. *Asterisk indicates corresponding author.*

*D. Waltemath is with the Institute of Computer Science, University of Rostock, Rostock 18051, Germany (e-mail: dagmar.waltemath@uni-rostock.de).

O. Wolkenhauer is with the Stellenbosch Institute for Advanced Study, Wallenberg Research Centre, Stellenbosch University.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2016.2555481

In this paper, we focus on reproducibility issues related to computational modeling in systems biology and medicine. We first survey challenges in Section II and then discuss solutions in Sections III– V. As most work relies on previous works, science benefits from reusability saving time and money for experimentation and modeling. Furthermore, sharing results openly and making the methods transparent will broaden the impact of the work.

II. Reproducibility Challenge

In systems biology and systems medicine, a lack of suitable standards, limited coverage of standards through software tools [19], use of proprietary or unavailable software [11], and missing information on software configurations [11] can be summarized as a lack of infrastructure in support of reproducibility. Furthermore, a lack of openness to share data; the unavailability of complete models, including all parameter values and all equations; and the unavailability of integrated data repositories impede access to knowledge [20]. A lack of provenance in the management of data is yet another source of concern. Sadly, there is also a lack of recognition for the extra effort required to support the reproducibility of scientific results. Scientists are often not given the time and resources to fully document their results. Neither are there incentives to publish negative results [21]. Also, there are not any incentives to share the original data sets, instead of sharing only the interpretation of data [22], [23].

In systems biology, the reproducibility of results depends on two aspects [24]. First, materials, methods, and standard operating procedures used to generate experimental data should be provided. Second, detailed information on the software, algorithms, and parameters used to analyse and simulate the model should be given. In systems medicine, the reproducibility challenge is becoming even greater, because the diversity of data is immense. The types of data used for decisions related to prognosis, diagnosis, and therapy range from molecular concentrations to sequence information to conventional clinical parameters. Furthermore, access to primary data, software, and documentation is often critical [25].

The challenge is thus to provide researchers with resources and standards to make data and models available over a long period of time. Luckily, the awareness of these problems increases and has led to initiatives that focus on community infrastructure, data availability, and model management. Efforts to improve reproducibility are gradually recognized.

III. Infrastructures

The attention given to the irreproducibility of scientific results has led to the launch of various initiatives tackling the problems from different angles. The *Reproducibility Initiative* [26] offers scientists a blind and independent validation of published findings. Experts from the Science Exchange network of more than 1000 providers at core facilities and contract research organizations try to replicate scientific results [26]. The initiative is a collaboration between Science Exchange, PLOS ONE, Figshare, and Mendeley. *Publishers Weekly* referred to it as the "Dropbox for scientific work." Figshare (http://figshare.com/) is a platform for the management of research data in the cloud. It provides researchers with full control of what data to share, with whom, and when to make it publicly available and citeable. With such platforms, academic institutions get support in sharing and managing their research results, duplication of work is avoided, and thereby time and effort saved. On Figshare, for example, researchers not only share data sets, but also presentations, papers, posters, or software code. Contrary to the *Reproducibility Initiative*, Figshare relies on the feedback of its users to detect nonreproducible results. Observations can directly be reported through a website. The use of information on such servers and the protection of data are a matter of concern, if the services are provided by private companies. Free services, such as Dropbox and Figshare, should therefore be available from public services, dedicated to science.

The following sections focus on networks, standards, and software for enhanced reproducibility of simulation studies. Systems biology has a long tradition of community-driven standard development. Since the field emerged, interoperability of software tools and reuse of modeling data were on the agenda. It is widely accepted that the definition of standards and guidelines maximizes a model's value and impact [27].

A. Community Networks and Standards

Since 2009, the *Computational Modeling in Biology Network* (COMBINE, http://co.mbine.org) oversees the development of open standards for modeling in computational biology. COMBINE fosters communication between the various standardization efforts. It thereby helps to coordinate common activities and to establish a common infrastructure [28]. COMBINE supports both, mature standards and emerging efforts, in covering the current needs in the interoperability landscape [29]. Finally, the network identifies missing standards and initiates further developments. One example of a COMBINE-related activity is the 2015 Whole Cell Summer School, where COMBINE experts advised the participants on the use of COMBINE standards and open software [19].

Currently, COMBINE covers standards for models (CellML [30], SBML [31]), synthetic designs (SBOL [32]), and pathways (BioPAX [33]); for simulation descriptions (SED-ML [34]); and for graphical representations of biological knowledge (SBGN [35]), cf., Fig. 1. The standards are developed by the community, freely available, and implemented in open software [36]. Besides the core COMBINE standards, a number of *associated* efforts add a layer of semantics that facilitates the use, interoperability and enhancement of COMBINE standards. One example is the Systems Biology Ontology (SBO [37]). A third type of COMBINE standards are *related* standards. They are mostly candidate standards, or similar efforts in related domains of research. One example is NeuroML [38], a language to represent neuronal models. Another example is PharmML [39], which is used in pharmacometrics to encode models, associated tasks, and their annotations. PharmlML is also a project of the European *Innovative Medicines Initiative* (IMI, http://www.imi.europa.eu/). IMI brings together industrial and academic partners to build efficient networks for pharmaceutical research.

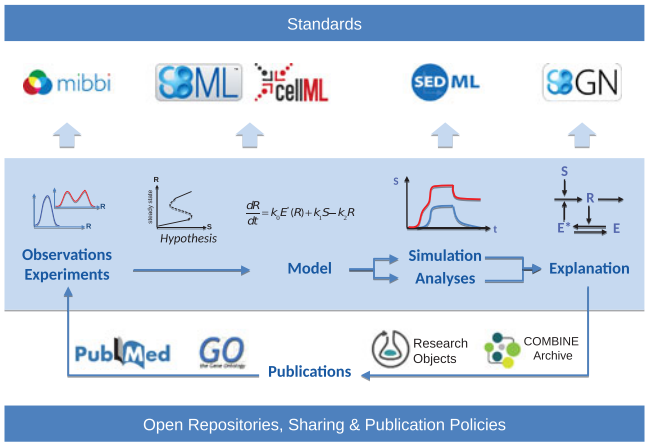

Fig. 1. *One sample workflow in modeling and simulation.* In this example, a computational experiment leads to observations, which in turn lead to hypotheses, which are then encoded by a model. The simulation can then be used to explain the phenomena, but may also lead to new predictions. These predictions help determining a network that captures a mechanism explaining the system's behavior. To make this work reproducible, (1) all steps of this workflow must be documented and enriched with provenance information; (2) the data must be generated following standardized operating procedures and be stored in open repositories; (3) the entities of the model must be stored in standard formats and enriched with semantic information; (4) and the workflow itself must be made transparent. The figure provides examples of different standardization initiatives in this workflow.

The *FAIRDOM* project is a European effort to establish data and model management service facilities for systems biology (http://fair-dom.org). The initiative works together with stakeholders towards FAIRer management, that is to make data, operations, and models better Findable, Accessible, Interoperable, and thereby Reusable. A FAIRDOM Commons is an instance of the SEEK Systems Biology asset platform (http://www.seek4science.org). It comes in two flavours: independently managed Commons installations and the FAIRDOM Public Commons FAIRDOM Hub (http://www.fairdomhub.org). The FAIRDOM project also offers curation, training, and data management plans for European research projects; runs workshops and summer schools to expand the expertise and knowledge about management strategies within the systems biology community; and contributes to public policy and standards setting for research asset management.

B. Software

Access to a study's raw data is one component of reproducible research, but another equally critical component is access to code and software [40], [41]. Indeed, one can observe a trend towards the exchange of the full computational setup, together with the experimental results, in order to ensure reproducibility. Platforms such as *Docker* (https://www.docker.com/) allow developers to distribute complete applications. Users can build, execute, and share software containers for Linux-based operating systems [42]. Docker images help to avoid dependency problems in software, imprecise code documentation, barriers to adoption of existing solutions, or problems due to bug fixes in the software [43]. Already in 2012, the ENCODE consortium showcased how resources used in integrative analyses of the human ENCODE data can be distributed through a virtual machine to help share data, tools, and pipelines in a reproducible manner [44]. Recently, standardized workflow descriptions gained popularity as a mean to exchange repeatable analyses. For example, Galaxy [45] and Knime [46] allow for the reuse of scientific workflows, enabling consistent and repeated analysis of data. Specifically for modeling in systems biology, we observe that simulation software is accessible online. For example, Sycamore [47] is an online collection of tools and methods to build models of biochemical systems, to view, analyse, refine, and simulate them. Another example is JWS Online [48], a model database integrated with a simulation and analysis environment. In addition, more specific, smaller online services offer annotation and clustering of models (SemanticSBML [49]); conversion between formats (Systems Biology Format Converter [50]); retrieval of models and associated files (M2CAT [51]); or export of simulation descriptions and archives (http://sysbioapps.dyndns.org/SED-ML%20Web%20Tools). If tasks cannot be handled online, say, they are too complex, client-side tools such as COPASI [52] and toolboxes such as the Systems Biology Workbench [53] offer export of simulation studies in a standard-compliant format. In their latest survey, the SBML team reported more than 200 supporting software tools [36].

Reproducibility can also be enhanced if existing data in separate repositories are integrated, thereby avoiding data silos. For example, *I2B2* (https://www.i2b2.org/) develops a framework for integrating biology and the bedside [54]. The assumption is that research will benefit from existing medical record data that can be queried across multiple patients. I2B2 focuses on the development of methods for data integration, and it develops into a valuable pool of data. The initiative is funded by the US National Institute of Health.

IV. Model Management

Quantitative models will be only useful when access and reuse are easy [55]. Standards make model-related data interoperable, but documentation, accessibility, curation, and openness are equally important. To this end, data management strategies are required, and methods and tools to access models and related data must be developed [27], [55], [56].

A. Repositories and Management Systems

Open model repositories provide curated and reusable model code in standard formats. They are therefore crucial tools to publicise models [57]. BioModels [58] and JWS Online [48] are two repositories that provide model code in an SBML format. These repositories contribute to the reproducibility of scientific results, because they provide ready-to-reuse model code. They also offer useful online services. BioModels, for example, features the conversion of SBML models between the different available levels and versions [59]. Consequently, the models are in standard formats and can be read by SBML-compliant software tools, e.g., for analysis, parametrization, visualization, exploration, etc. Furthermore, links point to the original publication and eventually to online simulation facilities [48],

[60]. Thereby, the model code is directly accessible from the publication and can be referenced through a unique identifier [61]. Other examples of open repositories include the ModelDB [62] and Opensourcebrain [63] for neuroscience models, or the Physiome Model Repository for models encoded in CellML (PMR2 [64]). Recent efforts aim to standardize more model-related data. The need to search and retrieve models, to enable version control, and to document computational experiments [56] gave birth to the term "model management" [65]. It fostered the development of tailor-made solutions for classical management tasks [49], [56], [57], [66] and model management systems. The SEEK platform, for example, offers rich data and model management facilities for systems biology and systems medicine projects [67].

B. Curation

If the validity of a model has been guaranteed by an independent party, then the model can be trusted and thus reused. In an ideal scenario, repositories offer services for model curation. Curators verify that the submitted model code complies with what was reported in the original publication. This procedure consists of several sequential steps and leads to the publication of the model [58]. If reproduction fails, however, the authors are contacted and the model code is fixed. Curated models in BioModels contain detailed annotations to bio-ontologies. These annotations improve model retrieval, e.g., through a Gene Ontology-based browser. Models in PMR2 are marked with a "star"-system indicating a model's level of curation [64]. A current bottleneck is the effort required to manually curate the models. This could be reduced through thorough documentation and reporting.

C. Documentation and Reporting

Even reproducible analyses can suffer from problems, including omitted variables, or missing data [68]. Improved standards in peer review, reporting and dissemination of research, and training of the scientific workforce can help [86]. A detailed lab notebook is a key for good record-keeping [69] and eases the later documentation of a scientific study. The lab notebook, whether electronic or not, should contain records for all results that were produced in a study, including version control for models, simulation experiments, and computer code in general. Furthermore, best practices in documenting published research results should be followed. For example, the *Reproducibility Initiative* advices researchers to first conduct a direct replication (using the same materials and methods, including any additional controls as necessary), pre-register protocols, use positive and negative controls , etc., [70]. Reproducible computational research in biology could, for example, follow guidelines such as the *Ten Simple Rules for reproducible computational research* by Sandve *et al.* [71].

Reporting Guidelines, also referred to as Minimum Information (MI) guidelines, provide checklists for the publication of data and models in systems biology and systems medicine. When providing all information requested by the corresponding MI, researchers have already contributed to making their results reproducible. The *Minimum Information Requested in the Annotation of Models* (MIRIAM [72]), for example, contains a checklist of information to be provided together with a computational model describing a biological system. The *Minimum Information About a Simulation Experiment* (MIASE [73]) lists all information necessary to describe a reproducible simulation experiment on a MIRIAM-compliant model. Software infrastructure supports these reporting steps, e.g., through tools that bundle all files necessary to reproduce a modeling result [51].

D. Sharing Reproducible Simulation Studies

Even if the raw data are deposited in public archives, the essential analysis intermediaries, scripts, or software are frequently not made available, meaning the science is not reproducible [74]. One approach to transfer research results are *Research Objects* (RO [75]), which allow scientists to group and associate the resources and data used to generate a scientific result. Thereby, ROs guarantee reproducibility of the encapsulated research result, and they increase transparency of the procedure how the result was achieved. ROs are implemented in data management systems such as SEEK and platforms such as Figshare, for example.

Encoding and sharing reproducible simulation experiments in systems biology is the mission of two recent COMBINE efforts: SED-ML and COMBINE Archive. The *Simulation Experiment Description Markup Language* (SED-ML [34]) is a computer-readable format for encoding virtual experiments, predominantly simulations. A simulation experiment is considered reproducible if simulations performed with the same methods on identical models but using different software tools lead to the same results. To this end, SED-ML defines the elements to capture preprocessing, including alternative parametrizations; to describe the simulation procedure; to capture postprocessing of the simulation result; and to encode details about the output displayed to the user. SED-ML Level 1 Version 2 [76] covers most of the experiments typically performed on models in BioModels (cf., Fig. 1 for a systems biology workflow that involves simulation experiments). However, more complex, multistep operations that involve a calibration of the model with experimental data, for example, are not yet possible. If published, the models used in an SED-ML file can be pointed to via perennial identifiers [61]. Sometimes, however, it is useful to ship the complete set of model-related data to a different location, for example in large-scale projects. For this application, the COMBINE Archive [24] was developed. An archive is a single, zip-like file to exchange a model together with its associated simulation experiments and data, graphical representations, result plots, or any other files related to a modeling study.

V. Recognition for Best Practice

How work is conducted in a research project is to a large extent determined by conditions set by funding agencies, requirements set by journals, and constraints from the institutional or industrial environment [77]. While traditional paper publications often present intellectual arguments only, reinforcement of results demands inclusion of data, methods and results in our

publications [78]. The efforts described in the previous sections, specifically documentation and reporting, or the curation process, require additional time and money. We observed ourselves, while writing this survey, that we are not always doing what we ourselves preach: A main problem is that there is little or no recognition for the extra effort that is required. While individuals may not immediately recognize the benefit of the extra effort for themselves, from the perspective of the science community and funding bodies, "good habits of reproducibility" do save time and money in the longer run [71]. There is thus a need either to encourage best practice through the community or to enforce it through requirements by funders and journals.

A. Publishing Reproducible Research

The *Reproducibility Initiative* generated attention for the reproducibility challenges described in Section II. Also the Nature special issue on *Challenges in Irreproducible Research* (http://www.nature.com/nature/focus/reproducibility/) highlighted the importance of avoiding sloppy mistakes when publishing results. The special issue features a collection of articles discussing common reporting principles for journals, code sharing, and pitfalls of working on massive amounts of data. Nature also introduced checklists of reporting requirements and eliminated length limits on methods sections. Together, these efforts encourage new data management and submission guidelines for journals. When submitting a modeling result to the BMC journals, FEBS or PLOS, for example, the authors are asked to provide the model code through an open model repository. Online services support researchers in finding relevant guidelines quickly. BioSharing [79] is a catalogue of standards, minimum information guidelines, and formats for the biosciences. The website allows researchers to browse through the landscape of standards and the systems implementing them. In health research, the EQUATOR initiative serves as an umbrella organisation for all areas of health-research reporting [80]. The initiative aims to improve the reliability in existing reporting guidelines and to make published health literature transparent and more accurate.

The call for open data and reproducible science also resulted in new journals that left the traditional path of publishing. *Nature Scientific Data* (www.nature.com/sdata/) is an open-access journal (and deposit) for scientifically valuable datasets. The descriptions are narrative, as are traditional journal articles. At the same time, the descriptions are enriched with structured descriptions (metadata) of the relevant data. Thus, the data items themselves become discoverable, reusable, citable, and they are open. GigaScience (http://www.gigasciencejournal.com) and F1000Research (http://f1000research.com/) are examples of online, open-access, and open-data journals for life sciences and biomedical research. A publication in GigaScience consists of standard manuscripts linked to an extensive database of all associated data and to the data analysis tools and cloud-computing resources used in the study.

Data sharing and good practice in data management require efficient infrastructure and training of researchers [81], [82]. On the funder's site, the lack of structured information about research outcomes is addressed by new guidelines for long-term data preservation and data management strategies, e.g., the ERASYS-APP guidelines in Europe, or the NIH guidelines in the U.S. This bottom-up approach to data management requires researchers to follow and develop new best practices collaboratively [79]. Already now, stakeholders explicitly fund infrastructure projects that aim at improving data management and thus reproducibility of scientific results. As part of their Big Data initiative, the NIH now explores ways to make data transparent and more accessible, for example through their Data Discovery Index (DDI [83]).

B. Depositing Reproducible Models

Finally, reproducibility is in the interest of the scientists themselves. Recognition can be gained through BioModels' so-called Model of the month. The BioModels team selects one model each month and describes it in full detail online. This not only increases the visibility of the model, but also the visibility of the lab doing the research. Similarly, the SBGN website has a "Symbol of the month" to teach SBGN glyphs. The Cancer Biology project started to give credit to scientists who publish reproducible results [12].

The thorough documentation and the provision of all necessary files reduce the effort of curating simulation studies. Penkler *et al*. [84] nicely showed how a modeling study can be described and published in a reproducible manner, using standard formats and the SEEK platform. Their model code is available, together with raw data, simulation setups, and result data. In their publication, all processing steps are fully described. The necessary steps to prepare a model for publication include: encoding a model in standard format, annotating a model with terms from bio-ontologies, encoding the simulation recipes in standard format, and depositing the result data in standard format [27]. All files should then be bundled as ResearchObjects or COMBINE Archives, and the main execution file should be clearly marked [24]. For an example of how to prepare a COMBINE Archive, we recommend the interested reader to have a look at the comprehensible, fully-featured archive of a Cell Cycle model [85].

VI. Discussion

The impression that many published research findings are false or exaggerated is a matter of concern [86]. The retraction of scientific results that were found to be erroneous after publication (http://retractionwatch.com/) receive attention by the media in the general public [87]. While a large number of such cases can damage the reputation of the science community, finding errors is key to scientific progress. Efforts that encourage reproducibility do therefore make an important contribution to science in general.

The *Reproducibility Initiative* and similar efforts, special issues in journals, and data management policies introduced by funders contribute to better reproducible science. Data and model management guidelines request long-term availability of reproducible scientific results, often in open repositories. These repositories store data in standard formats. In systems biology,

the COMBINE initiative develops such formats for the exchange of modeling results. In addition, projects such as FAIRDOM develop management guidelines and infrastructure for collaborative modeling.

Although publication of model code through open repositories has become common practice, simulation descriptions remain a rare event. To encourage the wide-spread use of SED-ML, this standard will be extended to cover more types of simulation experiments. The storage and retrieval of simulation studies will become more effective when there is greater SED-ML support by software tools. This support will therefore improve the reproducibility of simulation results.

To make simulation studies transparent, the provenance of all data related to a simulation study, as well as who, when, where, and why the person generated the data should be tracked. This task requires better software tools and repositories to record a model's provenance, and appropriate ontology terms to encode provenance of models and simulations.

Generally, modelers expect their software tools to manage model code correctly in a standard format. They should furthermore receive better support in archiving model-related files, and the process of sending the model code to an open repository should be automated. The current situation could be improved by funding calls specifically dedicated to the development of model management infrastructure and standards support in software.

The uptake of Linked Data and Semantic Web technologies provide compelling approaches to reuse scientific data [78], [79]. While it has been shown that sharing detailed research data is associated with increased citation rates [88], the question remains open whether this also leads to more collaboration, or greater reuse, as well as better research results.

Recent success stories of reproducibility initiatives can motivate and inspire. For example, the *Reproducibility Project: Cancer Biology* is a collaboration between Science Exchange and the Center for Open Science. It independently replicates a subset of experimental results from 50 high-impact cancer biology studies published between 2010 and 2012 using the Science Exchange network of expert scientific labs (https://osf.io/e81xl/wiki/home/). Furthermore, researchers at the University of Oxford recently launched the *Cardiac Physiology Web Lab*, an online system that provides comparable and reproducible virtual experiments on sets of models [89]. Together with the models, the study results are shown and compared against different models and versions thereof. All data are open access and can be downloaded in COMBINE formats.

VII. Conclusion

Most scientists are aware of the reproducibility challenges and the costs this generates. But time passes very quickly, and documentation is often the last and neglected step before publication. In this situation, we need to teach young scientists to integrate documentation and data management tasks into their daily schedule. This practice will also help to reward reproducible research results with extra funding or other forms of recognition [15]. If then an effort towards reproducibility is made, we should acknowledge the value of negative results and finding errors. If authors make their work open and transparent, they should not be blamed for the errors found, because even then the work can make a valuable contribution to the progress of science.

Reproducibility is of great importance. "Nonreproducible single occurrences are of no significance to science" [see Karl Popper [90]].

References

[1] B. Nosek *et al.*, "Promoting an open research culture," *Science*, vol. 348, no. 6242, pp. 1422–1425, 2015.

[2] J. Lehrer, "The truth wears off," *New Yorker*, vol. 13, 2010, Art. no. 52.

[3] B. Richard, "Reproducibility undergoes scrutiny," *BioScience*, vol. 64, no. 4, 2014, Art. no. 368.

[4] E. García-Berthou and C. Alcaraz, "Incongruence between test statistics and *p* values in medical papers," *BMC Med. Res. Methodology*, vol. 4, no. 1, 2004, Art. no. 13.

[5] J. P. Ioannidis *et al.*, "Repeatability of published microarray gene expression analyses," *Nature Genetics*, vol. 41, no. 2, pp. 149–155, 2009.

[6] A. W. Bell *et al.*, "A HUPO test sample study reveals common problems in mass spectrometry–based proteomics," *Nature Methods*, vol. 6, no. 6, pp. 423–430, 2009.

[7] J. Arrowsmith, "Trial watch: Phase II failures: 2008–2010," *Nature Rev. Drug Discovery*, vol. 10, no. 5, pp. 328–329, 2011.

[8] F. Prinz, "Believe it or not: How much can we rely on published data on potential drug targets?" *Nature Rev. Drug Discovery*, vol. 10, no. 9, pp. 712–712, 2011.

[9] A. Mullard, "Reliability of 'new drug target' claims called into question," *Nature Rev. Drug Discovery*, vol. 10, no. 9, pp. 643–644, 2011.

[10] C. G. Begley and L. M. Ellis, "Drug development: Raise standards for preclinical cancer research," *Nature*, vol. 483, no. 7391, pp. 531–533, 2012.

[11] D. Garijo *et al.*, "Quantifying reproducibility in computational biology: The case of the tuberculosis drugome," *PLoS ONE*, vol. 8, no. 11, 2013, Art. no. e80278.

[12] T. M. Errington *et al.*, "An open investigation of the reproducibility of cancer biology research," *eLife*, vol. 3, 2014, Art. no. e04333.

[13] M. Topalidou *et al.*, "A long journey into reproducible computational neuroscience," *Frontiers Comput. Neurosci.*, vol. 9, 2015, Art. no. 30.

[14] A. Mobley *et al.*, "A survey on data reproducibility in cancer research provides insights into our limited ability to translate findings from the laboratory to the clinic," *PLoS ONE*, vol. 8, no. 5, 2013, Art. no. e63221.

[15] J. P. Ioannidis *et al.*, "Increasing value and reducing waste in research design, conduct, and analysis," *Lancet*, vol. 383, no. 9912, pp. 166–175, 2014.

[16] R. Nuzzo, "How scientists fool themselves-and how they can stop," *Nature*, vol. 526, no. 7572, pp. 182–185, 2015.

[17] R. MacCoun and S. Perlmutter, "Blind analysis: Hide results to seek the truth." *Nature*, vol. 526, no. 7572, pp. 187–189, 2015.

[18] H. A. Piwowar *et al.*, "Data archiving is a good investment," *Nature*, vol. 473, no. 7347, pp. 285–285, 2011.

[19] D. Waltemath *et al.*, "Toward standards and tools for whole-cell modeling," vol. 63, no. 10, pp. 2007–2014, 2016.

[20] B. Mons, "Which gene did you mean?" *BMC Bioinform.*, vol. 6, no. 1, 2005, Art. no. 142.

[21] J. P. Ioannidis, "Why most published research findings are false," *Chance*, vol. 18, no. 4, pp. 40–47, 2005.

[22] Anonymous, "Credit where credit is overdue," *Nature Biotechnol.*, vol. 27, 2009, Art. no. 579.

[23] V. Stodden *et al.*, "Toward reproducible computational research: An empirical analysis of data and code policy adoption by journals," *PLoS ONE*, vol. 8, no. 6, 2013, Art. no. e67111.

[24] F. T. Bergmann *et al.*, "COMBINE archive and OMEX format: One file to share all information to reproduce a modeling project," *BMC Bioinform.*, vol. 15, no. 1, 2014, Art. no. 369.

[25] C. G. Begley and J. P. Ioannidis, "Reproducibility in science improving the standard for basic and preclinical research," *Circulation Res.*, vol. 116, no. 1, pp. 116–126, 2015.

[26] D. Pattinson, "PLOS ONE launches reproducibility initiative," PLOS ONE Community Blog, AUG. 2012, published online at http://blogs.plos.org/everyone/2012/08/14/plos-one-launches-reproducibi lity-initiative/.

[27] D. Waltemath *et al.*, "Reproducibility of model-based results in systems biology," in *System Biology*. Berlin, Germany: Springer, 2013, pp. 301–320.

[28] D. Waltemath *et al.*, "Meeting report from the fourth meeting of the computational modeling in biology network (COMBINE)," *Standards Genomic Sci.*, vol. 9, no. 3, 2014.

[29] M. Hucka *et al.*, "Promoting coordinated development of community-based information standards for modeling in biology: The COMBINE initiative," *Frontiers Bioeng. Biotechnol.*, vol. 3, 2015, Art. no. 19.

[30] A. A. Cuellar *et al.*, "An overview of CellML 1.1, a biological model description language," *Simulation*, vol. 79, no. 12, pp. 740–747, 2003.

[31] M. Hucka *et al.*, "The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models," *Bioinformatics*, vol. 19, no. 4, pp. 524–531, 2003.

[32] M. Galdzicki *et al.*, "The synthetic biology open language (SBOL) provides a community standard for communicating designs in synthetic biology," *Nature Biotechnol.*, vol. 32, no. 6, 2014, Art. no. 545.

[33] E. Demir *et al.*, "The BioPAX community standard for pathway data sharing," *Nature Biotechnol.*, vol. 28, no. 9, pp. 935–942, 2010.

[34] D. Waltemath *et al.*, "Reproducible computational biology experiments with SED-ML—The simulation experiment description markup language," *BMC Syst. Biol.*, vol. 5, no. 1, 2011, Art. no. 198.

[35] N. Le Novère *et al.*, "The systems biology graphical notation," *Nature Biotechnol.*, vol. 27, pp. 735–741, 2009.

[36] M. Hucka *et al.*, "A profile of today's SBML–Compatible software," in *Proc. IEEE e-Science Workshops*, 2011, pp. 143–150.

[37] M. Courtot *et al.*, "Controlled vocabularies and semantics in systems biology," *Molecular Syst. Biol.*, vol. 7, no. 1, 2011, Art. no. 543.

[38] P. Gleeson *et al.*, "NeuroML: A language for describing data driven models of neurons and networks with a high degree of biological detail," *PLoS Comput. Biol.*, vol. 6, no. 6, 2010, Art. no. e1000815.

[39] M. J. Swat *et al.*, "Pharmacometrics markup language (PharmML): Opening new perspectives for model exchange in drug development," *CPT: Pharmacometrics Syst. Pharmacology*, 2015, pp. 316–319.

[40] R. D. Peng, "Reproducible research in computational science," *Science*, vol. 334, no. 6060, 2011, Art. no. 1226.

[41] J. B. Buckheit and D. L. Donoho, "Wavelab and reproducible research," in *Wavelets and Statistics*, vol. 103, Berlin, Germnay: Springer, 1995, pp. 55–81.

[42] S. R. Piccolo *et al.*, "Tools and techniques for computational reproducibility," *bioRxiv*, 2015.

[43] C. Boettiger, "An introduction to Docker for reproducible research," *ACM SIGOPS Oper. Syst. Rev.*, vol. 49, no. 1, pp. 71–79, 2015.

[44] ENCODE Project Consortium, "An integrated encyclopedia of DNA elements in the human genome," *Nature*, vol. 489, no. 7414, pp. 57–74, 2012.

[45] J. Goecks *et al.*, "Galaxy: A comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences," *Genome Biol*, vol. 11, no. 8, 2010, Art. no. R86.

[46] S. O'hagan and D. B. Kell, "Software review: The KNIME workflow environment and its applications in genetic programming and machine learning," *Genetic Program. Evolvable Mach.*, vol. 16, no. 3, pp. 387–391, Sep. 2015. [Online]. Available: http://dx.doi.org/10.1007/s10710-015-9247-3

[47] A. Weidemann *et al.*, "SYCAMORE—A systems biology computational analysis and modeling research environment," *Bioinformatics*, vol. 24, no. 12, pp. 1463–1464, 2008.

[48] B. G. Olivier and J. L. Snoep, "Web-based kinetic modelling using JWS Online," *Bioinformatics*, vol. 20, no. 13, pp. 2143–2144, 2004.

[49] F. Krause *et al.*, "Annotation and merging of SBML models with semanticSBML," *Bioinformatics*, vol. 26, no. 3, pp. 421–422, 2010.

[50] N. Rodriguez *et al.*, "The systems biology format converter," *BMC Bioinform.*, vol. 17, no. 1, 2016, Art. no. 1.

[51] M. Scharm and D. Waltemath, "Extracting reproducible simulation studies from model repositories using the combinearchive toolkit," in *Datenbanksysteme für Business, Technologie und Web*, N. Ritter, A. Henrich, W. Lehner *et al.*, Eds., 2015.

[52] S. Hoops *et al.*, "—A complex pathway simulator," *Bioinformatics*, vol. 22, no. 24, pp. 3067–3074, 2006.

[53] F. T. Bergmann and H. M. Sauro, "SBW—A modular framework for systems biology," in *Proc. 38th Conf. Winter Simul.* , 2006, pp. 1637–1645.

[54] S. N. Murphy *et al.*, "Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2)," *J. Amer. Med. Inform. Assoc.*, vol. 17, no. 2, pp. 124–130, 2010.

[55] N. Le Novère, "Model storage, exchange and integration," *BMC Neurosci.*, vol. 7, 2006, Art. no. S11.

[56] R. Henkel *et al.*, "Combining computational models, semantic annotations and simulation experiments in a graph database," *Database*, vol. 2015, 2015, Art. no. bau130.

[57] R. Henkel *et al.*, "Ranked retrieval of computational biology models," *BMC Bioinform.*, vol. 11, no. 1, 2010, Art. no. 423.

[58] N. Juty *et al.*, "BioModels: Content, features, functionality, and use," *CPT Pharmacometrics Syst. Pharmacology*, vol. 4, no. 2, pp. 1–14, 2015.

[59] C. Li *et al.*, "Biomodels database: An enhanced, curated and annotated resource for published quantitative kinetic models," *BMC Syst. Biol.*, vol. 4, no. 1, 2010, Art. no. 92.

[60] R. Machné *et al.*, "The SBML ODE solver library: A native API for symbolic and fast numerical analysis of reaction networks," *Bioinformatics*, vol. 22, no. 11, pp. 1406–1407, 2006.

[61] N. Juty *et al.*, "Identifiers.org and MIRIAM registry: Community resources to provide persistent identification," *Nucleic Acids Res.*, vol. 40, pp. D580–D586, 2012.

[62] M. L. Hines *et al.*, "ModelDB: A database to support computational neuroscience," *J. Comput. Neuroscie.*, vol. 17, no. 1, pp. 7–11, 2004.

[63] P. Gleeson *et al.*, "The open source brain initiative: Enabling collaborative modelling in computational neuroscience," *BMC Neurosci.*, vol. 13, 2012, Art. no. O7.

[64] T. Yu *et al.*, "The physiome model repository 2," *Bioinformatics*, vol. 27, no. 5, pp. 743–744, Mar. 2011.

[65] J. Hetherington *et al.*, "Addressing the challenges of multiscale model management in systems biology," *Comput. Chemi. Eng.*, vol. 31, no. 8, pp. 962–979, 2007.

[66] M. Scharm *et al.*, "An algorithm to detect and communicate the differences in computational models describing biological systems," *Bioinformatics*, 2015, Art. no. btv484.

[67] K. Wolstencroft *et al.*, "SEEK: A systems biology data and model management platform," *BMC Syst. Biol.*, vol. 9, no. 1, 2015, Art. no. 33.

[68] J. T. Leek and R. D. Peng, "Opinion: Reproducible research can still be wrong: Adopting a prevention approach," *Proc. Nat. Acad. Sci.*, vol. 112, no. 6, pp. 1645–1646, 2015.

[69] S. Schnell, "Ten simple rules for a computational biologists laboratory notebook," *PLoS Comput. Biol.*, vol. 11, no. 9, 2015, Art. no. e1004385.

[70] The Reproducibility Initiative, "Replication best practices," 2015, last accessed 2015-08-03.

[71] G. K. Sandve *et al.*, "Ten simple rules for reproducible computational research," *PLoS Comput. Biol.*, vol. 9, no. 10, 2013, Art. no. e1003285.

[72] N. Le Novère *et al.*, "Minimum information requested in the annotation of biochemical models (MIRIAM)," *Nature Biotechnol.*, vol. 23, no. 12, pp. 1509–1515, 2005.

[73] D. Waltemath *et al.*, "Minimum information about a simulation experiment (MIASE)," *PLoS Comput. Biol.*, vol. 7, no. 4, 2011, Art. no. e10011221.

[74] T. P. Sneddon *et al.*, "GigaDB: Promoting data dissemination and reproducibility," *Database*, vol. 2014, 2014, Art. no. bau018.

[75] S. Bechhofer *et al.*, "Research objects: Towards exchange and reuse of digital knowledge," *Future Web Collaborative Sci.*, http://dx.doi.org/10.1038/npre.2010.4626.1

[76] F. T. Bergmann *et al.*, "Simulation experiment description markup language (sed-ml) level 1 version 2," *J. Integrative Bioinform.*, vol. 12, no. 2, 2015, Art. no. 262.

[77] V. Stodden *et al.*, "Setting the default to reproducible," *Comput. Sci. Res. SIAM News*, vol. 46, pp. 4–6, 2013.

[78] S. Bechhofer *et al.*, "Why linked data is not enough for scientists," *Future Generation Comput. Syst.*, vol. 29, no. 2, pp. 599–611, 2013.

[79] S.-A. Sansone *et al.*, "Toward interoperable bioscience data," *Nature Genetics*, vol. 44, no. 2, pp. 121–126, 2012.

[80] D. G. Altman *et al.*, "Equator: Reporting guidelines for health research," *Open Med.*, vol. 2, no. 2, 2008, Art. no. e49.

[81] P. Lord *et al.*, Large-scale data sharing in the life sciences: Data standards, incentives, barriers and funding models, the joint data standards study (2006) [Online]. Available: http://www. mrc. ac. uk/pdfjdss_final_report. pdf

[82] W. Wruck *et al.*, "Data management strategies for multinational large-scale systems biology projects," *Briefings Bioinform.*, vol. 15, no. 1, pp. 65–78, 2014.

[83] F. S. Collins and L. A. Tabak, "NIH plans to enhance reproducibility," *Nature*, vol. 505, no. 7485, 2014, Art. no. 612.

[84] G. Penkler *et al.*, "Construction and validation of a detailed kinetic model of glycolysis in plasmodium falciparum," *FEBS J.*, vol. 282, no. 8, pp. 1481–1511, 2015.
[85] M. Scharm, "Combine archive show case," University of Rostock, Rostock, Germany, Tech. Rep., 2015. [Online]. Available github, https://github.com/SemsProject/CombineArchiveShowCase
[86] J. P. Ioannidis, "How to make more published research true," *PLoS MEDICINE*, vol. 11, no. 10, 2014, Art. no. e1001747.
[87] P. Ronald, "Lab life: The anatomy of a retraction," Scientific American, OCT. 2013, last accessed online: 2015-08-07 Available: http://blogs.scientificamerican.com/food-matters/lab-life-the-anatomy-of-a-retraction/.
[88] H. A. Piwowar *et al.*, "Sharing detailed research data is associated with increased citation rate," *PLoS ONE*, vol. 2, no. 3, 2007, Art. no. e308.
[89] J. Cooper *et al.*, "The cardiac electrophysiology web lab," *Biophys. J.*, vol. 110, no. 2, pp. 292–300, 2016.
[90] K. Popper, *The Logic of Scientific Discovery*. Evanston, IL, USA: Routledge, 2005.

Dagmar Waltemath received the Ph.D. degree in computer science from the University of Rostock, Rostock, Germany, in 2011.

Since 2012, she has been working as a Junior Research Group Leader in the Department of Systems Biology and Bioinformatics, University of Rostock. Her group focuses on developing methods to reproducibly simulate computational biology models including storing, annotating, searching and retrieving models and model-related data. Her research interests include strategies for the management, provenance, and integration of model-related data in systems biology and medicine.

Dr. Waltemath is also a COMBINE Coordinator and an SBML Editor.

Olaf Wolkenhauer received the first degrees in control engineering from the University of Applied Sciences, Hamburg, Germany, and the University of Portsmouth, Portsmouth, U.K., in 1994. He received the Ph.D. degree from the University of Manchester, Institute of Science & Technology (UMIST), Manchester, U. K., in 1997, for research on the application of possibility theory to data analysis. Following a research lectureship at the Control Systems Centre at UMIST and a research fellowship at the Technical University Delft, the Netherlands, he directed his research toward systems biology. This led to a joint senior lectureship with the Department of Biomolecular Sciences and the Department of Electrical Engineering and Electronics, UMIST.

In 2003, he moved to the University of Rostock, Germany, where he holds the Chair in Systems Biology and Bioinformatics. Since October 2004, he has been an Adjunct Professor in the Department of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, USA. Since 2005, he has been a Fellow of the Stellenbosch Institute for Advanced Study.