



## D8.3 How to organise and run an ISBE modelling service

*Deliverable D8.3*

*Work package 8*

*Modelling infrastructure and expertise*

Contact: Stig W. Omholt <[stig.omholt@ntnu.no](mailto:stig.omholt@ntnu.no)>

Project ref. no.	INFRA-2012-2.2.4:312455
Project title	ISBE - Infrastructure for Systems Biology Europe
Nature of Deliverable	R=Report
Contractual date of delivery	Month 36
Actual date of delivery	Month 36
Deliverable number	D8.3
Deliverable title	A blueprint document outlining a proposal for how to organise and manage a pan-European ISBE modelling service which has broad support in the European biological research community
Dissemination level	PU=Public
Status & version	Final
Number of pages	23
WP relevant to deliverable	WP8
Lead participant	UMB, Stig Omholt
Author(s)	Jon Olav Vik, Stig Omholt, Frans van Nieuwpoort, Jacky Snoep, Martijn J. Moné, Hans Westerhoff, David Nickerson, Nick Juty, Carole Goble, Natalie Stanford, Thomas Höfer, Vítor Martins dos Santos, Henning Hermjakob, Nicolas Le Novère
Project coordinator	Richard Kitney
EC Project Officer	Keji-Alex Adunmo

Dissemination level: PU = Public, RE = Restricted to a group specified by the Consortium (including Commission services), PP = Restricted to other programme participants (including Commission Services), CO= Confidential, only for members of the Consortium (including the Commission Services)

Nature of Deliverable: P= Prototype, R= Report, D=Demonstrator, O = Other.

## Executive summary

ISBE will empower the experimental research community to make modelling part of their lab routine. Here we discuss how to define and structure ISBE modelling services, building on D8.2's review of the purpose and overall design principles for modelling services. We identify the components of ISBE modelling services and propose an organizational structure to achieve uniform coverage of services from a heterogeneous network of expertises. A roadmap concludes the report, proposing initial low-budget services for selected modelling frameworks, then outlining a strategy for growth and scalability of ISBE modelling services.

How to define and structure ISBE modelling services is a many-faceted question. First and foremost, one must identify the *purpose* of each service (what will it achieve?), which also entails specifying its *inputs and outputs*. It can be useful to group *kinds of services* into broader categories, either based on their *function*, or based on their *organizational structure* including staffing and funding.

Modelling services will include automated analyses of published or user-provided models, fitting of parameters to user data, and refinement and adaptation of existing models. However, it also makes sense to include related services such as stewardship and training in the broad term "modelling services".

A coordinating Systems Biology Centre (cSBC) will interconnect national Systems Biology Centres (nSBCs) to make the collective expertise of participating nations easily accessible for all European researchers. Chapter 2.3 concretizes this with illustrations of the variety of modelling services that ISBE is likely to offer, and how to route requests from clients to ISBE tools and expertise. Levels of service may range from fully automatic tools to long-term personal involvement, with access to more demanding services prioritized in cooperation with funders. Likewise, there will be several kinds and levels of documentation and user support.

We advocate a focus on user needs in prioritizing what ISBE is to offer, in specifying services, and in routing expertise to clients. Clear descriptions of the data requirements for each modelling approach helps clients figure out their options and see what additional data they might require to open more options for modelling. Standardized virtual experiments can form an articulated link between modellers and experimentalists, and can be used to streamline the confrontation of models with data, e.g. in parameter estimation and model selection.

As candidates for pilot services to be offered by the interim "ISBE light", we have identified modelling frameworks that span a wide range of biologically interesting questions within a well-defined set of user inputs. These include ordinary differential equations (where the rates-of-change of state variables is a function of the current value of those state variables) and constraint-based models (based on stoichiometry in biochemical reaction networks).

The demand for modelling services is anticipated to increase sharply once ISBE succeeds in transforming the practice of biological research and application. To meet this demand, we identify five key issues of scalability: Recruitment of providers, routing of clients to providers, strategic focusing of effort, training and education of users, and standardization and curation to streamline knowledge management, each of which is detailed in Chapter 5.2.

## Contents

Executive summary.....	3
1 Introduction and overview .....	6
1.1 About this document .....	6
2 Services.....	7
2.1 Purpose of ISBE modelling services .....	7
2.2 Design principles .....	8
2.3 Organizational structure .....	8
2.4 Specification of services.....	11
2.5 Levels of service .....	11
2.6 Help desk and guidance .....	12
2.7 Routing requests.....	13
2.8 Integration with experiment .....	14
2.8.1 Semantically rich interfaces: ontological annotation .....	14
2.8.2 Virtual experiments .....	14
2.8.3 A continuous integration of knowledge.....	15
3 Clients.....	15
4 Providers .....	16
5 Roadmap .....	16
5.1 Low-hanging fruit: First modelling services .....	17
5.2 Preparing for growth: Scalability of modelling services.....	17
6 Empowering biologists: Some current examples.....	18
6.1 VPH-Share .....	18
6.2 OpenCOR.....	19
6.3 JWS Online.....	19
6.4 BioModels .....	19
7 References .....	20
8 Appendix: Case study JWS Online .....	21
8.1 Introduction .....	21
8.2 The JWS database .....	21
8.3 The JWS simulator.....	21
8.4 Model coding and curation .....	21
8.5 Model types .....	21
8.6 Model editing and model creation .....	21
8.7 Teaching and research .....	22
8.8 Data and model management in systems biology projects.....	22
8.9 Data and model linking .....	22
8.10 Web services .....	22
8.11 Sustainability and scalability .....	22



# 1 Introduction and overview

## 1.1 About this document

ISBE Work Package 8, *Modelling infrastructure and expertise*, aims to *provide the scientific and organisational foundation for a rapid and highly coordinated pan-European implementation of modelling as a service, for academia and industry across all major disciplinary fields of biology.*

This document, Deliverable D8.3, marks the fulfilment of Task 8.3, set out as follows: *Reach a pan-European understanding on how to organise and run an ISBE modelling service.* The Description of Work for this task was as follows:

- *Based on the results from Task 8.2, and by extensive guidance from the European modelling community in academia and industry through various mechanisms, propose required web services, levels of intellectual engagement of modellers, connections between modelling groups, how to ensure that modellers receive due intellectual credit for their service activity, how to establish smooth theoretical-experimental research activity based on a virtual systems biology service structure, and mechanisms concerning decisions on who qualifies for receiving support for providing service.*
- *Expose experimental communities to how to organise and run a modelling service and receive their feedback through international meetings, web site interactions and direct interaction with prominent experimental researchers in various disciplines.*
- *Try out blueprint versions of mission-critical service mechanisms.*
- *Produce final document that will serve as roadmap for constructing main proposal and integrating various national funding initiatives for modelling service.*

D8.3 is a continuation of the resource assessment in D8.2, and was developed in dialog with representatives of the model repositories at JWS Online, Biomodels.net, and Cellml.org, who are also central in standardization efforts. The end-user focus was linked to the persona modelling in the ISBE Business Case (November 2014). Existing modelling services were explored, with JWS Online included as a case study in the Appendix. ISBE surveys and interviews provided further guidance on how to tailor services to users' demands.

References to sections of the Business Plan follow the headings and section numbering of version "ISBE BP 24JUN15 AP.docx".

## 2 Services

This chapter discusses how to define and structure ISBE modelling services. This is a many-faceted question. First and foremost, one must identify the *purpose* of each service (what will it achieve?), which also entails specifying its *inputs and outputs*. It can be useful to group *kinds of services* into broader categories, either based on their *function* (Table 1), or based on their *organizational structure* including staffing and funding (discussed below and illustrated in Figure 1).

**Table 1. Various kinds of possible modelling services.**

- Advice on tools and modelling strategies
- Model refinement and adaptation
- Model development from scratch
- Training and education
- Analysis of models under given scenarios:  
sensitivity analysis of phenotype with respect to parameters,  
simulation of behaviour under given scenarios, ...
- Confrontation of models with data:  
parameter estimation, model selection, ...
- Model reduction, i.e. purposeful simplification

In this report, the term "modelling services" has a broad meaning when used alone, and refers to all ISBE services related to modelling. More specific kinds of modelling services are named as such, e.g. "web services", "model development services", or "parameter estimation services".

The term "service" emphasizes a focus on *clients* and their *interfaces* to service providers. Clients provide input in the form of questions, data, model scenarios or representations of models, and receive output from the provided service, for example parameter estimates and associated statistics, simulated time-courses or solutions to models, or refined model representations. Output formats may range from predefined summaries and figures, to customized reports, to detailed output for post-processing by the client (possibly with the assistance of other ISBE services).

### 2.1 Purpose of ISBE modelling services

As discussed in Deliverable D8.2, the purpose of modelling is to integrate data and knowledge into a unified, predictive whole, generalizing insights from the model to a broader class of systems. However, many modelling activities require a large and highly specialized skillset that few biologists have. The main goal of ISBE modelling services is to empower the experimental research community to make modelling part of their lab routine. Thus, there is a particular focus on industry and research fields that have not traditionally identified themselves with the "systems biology" community.

ISBE modelling services will also benefit model developers, by making it easier to explore alternative models or modelling frameworks, and in obtaining data for model building, testing and validation. This is becoming much easier thanks to advances in data discoverability and interoperability through semantic and ontological annotation (see ISBE Business Plan section 3.2.2 and Appendix 7). The generation and stewardship of "systems-biology-ready" (see Deliverable 2.2) data is important for efficient modelling.

ISBE will let biologists become modellers with less frustration, better reproducibility (Peng 2011) and higher throughput by outsourcing the technical tedium to ISBE. The aim of systems biology training and education (See ISBE Business plan section 3.2.6) should be to make biologists *articulate users* of modelling, not necessarily top-notch computational instrument makers.

## 2.2 Design principles

The service catalogue and implementation of ISBE modelling services will reflect the design principles proposed in D8.2:

- **Empowerment.** ISBE modelling services will lower the threshold for biologists to incorporate computational modelling in their scientific work.
- **A range of service levels.** ISBE should cater both to routine modelling needs and more involved model development.
- **Driven by user demands.** Focus first on modelling frameworks that have a large potential user base and can span a large space of questions with a well-defined set of query parameters. For more involved consultancy, ISBE can play a unique role in mediating contact between experimentalists and modellers.
- **Cooperation with existing communities.** Prominent members of COMBINE and of the major model repositories have been involved in discussions towards pilot "ISBE light" services.
- **Interoperability.** Common standards for the exchange and reuse of knowledge are a prerequisite for offering modelling as a service, and ISBE will be part of international initiatives for standardization. For example, the technical committee for biotechnology standards of the International Organization for Standardization (ISO) has recently established a new working group for data processing and integration<sup>1</sup>.

## 2.3 Organizational structure

A coordinating Systems Biology Centre (cSBC) will interconnect national Systems Biology Centres (nSBCs) to make the collective expertise of participating nations easily accessible for all European researchers. As detailed in the ISBE Business Plan, Section 3.3, the cSBC will be headed by an ISBE director and a supervisory board of stakeholders, whereas nSBCs will be financially supported by their national governments. For the central office a membership fee is proposed related to the gross national product of a country.

---

<sup>1</sup> <http://www.h-its.org/en-sdbv/new-standards-in-biotechnology/>

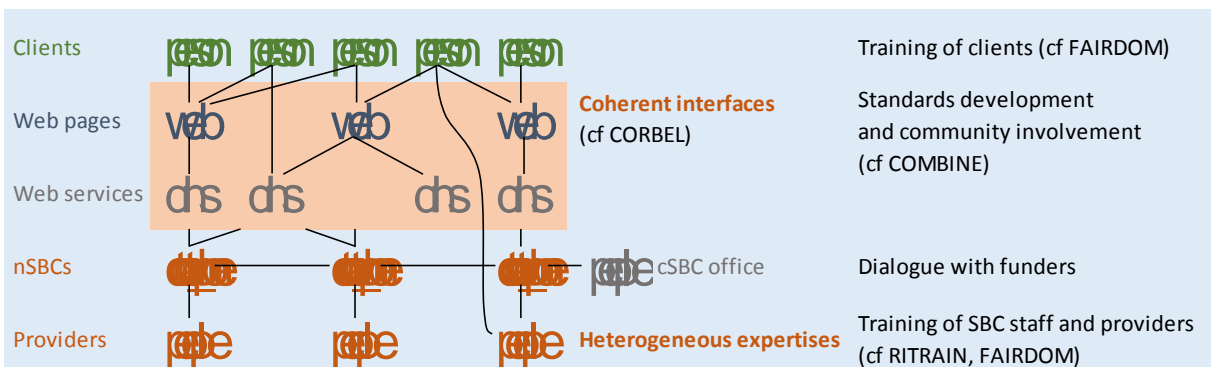


The staffing and funding of nSBCs remains for ISBE members to decide, but a tentative proposal might be:

- A full-time paid coordinator.
- Highly competent engineers/technicians to run a national help desk.
- A pool of ISBE postdocs splitting their time between own research and providing advanced service and training. This combination should make an attractive incentive for talented modellers, offering novel career paths supporting the growth of systems biology in Europe.
- Financial incentives to establish a pool of highly qualified tenured theoreticians, recruited from academic institutions that are willing to contribute with their expertise given that they become co-authors on papers generated.
- Resources to attend conferences and to interact closely with other systems biology centres.
- Resources to participate in large-scale and prestigious European projects

For further discussion of funding and incentives for modelling services, see Deliverable D8.2, Section 5.7.

The services offered by nSBCs will reflect their national priorities, and so service provision will be heterogeneously distributed across countries. However, this heterogeneity will as far as possible be hidden from end users<sup>2</sup>, thanks to the use of common standards, design of coherent interfaces, and coordination from the cSBC so that all nSBCs are aware of what the others have to offer. ISBE and its network of SBCs will work towards efficient cooperation and division of labour in close dialogue with funders, research communities, and standards bodies.



**Figure 1. Concept sketch of how ISBE will offer modelling services across Europe, built from heterogeneous expertise and priorities in participating countries. The diagram illustrates several ways in which ISBE expertise may be available in a manner suited to clients' needs, as described in the text.**

<sup>2</sup> The abstracting of underlying infrastructure from user-facing parts is exemplified by VPH-Share in Section 6.



The sketch in Figure 1 illustrates a few of the many forms of service that ISBE is likely to offer. Many clients will enjoy user-facing, interactive web pages for highly standardized analyses where users can pose a wide range of biological questions just by filling in a form and possibly uploading own data or models, similar to what is already commonplace in bioinformatics (Schattner 2008). More advanced users may automate operations by connecting to the web services underlying the interactive tools, either directly or using wrappers for their particular programming language or tool software. These web services and web pages will be developed by nSBCs with the help of experts in their countries, then made available as part of the coherent ISBE interface.

User guidance and assistance will be offered on several levels and through several channels. When online documentation and tutorials aren't enough, an online ISBE help desk will route requests to the relevant experts. nSBCs will advise users in their country on modelling strategies and feasibility of project proposals, and on what ISBE services may fit their needs. Questions that go beyond the current capabilities of automated tools may warrant more involved engagement of ISBE modelling expertise, either in model development or in connecting or refining existing tools. The resulting models or software solutions can then be offered more widely as new ISBE services. In agreement with the EU emphasis on open source tools, core software developed by ISBE will be made available and can then be installed locally or used to host resources for specific domains of research, either for reasons of confidentiality or privileged access. (Similar examples from bioinformatics include BLAST and genome browsers like Gbrowse.)

ISBE might also wish to facilitate access to resources and expertise outside ISBE, for instance matchmaking between experimentalists and a quality-controlled, categorized list of modelling experts. Similarly, ISBE may offer a certified guide to services available elsewhere that are widely used and may have wrappers in various programming languages. Many such directories exist in bioinformatics, as evident from a web search for [bioinformatics service directory](#); some examples are listed in Table 2 of the next section.

ISBE services and resources will complement existing national and European research infrastructures (RIs) as described in Deliverable D8.2, Section 1.4. Several countries are working to harmonize the operation of RIs. For example, the FAIRDOM project for interoperability in research, described in Appendix 7 to the ISBE Business Plan, ensures common standards among Germany, Switzerland, The Netherlands and the United Kingdom, as well as the ERASysApp programme.

## 2.4 Specification of services

Each service needs:

- A verbal description of what the service does and its domain of application. To enable users to quickly and reliably find further information, the description should refer to key publications or textbooks and use ontological identifiers<sup>3</sup> where available for biological systems, data types, models, and experimental protocols.
- A specification of the input data to be provided by the user or selected from available sources.
- A specification of the available operations offered by the service.
- A specification of the available outputs and data formats.

**Automated services** in particular will require more formal technical specifications<sup>4</sup> for their implementation. Such services may range from web forms to be filled in and submitted by users, to interactive tools to view data through the eyes of models specified on the fly, to web services<sup>5</sup> that expose an application programming interface for tool makers or end users fluent in programming. Such services are plentiful in bioinformatics; examples are given in Table 2 (see also Schattner 2008).

**Table 2. Some bioinformatics databases with interlinked information and programmable web services.**

Name	User-facing web pages	Web service documentation
KEGG (Kyoto Encyclopedia of Genes and Genomes)	<a href="http://www.kegg.jp">www.kegg.jp</a>	<a href="http://www.kegg.jp/kegg/docs/keggapi.html">www.kegg.jp/kegg/docs/keggapi.html</a>
NCBI (National Center for Biotechnology Information)	<a href="http://ncbi.nlm.nih.gov">ncbi.nlm.nih.gov</a>	<a href="http://eutils.ncbi.nlm.nih.gov">eutils.ncbi.nlm.nih.gov</a>
Ensembl	<a href="http://ensembl.org">ensembl.org</a>	<a href="http://rest.ensembl.org">rest.ensembl.org</a> , <a href="http://www.ensembl.org/info/data/biomart">www.ensembl.org/info/data/biomart</a>

## 2.5 Levels of service

Simple web pages will suffice for highly standardized analyses where users can pose a wide range of biological questions just by filling in a form and possibly uploading own data or models, as on the bioinformatics web sites listed in Table 2.

However, there is much that cannot be automated in application of mathematical models to new research questions. The particulars of a research question often requires tinkering, adaptation, simplification or refinement of the existing models. This is indeed how most new models arise (Smith et al. 2007, Waltemath et al. 2013), but doing it efficiently takes experience both in modelling and in the subject matter. It will be an important function of ISBE to support biologists in adding or refining model components, e.g. to reflect a newly-discovered ion channel in a heart cell (Smith et al. 2007, Nickerson & Buist 2008).

<sup>3</sup> See D8.2 Section 2 on description of modelling methodologies, and D2.2 on data and model management.

<sup>4</sup> [https://en.wikipedia.org/wiki/Specification\\_\(technical\\_standard\)#Information\\_technology](https://en.wikipedia.org/wiki/Specification_(technical_standard)#Information_technology)

<sup>5</sup> [https://en.wikipedia.org/wiki/Representational\\_state\\_transfer](https://en.wikipedia.org/wiki/Representational_state_transfer)

Even more labour intensive is the development of novel mathematical models, which places even higher demands on the skill and experience of both modeller and biologist. Such development is as much a craft as a science, usually with multiple options every step of the way, and lots of tacit knowledge about what may work well or not. It can be a demanding exercise with hard-to-predict payoffs, but nevertheless ISBE and funders may find it desirable for reasons of prestige to put some effort into enabling bleeding-edge modelling.

In support of the abovementioned activities, data and model stewardship is a very important kind of service for modelling. Putting clients' data (Salimi & Vita 2006) or models (Li et al. 2010) into annotated, standardized form makes them interoperable with other data and model resources and ISBE services.

## 2.6 Help desk and guidance

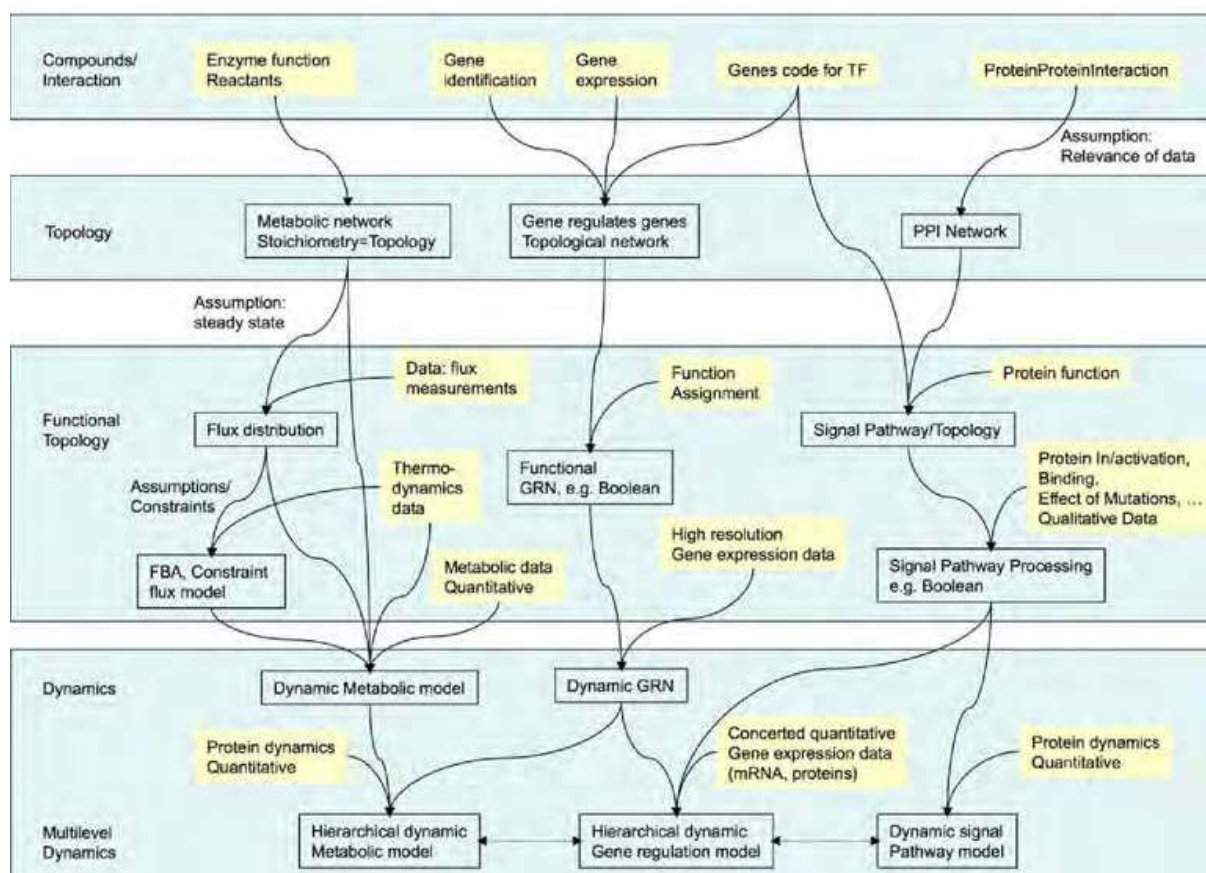
There will be several kinds and levels of documentation and user support, ranging from the fully automated to longer-lasting individual engagement of ISBE's human expertise. For example:

- Integrated documentation of interactive web pages, for instance captions, tooltips and links to details and specifications.
- Software wizards or "analysis assistants" to hide unneeded complexity from the user, presenting only options that are relevant for the user's chosen task.
- Textual and diagrammatic overviews of the various services offered, linking to more detail for each service.
- Technical documentation of web services, for use by advanced end-users and software developers, whether within or outside ISBE.
- A first-line help desk to route prospective clients to relevant services or user guides, or to the next level of human assistance. Help desk staff must be able to understand user needs and have a good overview of ISBE services, providers and resources.
- A second-line staff that provide in-depth guidance, acquiring a thorough understanding of their clients' needs. These providers can for instance advise on choice of tools, annotate and adapt client data for use with relevant tools, and tailor models to particular applications. This staff must be well trained in interoperability, standards and/or knowledge management within their various fields of expertise, though they need not be particularly expert in biology as such.
- Third-line service providers will be expert scientists participating directly in research projects, either as co-authors or consultants. Depending on client needs, these service providers may be experimental biologists, modellers, or mathematicians, though trained in effective communication with the rest of the ISBE system.

Many aspects of user support and service provision will be similar as in other research infrastructures such as ELIXIR. Adopting and adapting suitable existing systems will be part of ISBE's harmonization and cooperation with other European research infrastructures, including the projects CORBEL (for interoperability and coordination) and RITRAIN (for training of providers).

## 2.7 Routing requests

For modelling services to be useful, ISBE must succeed in adequately routing clients to services and providers. Knowing the purpose and data demands of each modelling approach is vital, whether this interface can be formalized (as it must for automated services, see Section 2.4) or can only be more loosely described. Once the data requirements are clear, it becomes easier for clients to figure out their options, and what additional data they might require to open more options for modelling (Figure 2). ISBE can then assist in referring clients to potential collaborators with expertise in how to acquire or interpret such additional data.



**Figure 2. "Decision tree" showing the data requirements for different kinds of metabolic and biochemical signalling models. This is also helpful for ISBE clients who "have data, want model" in identifying applicable modelling approaches, and points out what additional data would be required to include additional aspects of the system. (From Kahlem et al. 2011, Open Access.)**

The routing of clients to providers can initially rely on the overview of individual ISBE staff for a limited assortment of services. But eventually, ISBE will need to develop some kind of browsable product catalogue and directory of ISBE expertises, data and model resources. This extends to developing a vocabulary and ontology for modelling frameworks, including a



systematic overview of the problem domains and data requirements of given models. (See Deliverable D8.2, "A language to describe modelling needs and competences".)

## 2.8 Integration with experiment

Mathematical modelling is a powerful tool for summarizing, integrating and comparing experimental data. Conversely, mathematical models are of little use without data to support them. Confrontation with data through experiment is the primary means of evaluating the performance and applicability of candidate models for a given use. A testament to the power of tightly integrated theoretical-experimental research programmes can be found in Deliverable D8.1, "Portfolio of showcases illustrating the impact of modelling" and the section titled "Why modelling?" in D8.2. Thus, streamlining the confrontation of models with data is an important service of ISBE.

### 2.8.1 Semantically rich interfaces: ontological annotation

To link information on data, models and experiments, the information must be formatted and labelled in a common manner. *Ontological annotation* provides unambiguous identification of every variable, parameter or process in a model, data set or experimental procedure<sup>6</sup>. Many models in the BioModels database already have such annotation (Li et al. 2010), and data curated by ISBE will similarly be "model ready". The documentation, sharing and reuse of experimental protocols is detailed in Deliverable 2.2 and in Appendix 7 of the ISBE Business Plan.

### 2.8.2 Virtual experiments

A common strategy for model validation is to mimic real-world experiments in computer simulations, termed *virtual experiments* (Cooper et al. 2015). Similarly as for real-world experiments, there are several advantages to sharing and reusing protocols for virtual experiments:

- **Specifying, assaying, and comparing** the behavioural repertoires of models.
- **Reproducible research**, facilitating independent verification and critical scrutiny of scientific knowledge.
- **Guiding model reuse and composition**, by assessing whether a model performs as required in a given physiological context.
- **Quality assurance** in the application of computational biology models.

However, in current practice virtual experiments are rarely designed with reusability in mind. For example, most heart cell models in repositories implement a single hard-coded protocol, and substantial manual intervention is required to investigate how the system would perform under a different virtual experiment. Cooper et al. (2015) proposed a more flexible architecture for virtual experiments, based on two principles:

---

<sup>6</sup> Services such as [identifiers.org](http://identifiers.org) can help resolve identifiers in the form of URLs to specific ontologies and entries therein (Juty et al. 2012).

- **Separate experiment from model.** This involves identifying clear interfaces between the experimental and model components.
- **Annotate data, models and protocols** with unambiguous identifiers for variables and parameters.

The authors further discuss how this can be implemented by combining and extending standards such as SED-ML (simulation experiment description markup language) and ontologies for identification of model components (see above).

### 2.8.3 A continuous integration of knowledge

Such an architecture of virtual experiments would enable the routine application of experimental assays to new models, verifying that they work as intended (functional curation, Cooper et al. 2011) and comparing their performance against competing models and available datasets. It would also facilitate the updating of models and parameter estimates in light of new data, pointing the way towards a continuous integration of knowledge. Once in the system, each piece of information added provides new context for the existing ones, and thus the invested efforts bring rewards that continue to grow cumulatively. The high-throughput screening of data versus models will rapidly identify gaps in understanding, limitations in the data, and potential contradictions in current hypotheses.

By integrating functional curation as a public service, both for model repositories and model developers, researchers may easily characterise or benchmark their models under a wide range of scenarios (including easily varying published experiments to address their own specific setup), while checking whether standard behaviours are retained.

## 3 Clients

**Eligibility:** Work Package 3 (Overall infrastructure, eligibility and accessibility) is mandated with outlining criteria for eligibility of clients for ISBE services, and so here we only make a few observations and proposals. We expect that web tools with modest computational requirements will be freely available, as this would both benefit society and raise awareness of ISBE. Services involving substantial manual labour and intellectual effort can be allocated as part of funding agencies' grants, or paid for directly by industry. To attract users, initial services will be free for some time (see "Engaging clients" below).

**Training and education:** Here, "education" refers to teaching at universities, while "training" refers to post-graduate and life-long competence and skill acquisition. Training options for different user groups (academic, clinical, industrial) are outlined in section 3.2.6 of the ISBE Business Plan. Training was a primary concern of respondents to ISBE's survey of the industrial community. ISBE will inform about available training through the ISBE community website, and train clients in the use of ISBE tools and services. End-user training will be developed in collaboration with other research infrastructures such as ELIXIR, because systems biology ties together multiple disciplines and data sources. See also Work Package 10 (Training and education) and D8.2 Section 3.2.

**Engaging clients:** To encourage new clients to try out ISBE services, the first hours of service could be free for the client. For example, the first week of support in an initial marketing period, and perhaps the first day for routine operation. Once ISBE services have established a strong reputation, ISBE might establish a pricing model where end users pay for services and budget for these in grant proposals. This question is similar to that faced by other research infrastructures such as ELIXIR.

## 4 Providers

**Engaging and training modellers:** D8.2 Section 5.7 discusses how to motivate highly qualified modellers to provide ISBE services instead of pursuing more traditional careers. With adequate funding, modelling could become a new career path similar to biocuration, with schooling in a standard curriculum of methods and no expectation of co-authorship. While this will cover many common demands, many if not most applications of modelling in biology require substantial intellectual input, as seen in the "modelling success stories" reviewed in Deliverable D8.1.

Another option is for modellers to provide service while also building their own competence and scientific track record, implemented as follows. National centres award ISBE postdoc fellowships (competitively) to people who receive "ISBE-certified" training and then work in various experimental labs in academia or industry. This ensures that the national centres support end-users in their country, while adhering to common standards and remaining interoperable with international data and model resources, and keeps the modellers close to the wet-lab experts.

Thus, ISBE modelling services might involve three categories of people, each funded in different ways: Curators and developers that belong to the hub; full-time ISBE modellers that train others and can also be assigned to projects; and the ISBE-approved postdocs outlined above.

For the training of its managers and providers, ISBE will collaborate with other research infrastructures in the biomedical sciences (such as ELIXIR) through the projects CORBEL (Coordinated research infrastructures building enduring life-science services) and RITrain (Research infrastructures training programme).

## 5 Roadmap

The long-term aim of ISBE is to become the preferred infrastructure for systems biology whenever such research is funded. The actual ISBE service catalogue will depend on national willingness to fund, and strategies priorities of the interim and future ISBE. The channelling of ISBE services through funding calls will influence what services ISBE develops, and will be determined in dialogue between ISBE participants, funders and their advisors in industry, healthcare and other sectors. Here we outline how ISBE can get started with pilot services with relatively small investments, then appraise how ISBE can grow towards its ambition to cover systems biology across Europe.



## 5.1 Low-hanging fruit: First modelling services

As candidates for pilot services to be offered by the interim "ISBE light", we have identified modelling frameworks that span a wide range of biologically interesting questions within a well-defined set of user inputs. Importantly, many existing tools are already online and maintained by people involved in ISBE, so that ISBE light services can be established with relatively little effort.

- Ordinary differential equations<sup>7</sup>, where the rates-of-change of state variables is a function of the current value of those state variables.
- Constraint-based models<sup>8</sup>, based on stoichiometry in biochemical reaction networks.
- Knowledge integration across spatial scales, and the ability to integrate models and data at one spatial scale into larger scale models.

For each of these frameworks, there exist tools for common operations such as parameter estimation and sensitivity analysis. Many of these are exemplified at the JWS Online "live repository" (see Section 6 and the Appendix in Section 8). An important showcase of integration of modelling with wet-lab experiments (see Section 2.8) through reusable "virtual experiments" (Cooper et al. 2015) is the Functional Curation prototype for cardiac electrophysiology<sup>9</sup> at the University of Oxford.

The Open Physiology workflow<sup>10</sup> (de Bono et al. 2015) demonstrates the integration of sub-cellular mechanisms into a whole-body circulation framework using a web-based interface. This is made possible through the use of standard formats for encoding the mathematical models (CellML in this example) and the semantic annotation of the models at all spatial scales with biologically relevant information to ensure unambiguous mapping of data across the disparate models at each scale. Exposing this workflow with a collection of pre-defined standard analyses and providing users with the ability to select their models of interest enables users to rapidly investigate and explore their model's behaviour in an integrative biological system.

## 5.2 Preparing for growth: Scalability of modelling services

It is the ambition of ISBE to transform the practice of biological research and application. Success in this endeavour implies that the demand for modelling services will snowball, driven by many factors such as perceived usefulness to researchers, encouragement by funders, growing mindshare, and easier availability. How will ISBE deliver on this demand, assuming it gets to enjoy increased funding?

---

<sup>7</sup> Ordinary differential equation model: [http://identifiers.org/mamo/MAMO\\_0000046](http://identifiers.org/mamo/MAMO_0000046)

<sup>8</sup> Constraint-based model: [http://identifiers.org/mamo/MAMO\\_0000009](http://identifiers.org/mamo/MAMO_0000009)

<sup>9</sup> <https://chaste.cs.ox.ac.uk/FunctionalCuration>, source code at <https://bitbucket.org/ioncooper/fcweb>

<sup>10</sup> <http://open-physiology.org>

- **Recruitment of providers** depends on the development of study programmes based on a European core curriculum in systems biology (see D8.2, Section 4.1) as well as incentives for modellers to pursue careers within or in combination with ISBE (see D8.2, Section 5.7).
- **Routing of clients to providers** on a large scale will require a common language to structure the ISBE product catalogue, to describe modelling needs and competences, and to identify suitable models for given data and vice versa (see D8.2, Sections 2 and 5.6).
- **Strategic focusing of effort** (see D8.2, Section 5.3). For the bulk of services, it seems prudent to prioritize expressive frameworks<sup>11</sup> with a large potential user base. Likewise, involved personal support might prioritize solutions with scope for wider application, as these can be developed into more routine services. On the other hand, putting some effort into enabling bleeding-edge modelling might be desirable for reasons of prestige. This is a strategic choice for ISBE and funders to make.
- **Training and education of users** will reduce the overhead and cost for ISBE in offering services, as more biologists become *articulate users* of modelling (see D8.2, Sections 1.3 and 3.2).
- **Standardization and curation to streamline knowledge management** (see D8.2, Sections 5.4 and 5.5) will be important in expanding the range of research domains where modelling can be supported by services, e.g. in continuum dynamics and multiscale modelling of the heart (see showcase in Deliverable D8.1).

## 6 Empowering biologists: Some current examples

Implementing ISBE will involve some novel challenges, but also combining existing tools and learning from working solutions developed in various user communities. Below are a few examples that ISBE may connect with and/or learn from in empowering biologists to benefit from modelling in their daily work.

### 6.1 VPH-Share

**VPH-Share**<sup>12</sup> is a successful example of abstracting underlying infrastructure from the user-facing parts (cf. Section 2.3). Part of the Virtual Physiological Human endeavour, the project offers an online environment for the construction of biomedical workflows, including a "Master interface"<sup>13</sup> that abstracts underlying infrastructure from the user-facing parts. This project received a rating of "Excellent" in its final review.

---

<sup>11</sup> An expressive framework is one that can inform a wide range of biological questions, and accommodate a rich amount of user-provided or user-specified data.

<sup>12</sup> <http://www.vph-share.eu>

<sup>13</sup> <https://portal.vph-share.eu>

## 6.2 OpenCOR

**OpenCOR** (Garny & Hunter 2015) is a modelling and analysis program to work with CellML models, including those kept in the Physiome Model Repository (Yu et al. 2011)<sup>14</sup>. Open source and cross-platform, it has a plugin architecture that allows users to organize, edit, simulate and analyze CellML models. Its features facilitate model reuse and composition, structuring of models into submodels, and automatic unit checking. Besides model development and analysis, it can be used to annotate existing models and for educational and training purposes.

## 6.3 JWS Online

**JWS Online** is a web-based interactive tool for online analysis of SBML models<sup>15</sup>. In contrast to the cellml.org and biomodels.net repositories, which require users to download models for use with local software, JWS is a "live repository". As such, it exemplifies many features that ISBE may want to streamline. These include uploading and on-the-fly modification of models, interactive pathway diagrams and simulations, and ready-to-use procedures e.g. for sensitivity analysis. It also exemplifies integration with journal publishing and with the knowledge management system "The SEEK" (see WP2 Deliverable D2.2).

As an example of technical implementation, organization and interaction with journals, this last example is described in more technical detail in the Appendix (Section 8).

## 6.4 BioModels

**BioModels** (<http://www.ebi.ac.uk/biomodels>) is a repository of models of biological processes which can be accessed through a web interface or web services. All models are provided in the public domain, and may be downloaded, modified and re-purposed. Besides providing an extensively annotated and curated set of peer-reviewed models which fully comply with MIRIAM guidelines, it also provides a significant number of automatically generated models, many of which are at the genome scale. Importantly, these automatically generated models can be used as a framework into which additional information can be incorporated to generate quantitative models. BioModels provides facilities to simulate models, the ability to download and inter convert models in multiple formats and versions, as well as providing manually and automatically generated reaction graphs. BioModels is currently transitioning to an improved infrastructure which will additionally facilitate nominated model sharing (for author/author, author/curator and author/publisher collaboration), provide full version tracking and permit live editing. This new infrastructure will additionally support model submission from a wider range of domain-specific formats.

---

<sup>14</sup> <https://models.physiomeproject.org>

<sup>15</sup> There also exists a separate JWS model format, though this is interconvertible with SBML.

## 7 References

- De Bono B, Safaei S, Grenon P, Nickerson DP, Alexander S, Helvensteijn M et al. (2015) The Open Physiology workflow: modeling processes over physiology circuitboards of interoperable tissue units. *Computational Physiology and Medicine* 6: 24.
- Cooper J, Mirams GR, Niederer SA (2011) High-throughput functional curation of cellular electrophysiology models. *Progress in Biophysics and Molecular Biology* 107: 11–20.
- Cooper J, Vik JO, Waltemath D (2015) A call for virtual experiments: Accelerating the scientific process. *Progress in Biophysics and Molecular Biology* 117: 99–106.
- Garny A, Hunter PJ (2015) OpenCOR: a modular and interoperable approach to computational biology. *Frontiers in Physiology* 6.  
<http://journal.frontiersin.org/Article/10.3389/fphys.2015.00026/abstract>
- Juty N, Le Novère N, Laibe C (2012) Identifiers.org and MIRIAM Registry: community resources to provide persistent identification. *Nucleic Acids Research* 40: D580–D586.
- Kahlem P, DiCara A, Klipp E, Segal E, Birney E, Xenarios I et al. (2011) Strengths and weaknesses of selected modeling methods used in systems biology. In: Yang N-S (ed) *Systems and computational biology: bioinformatics and computational modeling*, 77–98. INTECH Open Access Publisher. <http://dx.doi.org/10.5772/22126>
- Li C, Donizelli M, Rodriguez N, Dharuri H, Endler L, Chelliah V et al. (2010) BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology* 4: 92.
- Nickerson D, Buist M (2008) Practical application of CellML 1.1: The integration of new mechanisms into a human ventricular myocyte model. *Progress in Biophysics and Molecular Biology* 98: 38–51.
- Peng RD (2011) Reproducible Research in Computational Science. *Science* 334: 1226–1227.
- Salimi N, Vita R (2006) The Biocurator: Connecting and Enhancing Scientific Data. *PLoS Computational Biology* 2. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1626147/>
- Schattner P (2008) *Genomes, Browsers and Databases: Data-Mining Tools for Integrated Genomic Databases* 1 edition. Cambridge University Press, Cambridge UK ; New York
- Smith NP, Crampin EJ, Niederer SA, Bassingthwaight JB, Beard DA (2007) Computational biology of cardiac myocytes: proposed standards for the physiome. *J Exp Biol* 210: 1576–1583.
- Waltemath D, Henkel R, Hälke R, Scharm M, Wolkenhauer O (2013) Improving the reuse of computational models through version control. *Bioinformatics*.  
<http://bioinformatics.oxfordjournals.org/content/early/2013/01/18/bioinformatics.btt018>

Yu T, Lloyd CM, Nickerson DP, Cooling MT, Miller AK, Garny A et al. (2011) The Physiome Model Repository 2. *Bioinformatics* 27: 743–744.

## 8 Appendix: Case study JWS Online

This appendix was provided on request by Prof. Jacky Snoep of Stellenbosch University, to illustrate an actual implementation of a "live" model repository as described in Section 6.

### 8.1 Introduction

The JWS Online web site gives access to a database of curated models of biological systems and allows users to run these models in a web browser. The aim of this service is to give users quick and easy access to published models. All models are curated such that they reproduce the simulation results for the model as published. JWS Online is linked to a number of scientific journals for which it provides a technical curation service during the reviewing and publication cycle. In addition, JWS Online is linked to research initiatives for which it provides an online simulation tool and model curation services. JWS Online is available at three mirror sites: <http://jji.bio.vu.nl>; <http://jji.mib.ac.uk>; and <http://jji.biochem.sun.ac.za>.

### 8.2 The JWS database

In 2000, JWS Online started as a repository of models that could be simulated as Java applets, giving the user the opportunity to quickly test published models. Currently JWS Online uses a database implementation with a native format that mirrors the SBML specification. JWS Online supports uploading and modification of existing models in SBML or JWS format. A user can select a model from the database via a number of filters, such as: model name, organism, process, or model type. This set of filters will be extended to allow the user to search for specific model components and annotations.

### 8.3 The JWS simulator

The models in JWS Online can be interrogated via a browser interface. In this way the models are easily accessible and no additional software is required to simulate the models. The simulations run in a server so the user does not need a strong computer; access to a web browser is all that is needed. Help documentation is made available via a number of short instruction videos.

To keep simulation time low, the functionality of the JWS Online simulator is restricted to a few commonly used tasks, such as time integration,

steady state analysis, parameter scans and metabolic control analysis.

All simulations can be stored as SED-ML archives, which contain all the necessary information to reproduce the simulation on other SED-ML enabled simulators.

### 8.4 Model coding and curation

Models are coded from the literature, preferably from SBML files when these are made available by the authors, but also other electronic formats are used and sometimes models are coded straight from the equations in the paper. Models are curated to make sure that the simulations shown for the model in the literature can be reproduced. If model simulations cannot be reproduced, the authors are contacted to resolve the problem. This is often a labour-intensive and time-consuming process, especially when models are coded from the literature, as model descriptions in manuscripts are often incomplete or contain errors. During the technical curation in the review process of papers with mathematical models, the JWS Online team tests both model simulations and the description of the model in the paper, to minimize the chance of erroneous model publications and irreproducible model simulations.

### 8.5 Model types

Most of the models in the JWS Online database are models described in ordinary differential equations (ODEs, and differential-algebraic equations, DAEs). This is for the simple reason that most systems biology models are of this type. These model types are currently being extended to include constraint-based models, stochastic models and PDEs. The model types will be extended as different model formalisms are being used in computational biology studies.

### 8.6 Model editing and model creation

Originally JWS Online allowed users only limited possibilities to make changes to existing models; these were changes in values for parameters and



initial values of variables. Basically, the simulator gave access to the published model, which was rather restrictive, and later a model editing function was added to JWS Online. This editing functionality loads the original model in an editable interface that strongly follows the SBML model specification. The editor supports MIRIAM and SBO standards for annotation, and links to a useful online annotation tool (semantic SBML annotator). This editor replaces the OneStop tool that was used in earlier versions of JWS Online for model construction and annotation.

Upon request, users can obtain a login for JWS Online, which makes it possible to store user-defined models in the database, which are only accessible to the user. The model editor only stores models as session variables, so it is advisable to download models or store them as a registered user.

The model editor is a strong tool for editing of an existing model, but is a bit overpowering to use for model construction from scratch for novel modellers, and this will be improved.

### 8.7 Teaching and research

The ease of use of the JWS Online simulator and the immediate access to a large number of published models make it an ideal teaching tool for introductory modelling courses. Students can critically test hypotheses from literature and make changes to existing models to become familiar with the modelling process. More advanced modellers might want to quickly run some simulations with a published model, and can then decide to download the model and simulate it in a stand-alone simulator.

### 8.8 Data and model management in systems biology projects

Many systems biology projects are interdisciplinary and consist of several groups. Good data and model management tools are necessary for the exchange, analysis and linking of the different data and model files. The SEEK platform (used in the SysMO, Virtual Liver and ERASysApp projects) has integrated JWS Online as its model simulator from the very beginning.

### 8.9 Data and model linking

Model construction and validation is usually an iterative process between experimental analyses

and model simulations. To record such iteration cycles, a good versioning system and link between data and model files is important. The SEEK interface makes this possible in ISA structures and strong versioning of models and data files.

Upon publication of newly constructed models it is important to show how experimental data were used for parameterization of the rate equations. In JWS Online this can be done in rate plots for the individual processes. Such a plot shows the dependencies of a rate equation for its parameters and can be shown together with the experimental data used for the characterization of the reaction process.

### 8.10 Web services

JWS Online had a full set of SOAP web services but is now moving away from SOAP to RESTful services. A minimal set of web services is already available, and more functionality will be added in the near future.

### 8.11 Sustainability and scalability

JWS Online has always been a rather small initiative, driven by 2 to 4 researchers at Stellenbosch University, the Vrije Universiteit – Amsterdam and at the University of Manchester. At Stellenbosch University, two staff members (permanently employed) can maintain the infrastructure at the three mirror sites of JWS Online. This has been maintained for 15 years, and long-term guaranteed funding for these researchers is available to maintain this service for at least another 15 years.

Development of new functionality is performed by postdoctoral fellows at Stellenbosch and Manchester and is dependent on research grants that are typically for 3-5 year periods.

Model curation is of crucial importance to guarantee the quality of the models in the database. Currently, model curation is a very time-consuming process because many researchers do not provide their models in correct SBML formats, models are poorly described in manuscripts, and model simulations in figures are poorly annotated. This has limited the number of journals that JWS Online can effectively work with, and would also limit the support that can be provided for model curation in large systems biology initiatives. Although this can improve with time if researchers stick better to modelling



standards, at present this is a bottleneck for effective support in model management.