



FAIRDOM – FAIR Asset management and sharing experiences in Systems and Synthetic Biology

Prof Carole Goble

The FAIRDOM Consortium

carole.goble@manchester.ac.uk

<http://fair-dom.org>, <http://fairdomhub.org>



SCIENTIFIC DATA



OPEN

SUBJECT CATEGORIES

- » Research data
- » Publication characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson et al.*

Received: 30 December 2015

Accepted: 12 February 2016

Published: 15 March 2016

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measurable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

Supporting discovery through good data management

Good data management is not a goal in itself, but rather is the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse by the community after the data publication process. Unfortunately, the existing digital ecosystem surrounding scholarly data publication prevents us from extracting maximum benefit from our research investments (e.g., ref. 1). Partially in response to this, science funders, publishers and governmental agencies are beginning to require data management and stewardship plans for data generated in publicly funded experiments. Beyond proper collection, annotation, and archival, data stewardship includes the notion of ‘long-term care’ of valuable digital assets, with the goal that they should be discovered and reused for downstream investigations, either alone, or in combination with newly generated data. The outcomes from good data management and stewardship, therefore, are high quality digital publications that facilitate and simplify this ongoing process of discovery, evaluation, and reuse in downstream studies. What constitutes ‘good data management’ is, however, largely undefined, and is generally left as a decision for the data or repository owner. Therefore, bringing some clarity around the goals and desiderata of good data management and stewardship, and defining simple guideposts to inform those who publish and/or preserve scholarly data, would be of great utility.

This article describes four foundational principles—Findability, Accessibility, Interoperability, and Reusability—that serve to guide data producers and publishers as they navigate around these obstacles, thereby helping to maximize the added-value gained by contemporary, formal scholarly digital publishing. Importantly, it is our intent that the principles apply not only to ‘data’ in the conventional sense, but also to the algorithms, tools, and workflows that led to that data. All scholarly digital research objects²—from data to analytical pipelines—benefit from application of these principles, since all components of the research process must be available to ensure transparency, reproducibility, and reusability.

There are numerous and diverse stakeholders who stand to benefit from overcoming these obstacles: researchers wanting to share, get credit, and reuse each other’s data and interpretations; professional data publishers offering their services; software and tool-builders providing data analysis and processing services such as reusable workflows; funding agencies (private and public) increasingly

*Correspondence and requests for materials should be addressed to B.M. (email: benedikt.mons@ttu.edu).

¹A full list of authors and their affiliations appears at the end of the paper.

FAIR

Findable
Citable
Accessible
Track & Countable
Interoperable
Intelligible
Reusable
Reproducible

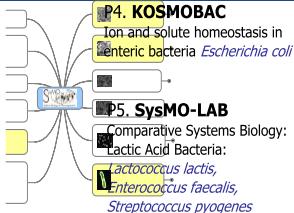




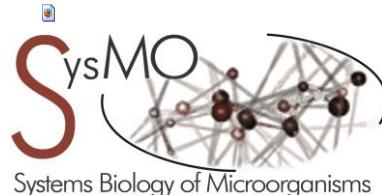
Findable
Accessible
Interoperable
Reusable Data
 Operations
 Models

Systems and Synthetic
Biology Projects

Sys Bio Projects



P6. **PSYSMO**
Systems analysis of biotech induced stresses: towards a quantum increase in process performance in the cell factory
Pseudomonas putida



2006
11 projects



Acronym	Project Title
SysMetEx	Systems Biology of acidophile biofilms for efficient metal extraction
SysMilk	Designer microbial communities for food production
SysVirDrug	Production of viral drugs
MetApp	Metabolic engineering of microorganisms
IMOMESIC	Microbial community engineering for industrial applications
WineSys	Systems biology approach to wine production
Cropclock	Microbial clock for crop improvement

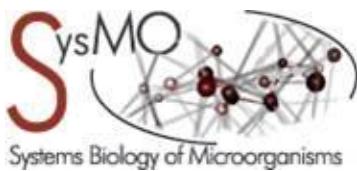


2015
12 projects



Stakeholders Stress...

Sponsored to Support Projects: National level and ERANets



SystemsX.ch
The Swiss Initiative in System Biology



Funders / Institutions

- Capitalising on investments
- Skills
- Justification, Audit, Compliance
- Showcase access

Publishers

- Reproducibility
- New publishable assets
- New services
- Showcase access



Projects....

Researchers' Productivity Rhetoric



Multi-partner
Distributed
Dynamic membership
Overlaps
Sensitivities

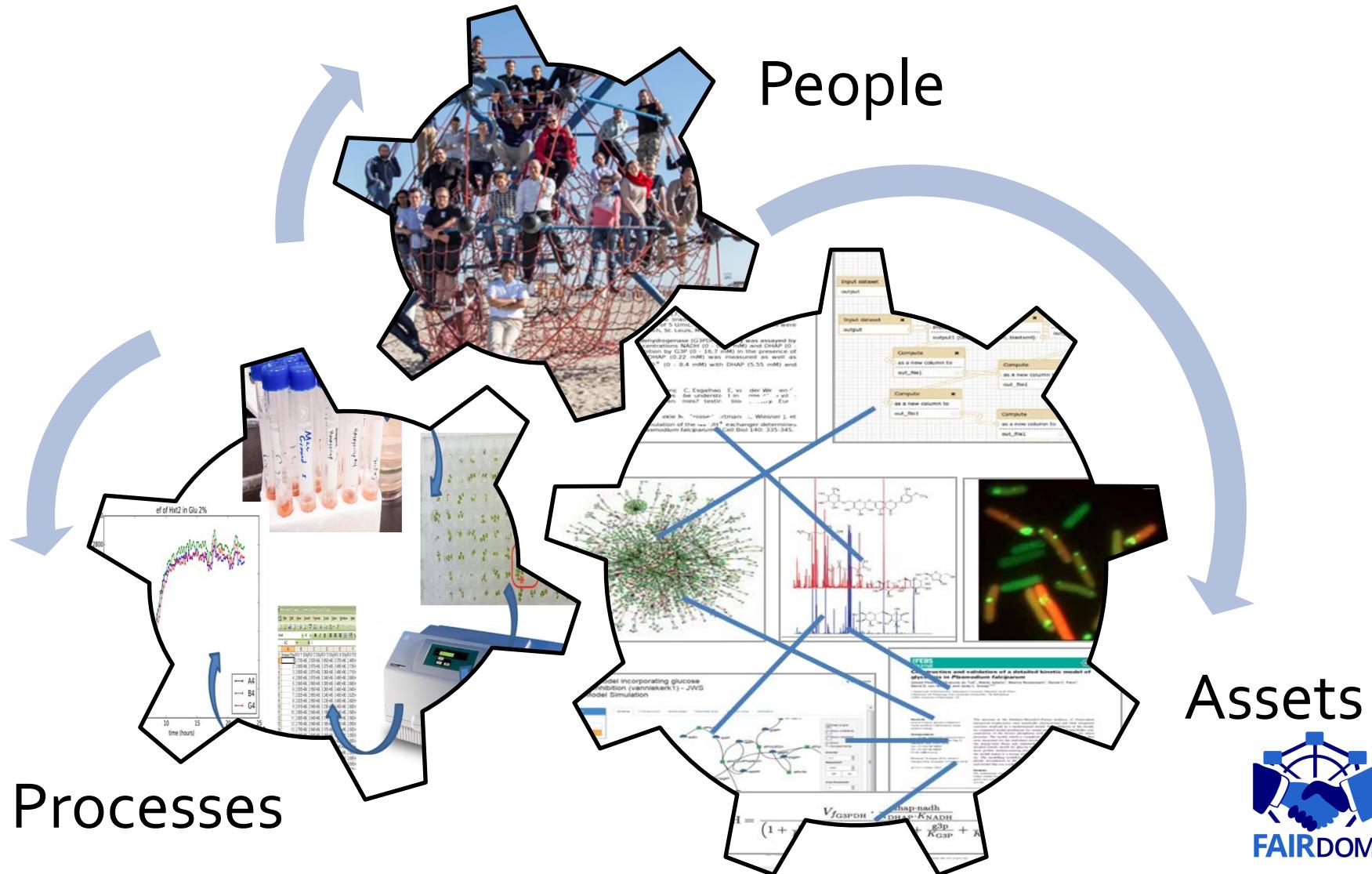
Collaboration
Publication
Showcasing
Compliance
Retention

Home-grown resources
Public resources
Mixed Skills



Projects' "PAP-PA"

People, Assets, Processes -> Publishing, Analysis



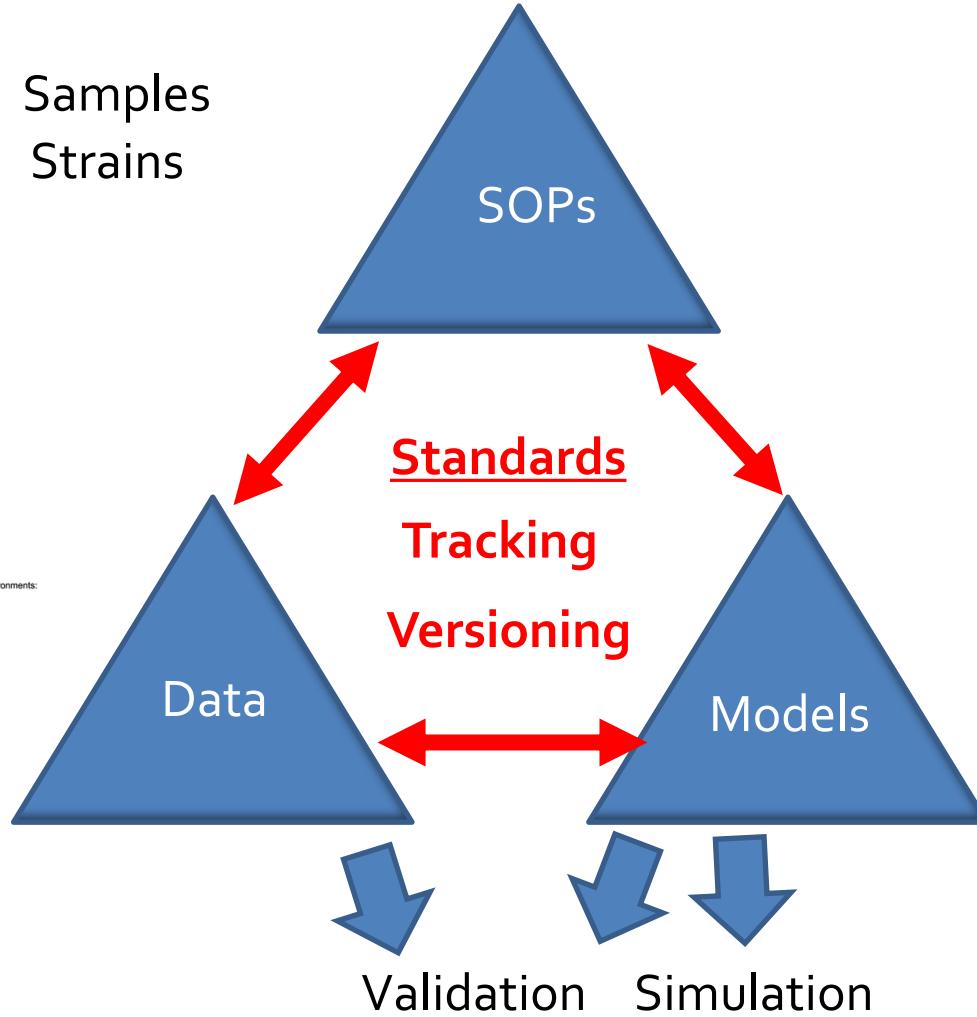
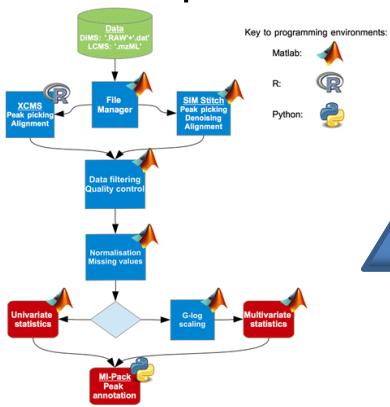
Projects' People, Assets, Processes

Organisation, Communication, Dissemination



Samples
Strains

Analytics
Pipelines

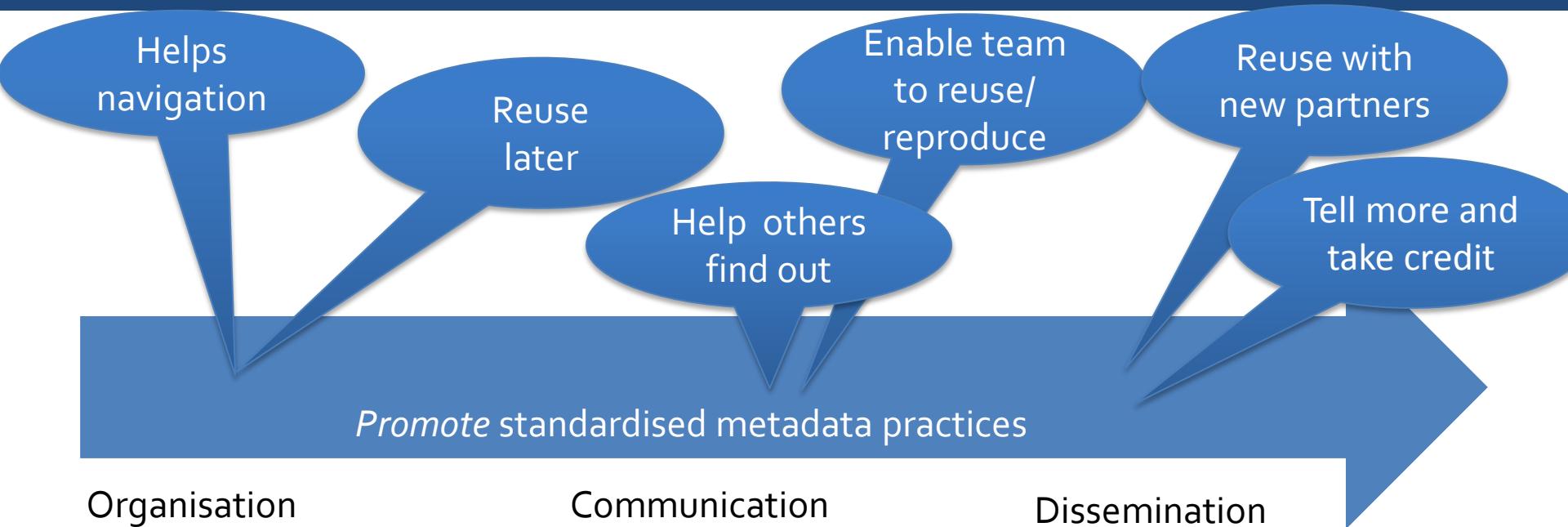


Articles

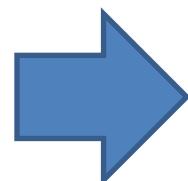


experimentalist
modeller
FAIRDOM

Upstream, downstream assets discovery



Collecting and tracking data/models
Choosing what to keep
Preparing what to share and when



Most data/models won't be shared

- Wrong experimental method
- Hidden parameter discovered
- Faulty experiment



FAIR Projects & Programmes

Is there any group generating kinetic data?

Track versions of my model

Who is working with which organism?

Which data belong to
which publications?

What methods are been used to determine enzyme
activity?

Whats the relationship between the data
and model

Under which experimental conditions are my partners
working on for the measurement of glucose
concentration?

What is the provenance of the parameters for this version
What SOP was used for this^{of} the model?
sample?

Where is the validation data for this model?

SOP Directory

Yellow Pages

Spreadsheet
tools

Auto-
harvesting

Is this data available?



FAIR Projects and Programmes



Findable

Exchange & find
assets and
people
Citation
Credit



Accessible

Share, disseminate and publish
assets *sensitively*
Gateway to third party tools,
archives
Store assets
Package assets



Interoperable

Standards!

Organise and link assets
Maintain the experimental context
Retain results beyond a project
Reuse results, tools, archives
Respect local solutions



Reusable

Standards!

Track collection of data and metadata
Consistent reporting for interpretation,
interop & comparison
Promote and support standardised metadata
practices.
Support reproducible publications

FAIRDOM Initiative



The University of Manchester



Community Actions



Project Support



Public Project Commons



Platforms, Tools



FAIRDOM Community, Knowledge Hub

<http://www.fair-dom.org>

The screenshot shows the FAIRDOM website homepage. At the top left is a video thumbnail titled "What is data management - and why do you need it in interdisciplin..." with a play button. Next to it is another video thumbnail titled "DATA & MODEL STANDARDS" with a hand writing "UP TO DATE" on a circular arrow. Below these are two news cards: one for "SBMC 2016 - 6th Conference on Systems Biology of Mammalian Cells" in Munich, Germany, and another for "BioSB 2016" in Lunteren, Netherlands. The "Upcoming Events" section lists "Wednesday, April 6th, 2016 - Friday, April 8th, 2016" and "Tuesday, April 19th, 2016 - Wednesday, April 20th, 2016". The "Latest News" section features a card for "Join our User Meeting" and another for "New Blog: Samples Club". The "Grow Your Community" section features a graphic of a molecular structure and links to "ISBE", "FORUMS", "PALS", "SAMPLES CLUB", "FOUNDRY", "WEBINARS", and "VIEW ALL".



Know-how, Guides, Templates, Workshops, Training,
Webinars, Standards and Policy Forums





BGRS\SB'2016

The 10th International Conference

On Bioinformatics of Genome Regulation
and Structure\Systems Biology

29 August - 2 September, 2016



Young Scientists School, SBB-2016

The 8th Young Scientists School “Systems biology and Bioinformatics” **SBB-2016** will be held before the BGRS/SB-2016 multi-conference on **22-25th of August**.



Policy and International Initiatives

STRATEGIC
RESEARCH AGENDA
SYSTEMS BIOLOGY
IN EUROPE 2016

July 2016

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 321567.

ERASys APP
ERA-Net for Applied Systems Biology

WORKSHOP REPORT

Career Pathways in Multidisciplinary Research:
How to Assess the Contributions of Single
Authors in Large Teams

BRUSSELS, 1 AND 2 DECEMBER 2015

ISBE Infrastructure for Systems Biology Europe

deNBI GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE

SCIENCE EUROPE Life, Environmental and Geo Sciences Committee

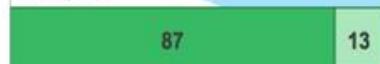
FAIRDOM

Surveys

Difficulty Reusing Models (% Researchers)

■ Yes ■ No

Modeller



Experimentalist



We found the following did not have enough annotations for reuse...

Model Parameters (67%)

Model Reactions (60%)

Model Species (58%)

Model Metadata (25%)



Hard Disk
(71%)



Shared File System
at Institute (58%)



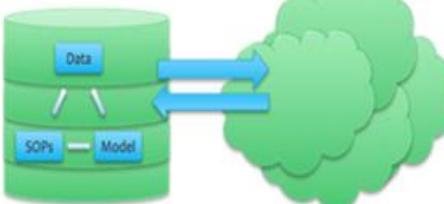
Project Database/Content
Management System (31%)



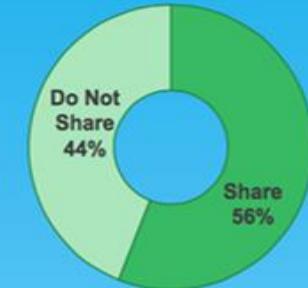
File Transfer
Server (18%)



Bespoke Systems
Biology System (7%)

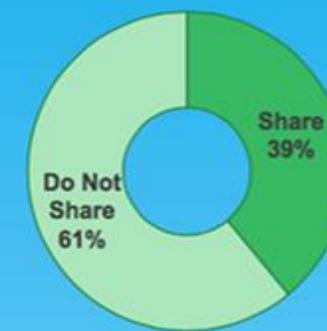


Models



BioModels	33%
JWSOnline	8%
CellML	7%

Data



GEO	19%
Array Ex.	11%
ENA	6%

Community Clubs

<http://www.fair-dom.org>

Samples Club with ELIXIR, BBMRI-ERIC, EBI...



Rework and harmonise
sample metadata
framework
bioschemas.org

Developers Foundry



Support developers of Systems Biology
tools and platforms

3rd Foundry meeting
Dec 1-2 2016, Frankfurt



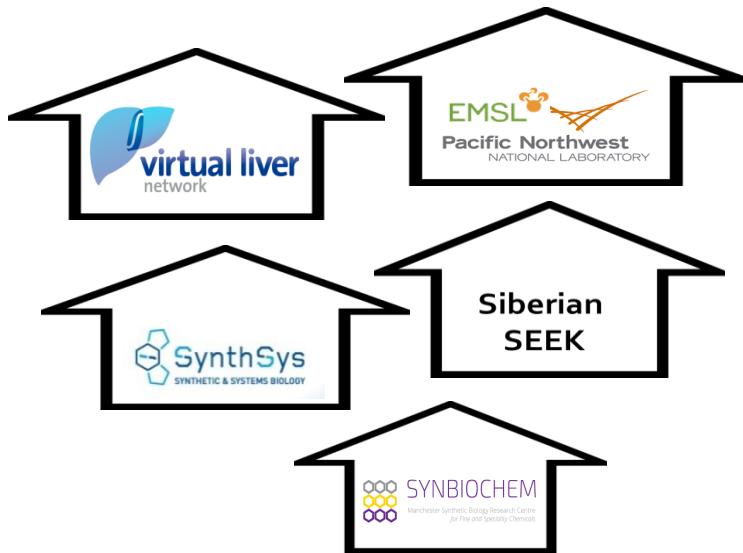
FAIRDOM Platforms and Tools



FAIR Sharing
Metadata Catalogue
Project Commons
Cross-repository gateway
Tool gateway
Results repository

Local Storage
Analytics
Tracking
LIMS, Auto-archiving
In flight repository

FAIRDOM Platform Installations



Local retention and In flight management,
Private sharing
Centres, large or national projects
Local skills



One stop showcase
Programmes
Post-project retention
Supplementary materials

*Troup, E.; Clark, I; Swain, P; Millar, AJ; Zielinski, T (2015) Practical evaluation of SEEK and openBIS for biological data management in SynthSys <http://hdl.handle.net/1842/12236>

People and Project Commons

FAIRDOMHub.org



Sandve Simen



Projects: GenoSysFat, DigiSal
Institutions: Norwegian University of Life Sciences

Disciplines: Bioinformatician, Experimentalist
Roles: Project Coordinator, Member
Expertise: evolution, Genomics
Tools: Not specified

Vitor Martins Dos Santos



Projects: PSYMS
Institutions: Helmholtz Zentrum München

My research area and strong interest group are to co

self-managed workspaces

GenoSysFat

Salmon farmed on modern feeds contains less of the healthy, long-chain fatty acids (EPA and DHA) than before. Up until the turn of the millennium, farmed salmon were fed fish oil as a replacement for their omega-3 rich natural prey. However, fish oil is now a scarce resource, and more than half of the fat in modern feeds comes from plant oils that are inexpensive, but devoid of long-chain omega-3 fatty acids. How can we increase the omega-3 content of salmon on sustainable feeds?

One option is

...
Programme: Independent Projects
Public web page: <http://www.nmbu.no/fo>

SysMetEx

Biomining is a biotechnological process called 'heap bioleaching' where typical base.

The role of the different types of microorganisms
...
Programme: ERASysAPP
Public web page: <http://sysmetex.eu/index.html>

SysVirDrug

Translating Systems Virology data

Programme: ERASysAPP
Public web page: Not specified

Sharing

Here you can specify who can view the summary of and edit the Investigation. More info

- Keep this Investigation private (only visible to you)

Or share it with..

- Members of Projects associated with this Investigation: Manage
... and all other registered users: View summary
 - All visitors (including anonymous visitors with) View summary
View and edit summary

Advanced permissions

Additional fine-grained sharing permissions

No one

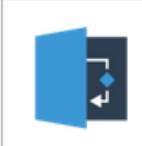
Build up the list of people and groups to share with. Select from the options below and click "Add" to apply your choice

Favourite groups
Projects and Institutions
Individual People

Sharing sensitivity

Whole body modelling of glucose metabolism in malaria patients

Hypoglycaemia and lactic acidosis are key diagnostics for poor chances of survival in malaria patients. In this project we aim to test to what extent the metabolic activity of Plasmodium falciparum contributes to a changed glucose metabolism in malaria patients. The approach is to start with detailed bottom up models for the parasite and then merge these with more coarse grained models at the whole body level.



Programme: SARCHI: Mechanistic modelling of health and epidemiology

Public web page: Not specified

Organisms: Plasmodium falciparum, Homo sapiens

FAIRDOM PALs: No PALs for this Project

Project administrators: Dawie Van Niekerk, Jacky Snoep

Related items

[People \(5\)](#)[Institutions \(3\)](#)[Investigations \(1\)](#)[Studies \(3\)](#)[Assays \(24\)](#)[Data files \(16\)](#)[Models \(23\)](#)[SOPs \(17\)](#)[Publications \(3\)](#)

Kathleen Green



Projects: Whole body modelling of glucose metabolism in malaria patients
Institutions: Stellenbosch University

Disciplines: Not specified
Roles: Not specified
Expertise: Not specified
Tools: Not specified

Dawie Van Niekerk



Projects: SysMO DB, Whole body modelling of glucose metabolism in malaria patients, Molecular Systems Biology
Institutions: University of Stellenbosch, Stellenbosch University
Email: ddvniekerk@sun.ac.za
Web page: Not specified
Phone: +27611305989
Skype: dawie.vanniekerk

Disciplines: Modeller
Roles: Not specified
Expertise: Not specified
Tools: Not specified

Erik Rust



Projects: Whole body modelling of glucose metabolism in malaria patients
Institutions: Stellenbosch University

Disciplines: Not specified
Roles: Not specified
Expertise: Not specified

- [Yellow pages](#)
- [Programmes](#)
- [Projects](#)
- [Institutions](#)
- [People](#)
- [Experiments](#)
- [Investigations](#)
- [Studies](#)
- [Assays](#)
- [Assets](#)
- [Data files](#)
- [Models](#)
- [SOPs](#)
- [Publications](#)
- [Biosamples](#)
- [Activities](#)
- [Presentations](#)
- [Events](#)
- [Other](#)
- [Organisms](#)

FAIRDOM [Browse](#) • [Help](#) • [Search](#)

Home Investigations Index Glucose metabolism in Plasmodium falciparum trophozoites

Glucose metabolism in Plasmodium falciparum trophozoites

The investigation entails the construction and validation of a detailed mathematical model for glycolysis of the malaria parasite Plasmodium falciparum in the blood stage trophozoite form.

ID:56

Projects: Whole body modelling of glucose metabolism in malaria patients

Selected item: Investigation: Glucose metabolism in Plasmodium falciparum trophozoites [Full graph \(8\)](#)

```

graph TD
    Inv[Investigation: Glucose metabolism in Plasmodium falciparum trophozoites] --> ModelConstruction[Study: Model construction]
    Inv --> ModelValidation[Study: Model validation]
    Inv --> ModelAnalysis[Study: Model analysis]
    Inv --> Publication[Publication: Construction and validation of a detailed kinetic model of glycolysis in Plasmodium falciparum trophozoites]
  
```

Related Items

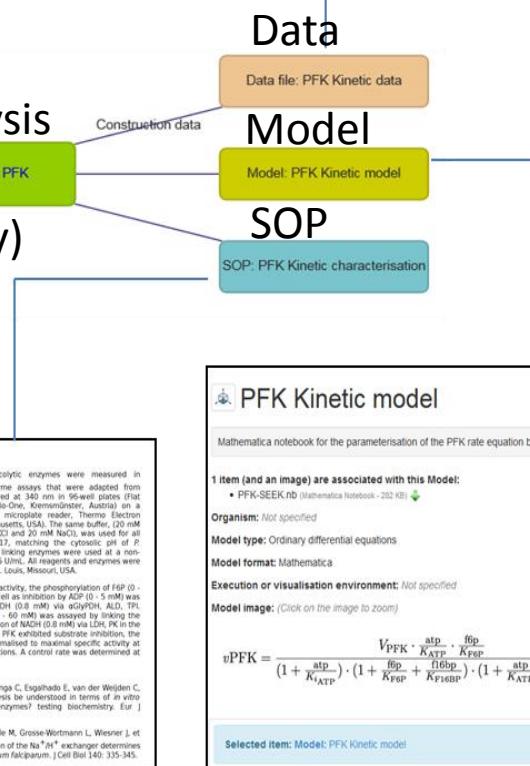
- [People \(1\)](#)
- [Projects \(1\)](#)
- [Studies \(3\)](#)
- [Assays \(24\)](#)
- [Data files \(16\)](#)
- [Models \(19\)](#)
- [SOPs \(13\)](#)
- [Publications \(1\)](#)

David Van Niekerk

Projects: SysMO DB, Whole body modelling of glucose metabolism in malaria patients
Institutions: University of Stellenbosch

Disciplines: Modeller
Roles: Not specified
Expertise: Not specified
Tools: Not specified

A	B	C	D	E	F	G	H	I
1. Metadata	Values (example)							
2. Assay SEEK ID	PFK_Kinetic_Data							
3. Assay Title	Darren van Niekerk							
4. Uploader								
5. Project SEEK ID								
6. Project								
7. ARRAY								
8. Assay SEEK ID	0							
9. Assay Title	PFK							
10. Assay Type	metabolicAssay							
11. Technology Type	initial rate experiment							
12. Description	Kinetic characterisation of PFK. Exper							
13. Experimentalist	Overall Predictor							
14. Date								
15. SOP								
16. Publication (optional)								
17. Temperature	pH							
18. pH	7.45							
19. Compound (of concern)								
20. User								
21. End value (optional)								
22. End value (optional)								
23. Comments								
24. Culture growth	Batch							
25. FACTORS_STUDIED								
26. User	concentration	concentration	concentration	concentration	concentration			
27. End value (optional)	ATP	ATP	ADP	ATP	ADP			
28. End value (optional)	nM	nM	nM	nM	nM			
29. SD (optional)	0.23475	0	0	0	0			
30. SD (optional)	10	5	5	5	50			



✓ Supply demand analysis on penkler1 model



Contributors: [Dawie Van Niekerk](#), [Jacky Snoep](#)

Model type: Ordinary differential equations

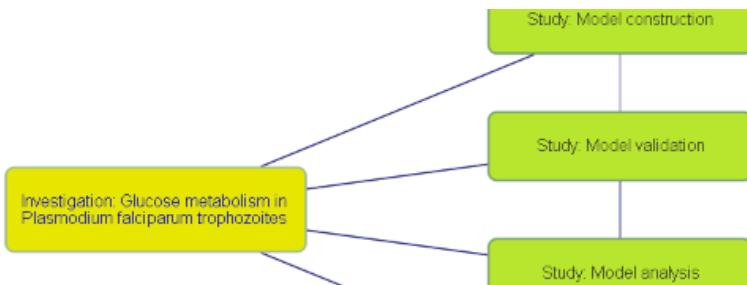
Model format: Mathematica



Organism: Not specified

Environment: Not specified

Created: 10th Jul 2015 at 12:30, Last updated: 23rd Nov 2015 at 14:09



Packaging

Related Items

Dawie Van Niekerk  



Projects: [SysMO DB](#), [Whole body modelling of glucose metabolism in malaria patients](#), [Molecular Systems Biology](#)

Institutions: University of Stellenbosch, Stellenbosch University

Disciplines: Modelle

Roles: Not specified

Expertise: Not speci

Tools: Not specified

[Home](#) / [Projects Index](#) / [GenoSysFat](#)

GenoSysFat

[Asset report](#) [Administration](#)

Salmon farmed on modern feeds contains less of the healthy, long-chain fatty acids (EPA and DHA) than before. Up until the turn of the millennium, farmed salmon were fed fish oil as a replacement for their omega-3 rich natural prey. However, fish oil is now a scarce resource, and more than half of the fat in modern feeds comes from plant oils that are inexpensive, but devoid of long-chain omega-3 fatty acids. How can we increase the omega-3 content of salmon on sustainable feeds?

One option is to breed salmon that are well adapted to the feeds of the future. There is heritable variation in salmon's ability to build EPA and DHA from shorter omega-3 fatty acids. The DNA sequence of salmon is now well known, allowing rapid characterization of heritable differences in nutrient utilization. A salmon family that appears promising on one feed, may not be the best on another. Therefore, we need to understand the salmon's body as a system: a functional whole made up of parts that mutually affect and also depend on, each other. A systems understanding of the interplay between genes and genetic variants will allow a tailoring of fish to feed and vice versa, based on the availability of feedstuff availability and pricing.

As a first step towards such a systems understanding of salmon metabolism, it involves two biological experiments. 1) A feeding experiment comparing high- vs low-omega-3 diets and salmon genotypes that differ in feed utilization. This will be followed by a detailed study of the metabolic pathways for each fish how different feeds affect metabolism and gene activity. This will allow a faster and more detailed exploration of the interplay between genetics and feeds. Results will be interpreted with the help of mathematical models for the biochemical reaction networks, which are well established for other species and will be adapted to salmon based on the new sequence of salmon genome.

Programme: The Digital Salmon

Public web page: <http://tinyurl.com/genosysfat>

Internal web page: Not specified

Organisms: *Salmo salar*, *Danio rerio*, *Oncorhynchus mykiss*

FAIRDOM PALS: No PALS for this Project

Project administrators: Jon Olav Vik

Asset housekeepers: No Asset housekeepers for this Project

Asset gatekeepers: No Asset gatekeepers for this Project

Related items

People (25) Institutions (5) Investigations (4) Studies (6) Assays (25) Data files (32) Models (2)
SOPs (7) Publications (5) Events (1)

[Home](#) / [Projects Index](#) / [DigiSal](#)

DigiSal

[Asset report](#) [Administration](#)

Towards the Digital Salmon: From a reactive to a pre-emptive research strategy in aquaculture (DigiSal)

Salmon farming in the future must navigate conflicting and shifting demands of sustainability, shifting feed prices, disease, and product quality. The industry needs to develop a flexible, integrated basis of knowledge for rapid response to new challenges. Project DigiSal will lay the foundations for a Digital Salmon: an ensemble of mathematical descriptions of salmon physiology combining mathematics, high-dimensional data analysis, computer science and measurement technology with genomics and experimental biology into a concerted whole.

DigiSal will focus on challenges of novel feedstuffs, collaborating with the Foods of Norway centre for research-based innovation at NBMU. Salmon are carnivores but today aquaculture provides more than half their fat and protein from plants, challenging the metabolic system and affecting fish health and nutritional value of salmon meat. The newly sequenced salmon genome and related resources will support a highly integrated theoretical-experimental study of mechanistic models of salmon metabolism, genetics, and feed factors.

Develop a systems biology framework for adapting salmon metabolism to modern feedstuffs, blazing the trail for a mechanistic framework for a model-based account of genetic and environmental variation in salmon metabolism

Travelling the systemic role of gut microbiota in adapting to new feeds
Provide and validate a theoretical framework for systematic identification of targets for steering EPA/DHA metabolism through concerted use of nutrition and genetics

Provide the foundation for a Digital Salmon knowledge base enabling adaption of a transformative pre-emptive research and development strategy

Digital part of the DigiSal life project by the BIOTEK2021 biotechnology programme of the Research Council of Norway.

Programme: The Digital Salmon

Public web page: <http://tinyurl.com/digisal>

Internal web page: Not specified

Organisms: *Salmo salar*, *Oncorhynchus mykiss*, *Danio rerio*

FAIRDOM PALS: No PALS for this Project

Project administrators: Jon Olav Vik

Asset housekeepers: No Asset housekeepers for this Project

Asset gatekeepers: No Asset gatekeepers for this Project

Related items

People (23) Institutions (4) Investigations (4) Studies (3) Assays (5) Data files (16) SOPs (5)
Publications (3)



Salmon feed-switch experiment vegetable ↔ fish oil 2015-2016

[Unsubscribe](#) [New Study based on this one](#) [Add an Assay](#) [Administration](#)

This experiment is designed to pinpoint where in the metabolic network there are differences between salmon of different genetic families and on different diets. Analyses of this material will help inform feeding and breeding strategies.

Salmon will be reared on feeds with contrasting levels of very-long-chain polyunsaturated fatty acids. Then some fish will be crossed over to the other diet while others remain as controls. This perturbation of diet should provoke changes in omega-3 metabolism that are evident in RNA expression, protein expression and eventually in the fatty acid composition of tissues. We also hypothesize that gut microbiota may be affected and/or modulate the effect of the feed switch.

Activities:

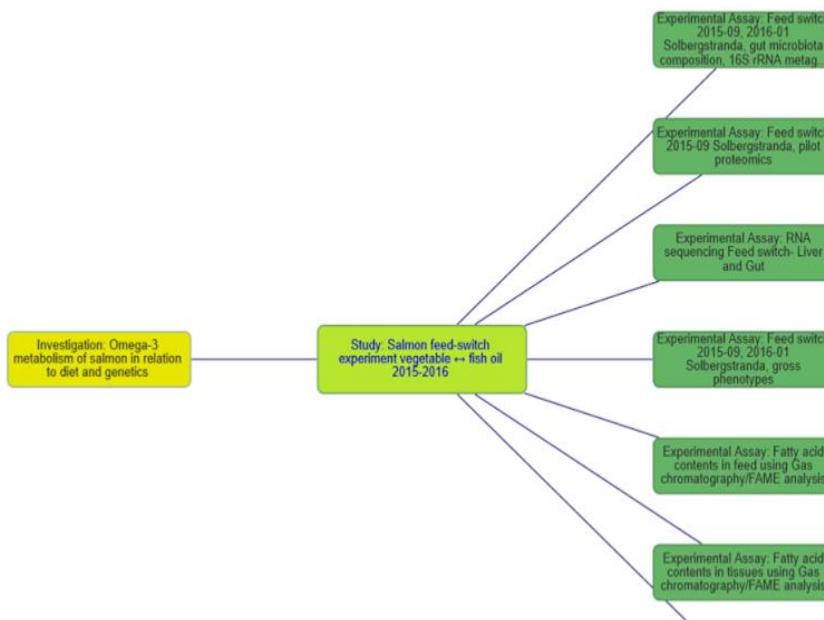
- Generate two fish groups that differ genetically and in EPA/DHA fillet content for feeding trial
- Carry out feeding trial to promote differences in FA metabolism and EPA/DHA phenotype
- Supply downstream WPs with tissue samples for RNA sequencing and proteomics with known FA phenotype.

Rationale: A dietary switch experiment in fresh- and seawater will provide data for a rich phenotypic characterization of the metabolic steady states under different diets. Although not directly relevant for aquaculture, we have chosen to include a diet switch at the parr stage (in fresh water) for two reasons:

Context: In hatchery rearing, it is hard to control exactly what happens to the fish during the transition from freshwater to seawater, with implications for growth and health.

Selected item: Study: Salmon feed-switch experiment vegetable ↔ fish oil 2015-2016

[Full graph](#)



Activity
Views: 219
Created: 2nd Feb 2016 at 11:25
Last updated: 6th Sep 2016 at 08:59

Programme

Overarching research theme ([The Digital Salmon](#))

Project

Research grant ([DigiSal](#), [GenoSysFat](#))

Investigation

A particular biological process, phenomenon or thing (typically corresponds to [plans for] one or more closely related papers)

Study

Experiment whose design reflects a specific biological research question

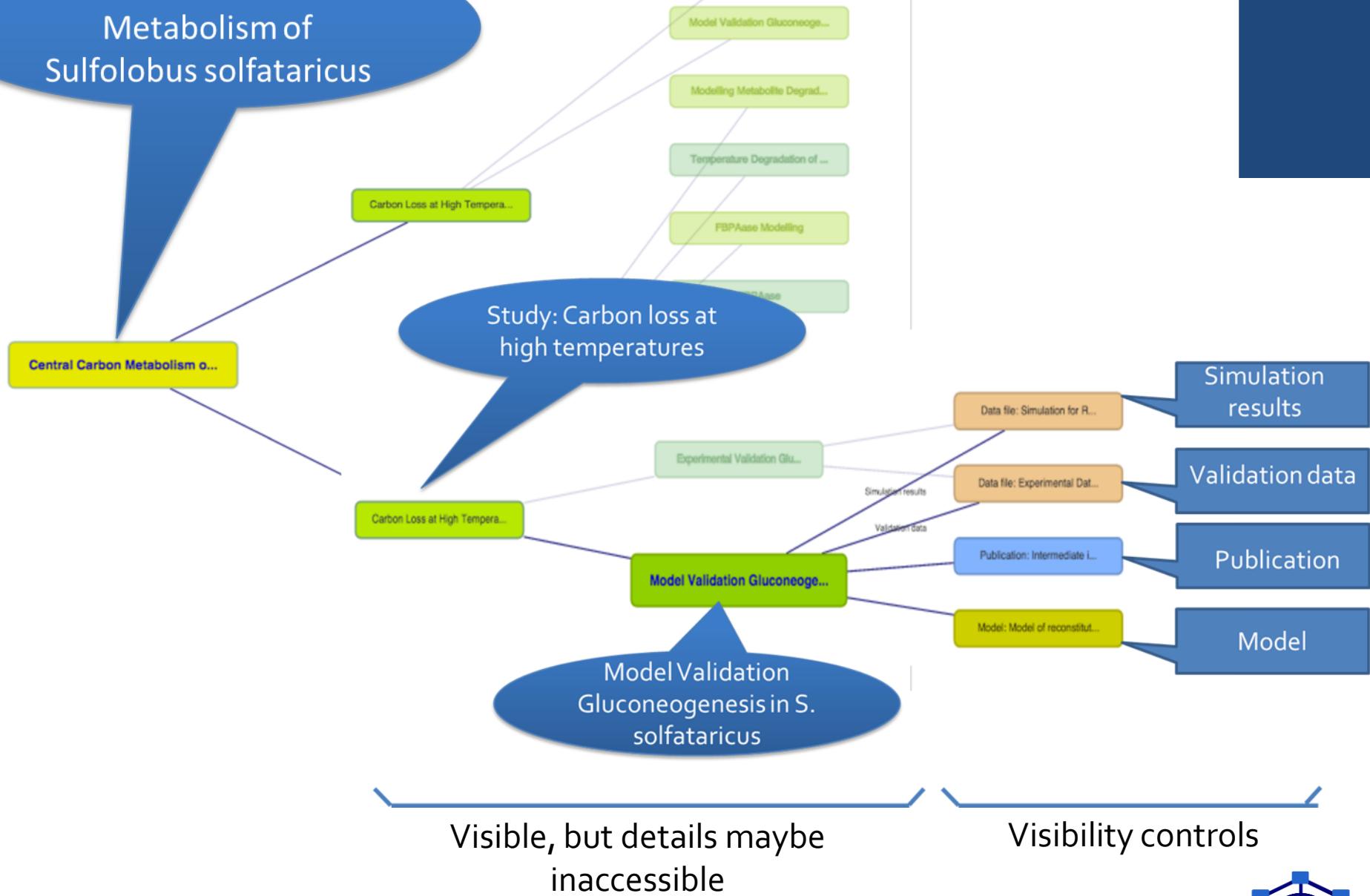
Assay

Standardized measurement or diagnostic experiment using a specific protocol

(applied to material from a study)

Jon Olav Vik,
Norwegian
University of
Life Science

Central Carbon Metabolism of Sulfolobus solfataricus



Kinetic analysis of metabolic system using transient perturbations

[Unsubscribe](#)[New Investigation based on this one](#)[Add a Study](#)[Administration ▾](#)Investigation of dynamics of t
conditions

ID:12

Projects: MOSES

Selected item: Investigation

Investigation: Kinetic analysis of
metabolic system using transient
perturbationsStudy: Metabolic perturbation of
the steady state culture by
glucose pulseExperimental Assay: Cellular size
and granularity during glucose
pulseExperimental Assay: Biomass
weight during glucose pulseExperimental Assay: Dynamics of
extracellular metabolites during
glucose pulseExperimental Assay: Dynamics of
intracellular metabolites during
glucose pulseExperimental Assay: Dynamics of
macromolecules during glucose
pulseModelling Analysis: MOSES:
dynamic model of glucose pulse

Publication

STUDY

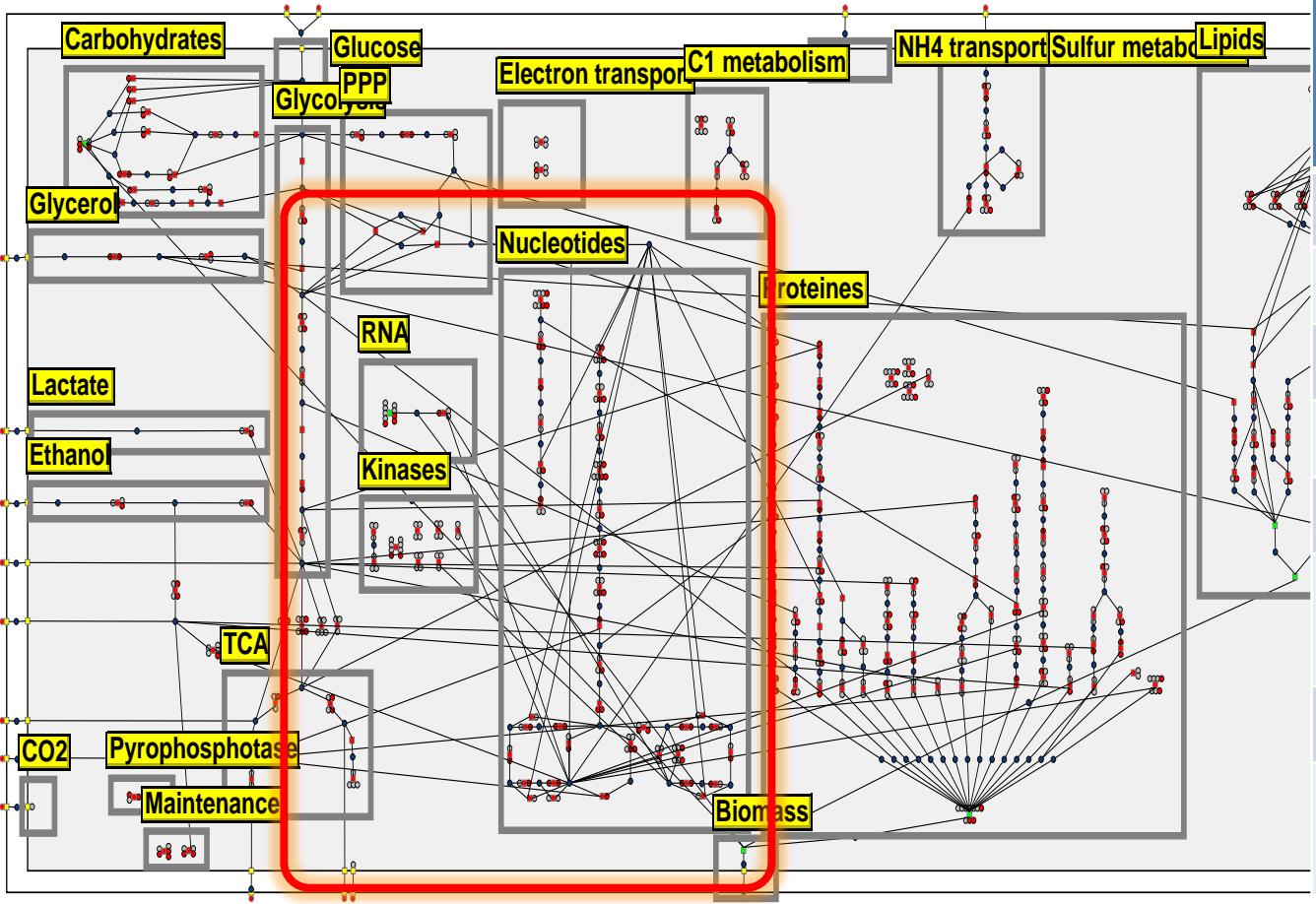
ASSAY

Experimental
assay

Modeling assay

INVESTIGATION

Stoichiometric model



Stoichiometric model of *Saccharomyces cerevisiae* growing anaerobically at $D=0.1 \text{ h}^{-1}$

Compartments	1
Pathways	35
Transformers:	254
reactions	214
transports	33
polymerizati on	7
Compounds	253
balanced comp.	234

Supplementary information



The screenshot shows the FluxAnalyzer software interface. On the left, there is a table titled "Stoichiometric Matrix" with columns for compartments A and B, and rows for metabolites. The metabolites listed include etoh_cell, NADH_cell, NAD_cell, acald_cell, NADH_cell, ADP_cell, ATP_cell, AMP_cell, gap_cell, dhap_cell, f16p_cell, IMP_cell, PRPP_cell, CO2_cell, pep_cell, dihydroxyacetone phosphate, and glycerate 3-phosphate. The matrix values show stoichiometric coefficients for each reaction. Below the table are two yellow folder icons labeled "SEEK".

In the center, a metabolic network diagram is displayed. Nodes represent metabolites and are labeled with their respective concentrations: etoh_cell (1.0), NADH_cell (1.0), NAD_cell (0.0), acald_cell (-1.0), ADP_cell (2.0), ATP_cell (-1.0), AMP_cell (-1.0), gap_cell (0.0), dhap_cell (0.0), f16p_cell (0.0), IMP_cell (0.0), PRPP_cell (0.0), CO2_cell (0.0), pep_cell (0.0), dihydroxyacetone phosphate (0.0), and glycerate 3-phosphate (0.0). Edges between nodes represent reactions, with some edges having numerical labels indicating flux values.

To the right of the network diagram, there are two windows for "Metabolic Flux Analysis". The top window shows a table of measured fluxes:

Reaction	Flux (mmol/gV*s)
T_ethyl	0.04599749500829
GLT	0.92417657318798
T_gly	0.139515753004979
#T_suc	0.022116887909356

The bottom window shows a "Standardize" dialog box stating "System is determined" and displaying the equation: $0.92^{glic} = 0.13^{gly} + 1.48^{etho} + 0.54^{ATP} + 0.64^{succ} + 0.64^{biomass} + 1.62^{f16p}$.

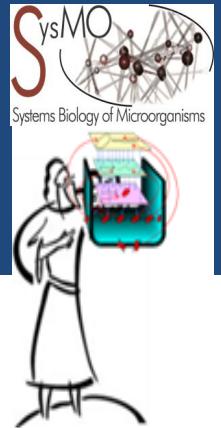
Annotation file

Stoichiometric matrix

SBML

Stationary fluxes





ODE-based model

example =
glc-permease

$$\frac{dC_{glc}^{ex}}{dt} = D \left(C_{glc}^0 - C_{glc}^{ex}(t) \right) - r_{Perm} \frac{C_x}{\rho} + a \frac{F_{pulse}}{V_r} C_{glc}^{pulse}$$

$$r_{perm} = r_{influx} - r_{efflux}$$

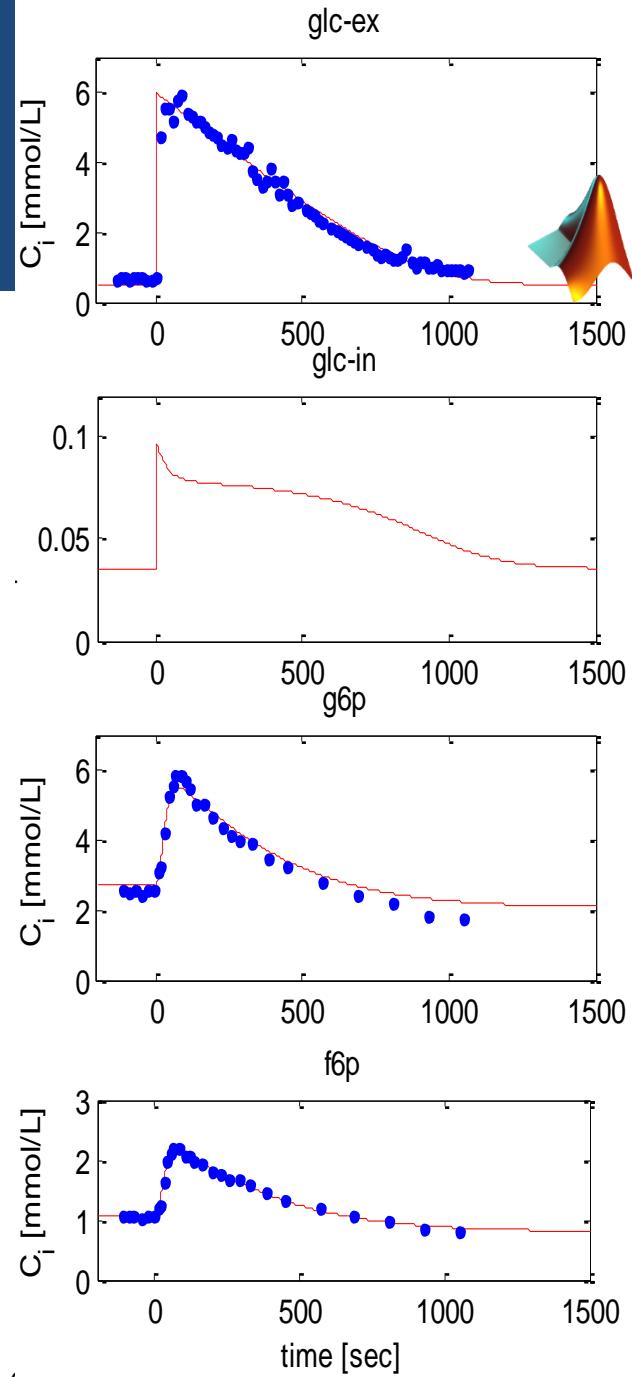
$$r_{influx} = \frac{R_{perm}^{\max} \frac{C_{glc}^{ex}}{K_{glc,1}}}{1 + \frac{C_{glc}^{ex}}{K_{glc,1}} + \frac{1 + \frac{C_{glc}^{ex}}{K_{glc,1}}}{1 + \frac{C_{glc}}{K_{glc,1}}} \left[1 + \frac{C_{glc}}{K_{glc,1}} + \frac{C_{g6p}}{K_{I-g6p,1}} + \frac{C_{glc} C_{g6p}}{K_{glc,1} K_{II-g6p,1}} \right]}$$

$$r_{efflux} = \frac{R_{perm}^{\max} \frac{C_{glc}}{K_{glc,1}}}{1 + \frac{C_{glc}}{K_{glc,1}} + \frac{1 + \frac{C_{glc}}{K_{glc,1}}}{1 + \frac{C_{glc}^{ex}}{K_{glc,1}}} \left[1 + \frac{C_{glc}^{ex}}{K_{glc,1}} + \frac{C_{g6p}}{K_{I-g6p,1}} + \frac{C_{glc} C_{g6p}}{K_{glc,1} K_{II-g6p,1}} \right]}$$

[Rizzi et al., 1997]



- variables: $C_{glc}^{ex}, C_{glc}, C_{g6p}$
- parameters: $R^{\max}, K_{glc}, K_{I-g6p}, K_{II-g6p}$

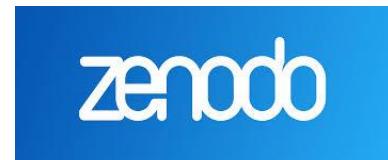


Model Publication

Publishing and data citation



The screenshot shows the FAIRDOM HUB interface. At the top, there's a circular navigation menu with nodes labeled 'HUB', 'FAIRDOM', 'IRDOM', 'CIBIO', 'Documentation', and 'Search here'. Below this, a search bar contains the query 'Glucose metabolism in Plasmodium falciparum trophozoites'. A main content area displays a summary of the project, mentioning 'The investigation entails the construction and validation of a detailed mathematical model for glycolysis of the malaria parasite Plasmodium falciparum in the blood stage trophozoite form.' and 'Project: Whole body modeling of glucose metabolism in malaria patients'. It includes a diagram showing 'Glucose metabolism in Plasmodium falciparum trophozoites' connected to 'Blood glucose' and 'Model construction'. Below this, a section titled 'Related Items' lists 'Models (13)' and a specific item: 'G3PDH Kinetic model' (Mathematica notebook). The URL at the bottom is <https://doi.org/10.15490/seek.1.investigation.56>.



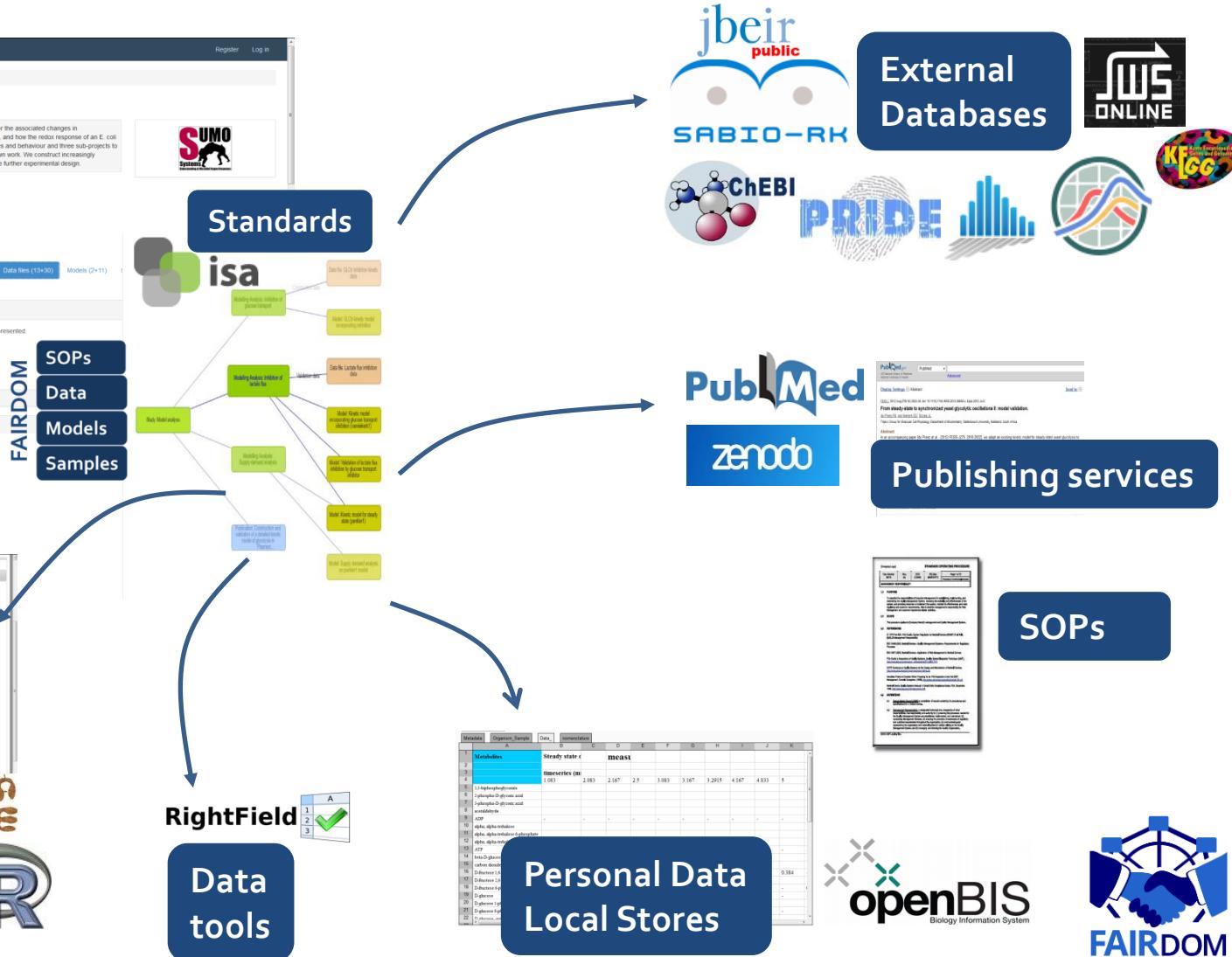
Institutional
Repository



FAIRDOM Catalogue, a web of resources

drawing together across resources; reusing tools and repositories
respect local project solutions, tool plugins

The screenshot shows the FAIRDOM catalogue interface. At the top, there's a search bar and navigation links for 'Browse', 'Help', and 'Log In'. Below the header, there are sections for 'Home', 'Projects index', and 'SUMO'. The 'SUMO' section contains a brief description of the Systems Understanding of Microbial Oxygen response (SUMO) project, mentioning Escherichia coli, Fnr and ArcA proteins, regulatory systems, and the redox response of an E. coli cell. It also notes that the project is part of the FAIRDOM project and has been developed by the same team. The 'Programme: SUMO' section includes a public web page link (<http://www.sumo.net/index.php?index=55>) and a GitHub link ([Organisms: Escherichia coli, Escherichia coli K-12](https://github.com/sumo-project/sumo)). The 'FAIRDOM PALS' section lists Sebastian Henke, Alireza Maleki-Dizaj, Maikel Verlicetti, David Koles. Below this, there's a 'Related Items' section with links to 'People (29)', 'Institutions (5)', 'Investigations (5+2)', 'Studies (7)', 'Assays (12+11)', 'Strains (22+1)', 'Data Sets (10+30)', 'Models (2+11)', and 'Presentations (0+4)'. A 'Gene expression rates at different aerobic levels via RT-PCR' section shows a plot of gene expression rates versus time. Another section shows 'Transcription of glucose transport systems under batch versus chemostat conditions'.



Repository Repertoire

Stanford et al The evolution of standards and data management practices in systems biology, Molecular Systems Biology (2015) 11: 851 DOI 10.15252/msb.20156053

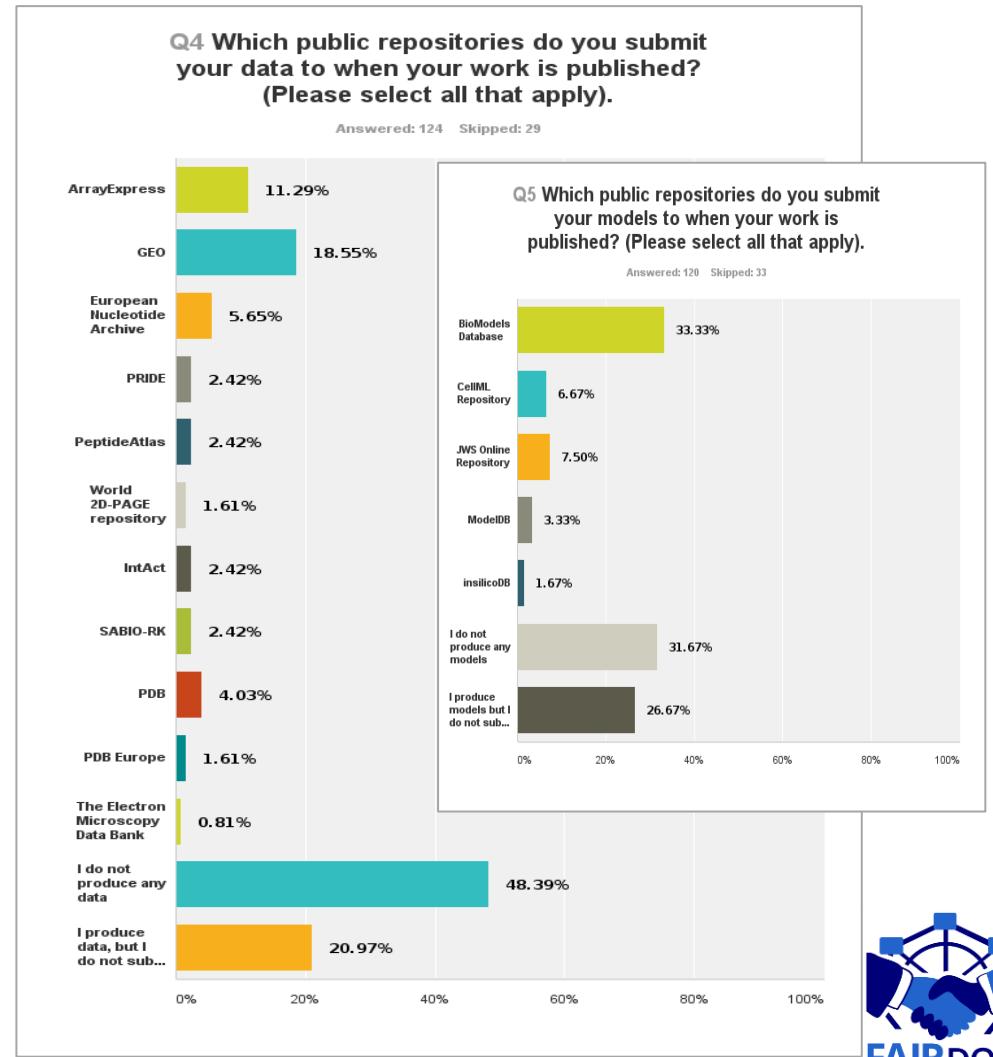
Specialist Public Repositories



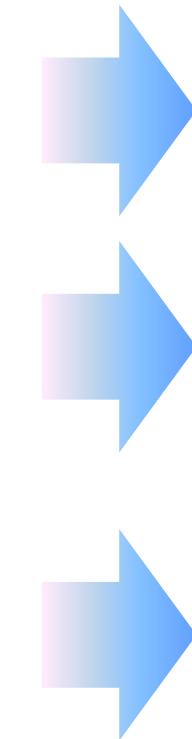
General archives



Local Data Stores



Laissez-Faire

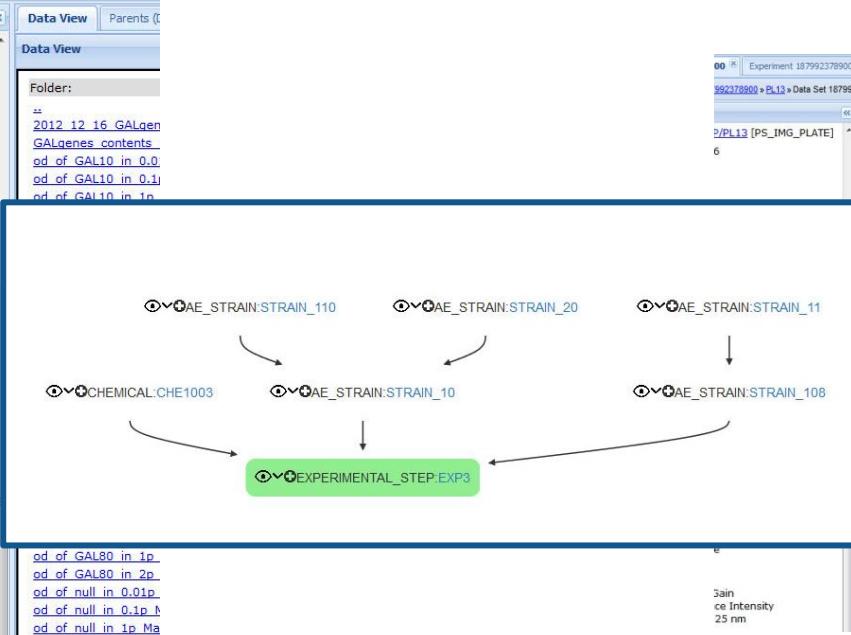


openBIS

metadata extraction
data relationship/linking
data processing
minimal user input

PS_GROUP > GROWTH RATE > 187992378900 > PL13 > Data Set 187992378900 [PS_PLATE_READER]

Data Set Properties	
Sample	/PS_GROUP/PL13 [PS_IMG_PLATE]
Measurement date	2012-12-16
Measurement time	19:20:40
.mth file	March2012_2EmWL_1gain_M2_80c.mth
Wsp file	2012_12_16.wsp
Instrument	infinite 200
Serial Nr	907001834
Plate desc.	[BD96ft_FluoroBlok] - BD Falcon 96 Flat Transparent/Black
Plate Cell Range	A1:H12
Temperature [C]	29.5
Min Temp. [C]	29.0
Max Temp. [C]	30.0
Shaking	Duration: 1000 sec; Mode: Linear; Amplitude: 6 mm; Frequency: 57.9 rpm;
Run Time	1days 45min 16s
Channels	OD 485-525mGain 485-585mGain
Channel1	OD Absorbance 595 nm reads:15



*Troup, E.; Clark, I; Swain, P; Millar, AJ; Zielinski, T (2015) Practical evaluation of SEEK and openBIS for biological data management in SynthSys <http://hdl.handle.net/1842/12236>

Modelling standards based in browser validation and simulation



Kinetic model incorporating glucose transport inhibition (vanniekerk1) - JWS Online Model Simulation

SBML Model simulation

vanniekerk1

vanniekerk1

Detail Download

Reactions Parameters Fixed species Initial values Functions and Rules

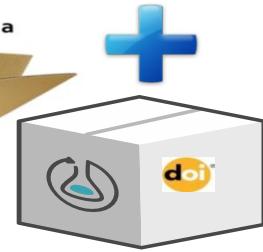
Time evolution Steady-state Parameter scan Reaction plots Information

Back to Model

SUCOME MORE

BiVes

A file that contains all the information needed to describe a modeling and simulation experiment.



Model name

Reactions

$v[v1] \{1.0\} \$s = \{1.0\} x_2$
 $v[v2] \{1.0\} x_2 = \{1.0\} x_3$
 $v[v3] \{1.0\} x_3 = \{1.0\} \p

Reactions help Show schema Hide schema

Zoom Move

vv1 vv2 vv3 x1 x2 x3 s p

Model comparison and versioning

FAIRDOM

[Stellenbosch,
Rostock]

Version History

Version 6 (latest) Created 25th Aug 2016 at 16:29 by Anton Chechushkov

Steady state from PD2 with cure and ROS oscillations

Version 5 Created 25th Aug 2016 at 16:11 by Anton Chechushkov

PD2 model with cure

Version 4 Created 25th Aug 2016 at 16:10 by Anton Chechushkov

Steady state from PD2 model and ROS oscillations

Version 3 Created 25th Aug 2016 at 16:10 by Anton Chechushkov

PD2 model

Version 2 Created 25th Aug 2016 at 16:09 by Anton Chechushkov

Steady state from basic model and ROS oscillations

Version 1 (earliest) Created 25th Aug 2016

No revision comments

SBML Differences

Both documents have same Level/Version: L2V4

Parameters

Kf_Apoptosis (Kf-Apoptosis)	Attribute value has changed: 0.724872855716881 → 0.197570659516782
ROSSynt (ROSSynt)	Attribute constant has changed: true → false

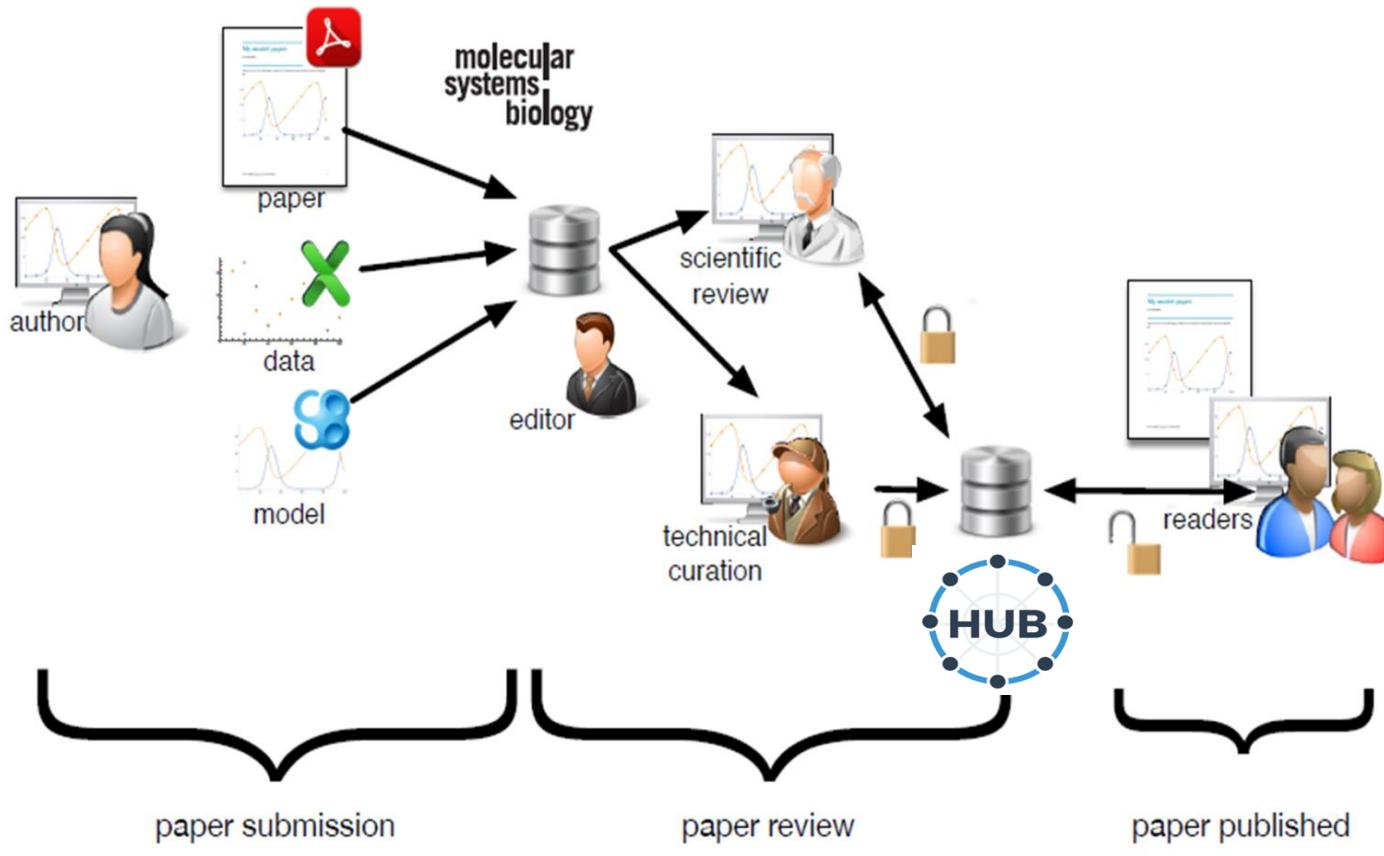
Species

s37 (Mtotal)	Attribute initialConcentration has changed: 71.9887518995473 → 27.0730581981201
CytC (CytC)	Attribute initialConcentration has changed: 516.494790148785 → 230.443022293088
s59 (AntioxidantsInactGene)	Attribute initialConcentration has changed: 0.520536587161077 → 0.678774240859421
s33 (p62cyt)	Attribute initialConcentration has changed: 586.300145609044 → 151.711834895516
s58 (p62actGene)	Attribute initialConcentration has changed: 0.985469109545648 → 0.640804278338365
s153 (KEAP1:p62)	Attribute initialConcentration has changed: 88.6458568095938 → 47.3278197938526

<http://seekdb.bionet.nsc.ru/>

**Alexey Kolodkin
Siberian Branch
Russian Academy of Sciences**

Reproducible model simulations in papers using COMBINE Archive & SED-ML



Three tiered service

★ store DOI citable supplementary files on FAIRDOMHub

★★ model and data curation

★★★ reproducible clickable figures in papers using SED-ML

COMBINE Archives support

The screenshot shows the FAIRDOM HUB interface. At the top, there is a navigation bar with the FAIRDOM HUB logo, search and browse buttons, and a help section. Below the navigation bar, the URL indicates the user is on the 'Assays Index' for 'GLC incubation'. The main content area features a large title 'GLC incubation' with a line graph icon. A green arrow points from the text 'SED-ML simulation: https://jjj.bio.vu.nl/models/experiments/penkler2aa_experiment-user/simulate' to the right. This text is part of a card-like element.

FAIRDOM HUB

Search here...

Browse ▾

Help ▾

Home / Assays Index / GLC incubation

GLC incubation

SED-ML simulation: https://jjj.bio.vu.nl/models/experiments/penkler2aa_experiment-user/simulate

ID: 260

Class: Modelling Analysis

Contributor: Dawie Van Niekerk

Projects: Whole body modelling of glucose metabolism in malaria patients

Investigation: Glucose metabolism in Plasmodium falciparum trophozoites

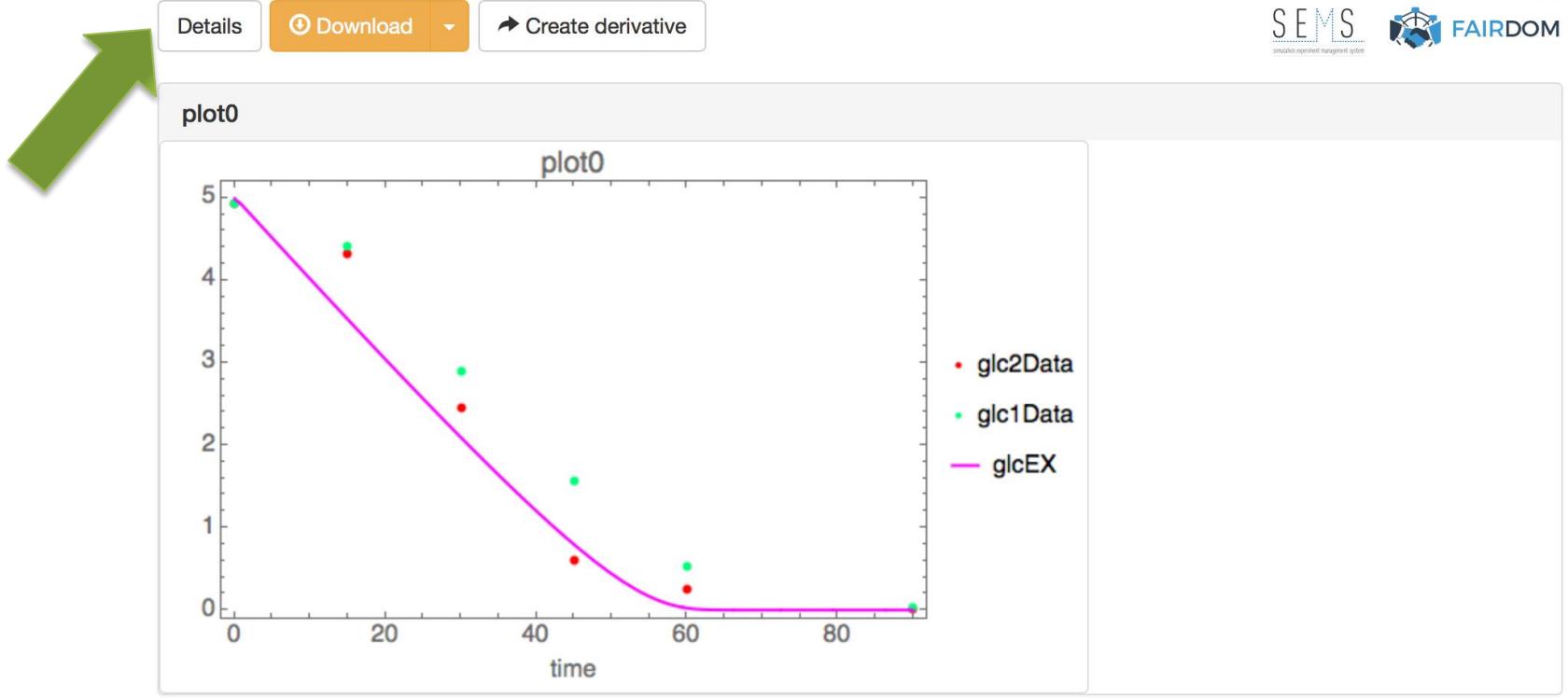
Martin Peters Talk Wednesday Morning



COMBINE Archives support

JWS Online Model Database ▾ Simulation Database ▾ Manuscript Database Help Servers ▾

SED-ML Simulation Result: penkler2aa_experiment



Martin Peters Talk Wednesday Morning



COMBINE Archives support

penkler2aa_experiment

▶ Simulate

⌚ Download

↗ Create derivative

powered by



Document

Manuscript information

Models

1

Data Description

1

+ Add Data Description



Id	Name	Source	Number Of Data Sources	
penkler_glc_data	penkler_glc_data	https://seek.sysmo-db.org/data_files/1159/download? version=1	9	

Simulation Tasks

1

Outputs

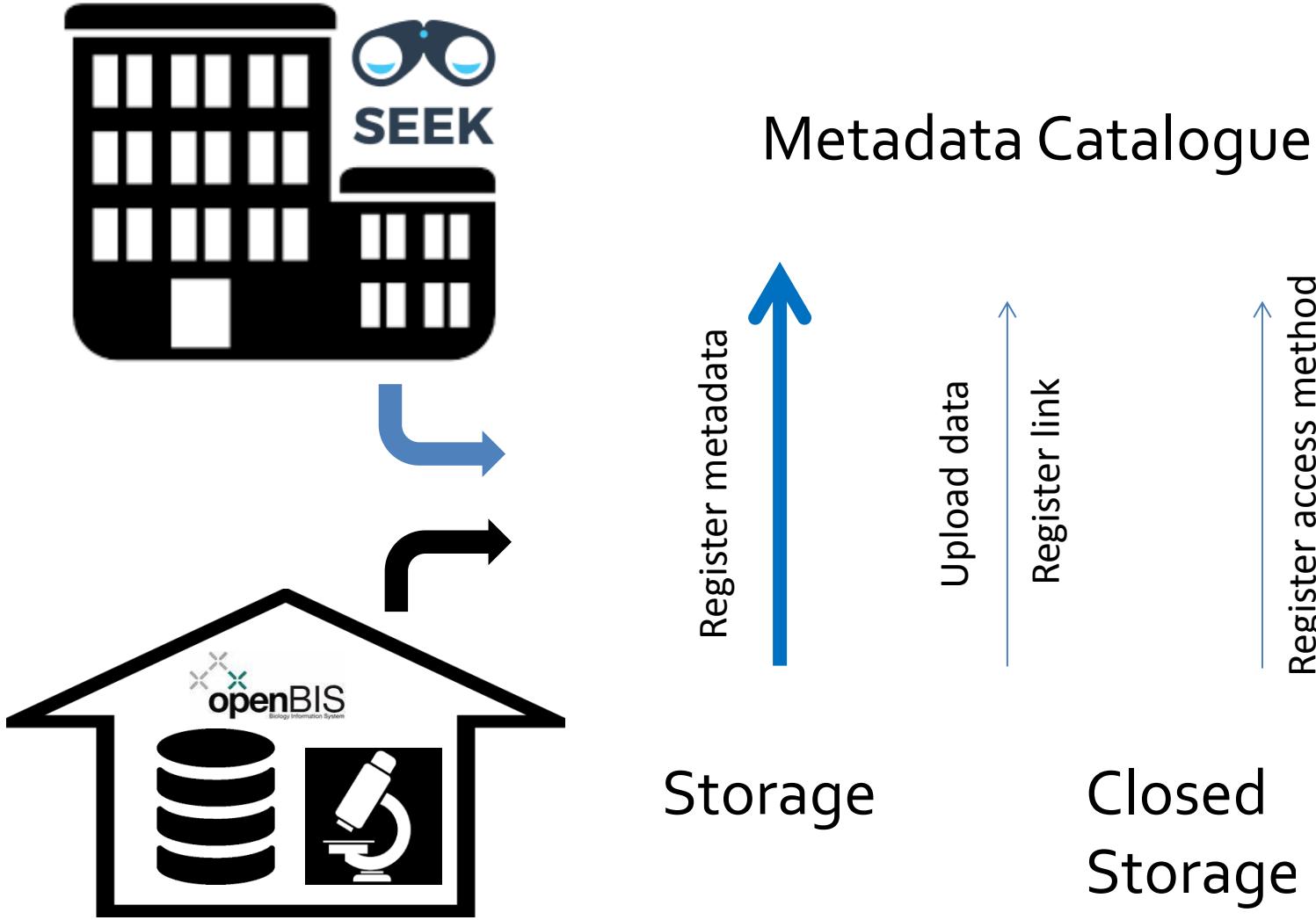
4 0

Martin Peters Talk Wednesday Morning



Standards-based Structured Description

templates, register, harvest, index, search



Hooking together templates, register, harvest, index, search

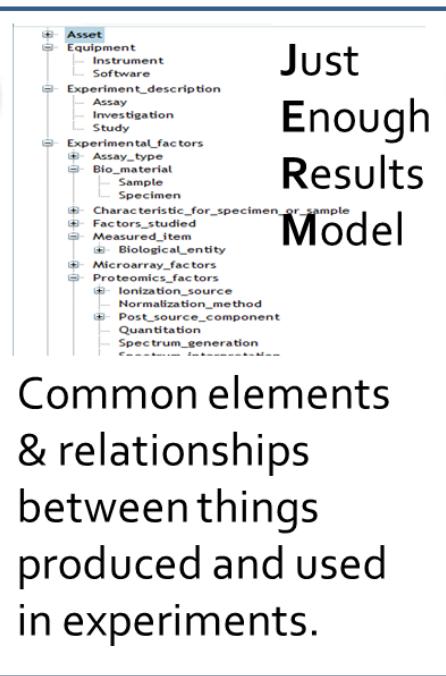
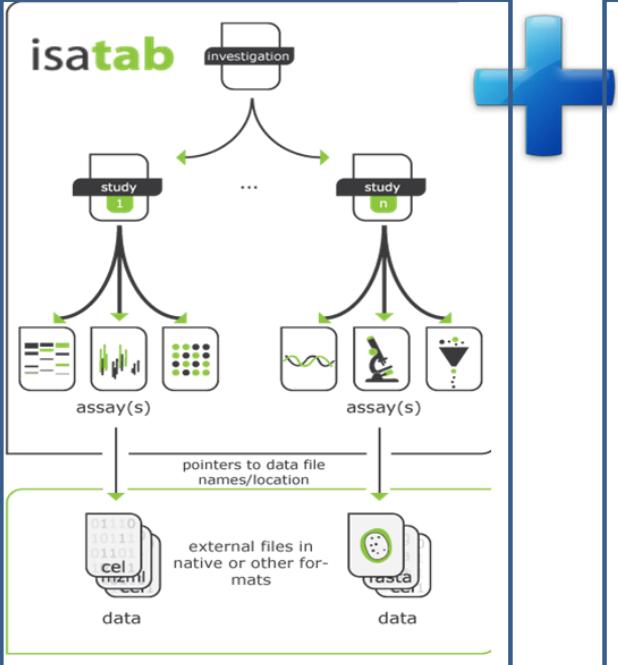


SPARQL endpoint release

Focus – joining stuff up!

- JSON read/write API
- ISA-TAB output





Specific elements
Community Standards

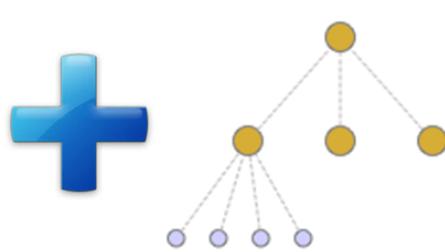
COMBINE
MIBBI
Etc....



RightField

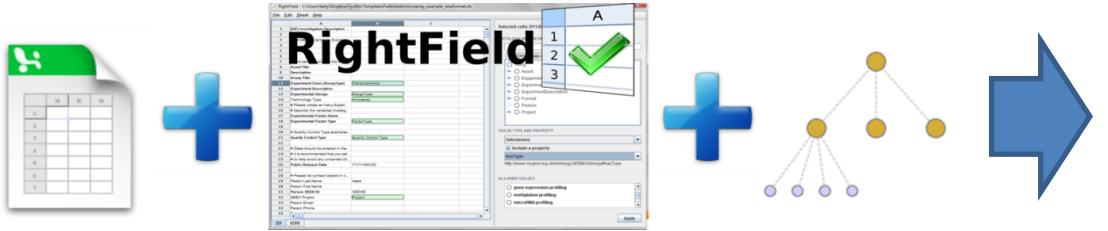
The screenshot shows a Microsoft Excel-like spreadsheet with several validation and metadata editing features overlaid:

- Validation Rules:** A sidebar on the left lists validation rules such as 'Experiment Class (Select type)', 'Assay Title', 'Experiment Design', etc.
- Value Type and Property:** A sidebar on the right shows 'Selected cell: B11B' with a dropdown menu for 'Value Type and Property' and 'Subclasses'.
- Allowed Values:** Another sidebar at the bottom shows 'ALLOWED VALUES' with options like 'gene expression profiling', 'metabolism profiling', and 'microRNA profiling'.



Spreadsheet tooling for metadata templates and metadata harvesting

Data annotation with standards



Embed ontologies into Excel templates

Excel spreadsheets enriched with ontology annotations

A screenshot of an Excel spreadsheet titled "IDFExcelExample2_tern.xls". The spreadsheet contains several rows of data with annotations. For example, in the first row, there is a note: "Transcription profiling of embryo and AFTS - Not NULL". In the second row, there is a note: "Experiments were conducted in the following conditions: Not NULL". The annotations are color-coded, with some in red and others in green.

A screenshot of an Excel spreadsheet titled "List Annotations". The table has columns labeled A, B, C, D, E, F, G, H, I, and J. The data includes annotations such as "% LAC flux (mea Error)" and "1.072". A large blue arrow points from this stage to the final outcome.

	A	B	C	D	E	F	G	H	I	J
1	GLC	cyt	% LAC flux (mea	Error						
2	mM	mM	1	1						
3	5	0	100	0						
4	5	1.5625	97.69837	1.072						
5	5	3.125	93.21616	4.103						
6	5	6.25	86.61292	4.743						
7	5	12.5	54.75364	3.622						
8	5	25	27.87788	6.512						
9	5	50	11.51952	0.877						
10	5	100	0	0						
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										
30										
31										
32										
33										
34										
35										

in browser viewing +
annotations

Upload, extract metadata
and register



Model annotation – in schema

seek.sysmo-db.org

JWS Online Webmail CYBERSMART Apple Amazon eBay Yahoo News

FAIRDOM HUB Search here... Search Jacky Snoep

Home / Models Index / Kinetic model for incubation (penkler2) / Simulate

Kinetic model for incubation (penkler2) - JWS Online Model Simulation Version 7

penkler2

glyEX
vPFvGLYtr

Reactions
Parameters
Fixed species
Initial values
Functions and Rules
Events
Constraints

Schema Time evolution Steady-state Parameter scan Reaction plots Information Documentation

Annotations Rate equation Structure Database query

HO-P-O
HO
O
HO
C00092

g3pPF vPFvG3PDI dhapPF vPFvTPI gapPF vPFvGAPDH b13pPF vPFvPGK p3gPF

glcEX vPFvGLCtr glcPF atpPF adpPF lacEX vPFvLACtr vPFvPYRtr vPFvLDH pyrPF vPFvATPASE vPFvPK pepPF vPFvENO p2gPF vPFvPGM

glyEX vPFvGLYtr vPFvHK vPFvPFK vPFvALD f16bpPF adpPF nadPF vPFvTPI gapPF vPFvGAPDH b13pPF vPFvPGK p3gPF

atpPF adpPF

Annotations Rate equation Structure Database query

Snap to grid Show wireframe Show modifiers Show compartments Gravity* Repulsion* Pool threshold* Species: Reactions:

0 Off On 4 Pin Unpin Pin Unpin

Back to Model

FAIRDOM

Model annotation – semanticSBML

Screenshot of a web browser showing the JWS Online Model Simulation interface for a "Kinetic model for incubation (penkler2)". The interface includes a sidebar with species lists and a central panel for annotating species.

The central panel displays the "Annotate Species" dialog box:

- MIRIAM** tab selected.
- Attribution** tab available.
- Annotations** section:
 - Three entries under "is":
 - `http://identifiers.org/3dmet/B01130`
 - `http://identifiers.org/kegg.compound/C00008`
 - `http://identifiers.org/pubchem.substance/3310`
 - Red trash can icons next to each entry.
- Search** section:
 - Text input: `adp`
 - Search button icon.
- Results** section:
 - Four entries:
 - `ADP`: `http://identifiers.org/cas/20398-34-9` with a green plus icon.
 - `ADP`: `http://identifiers.org/3dmet/B01130` with a green plus icon.
 - `ADP`: `http://identifiers.org/obo.chebi/CHEBI%3A16761` with a green plus icon.
 - `ADP`: `http://identifiers.org/kegg.compound/C00008` with a green plus icon.





Annotation and Harvesting by Stealth

Samples Club samples framework

Samples

Show 5 entries

Sample name	title	raw data file 1	raw data file 2	ch1: source name	ch1: organism
Sample 1	RAG H3K4me2	RAG_H3K4me2_Cy5.pair	RAG_H3K4me2_Cy3.pair	H3K4me2 ChIP DNA from Rag2-/- pro-B cells	Mus musculus
Sample 2	RAG H3K4me3	RAG_H3K4me3_Cy5.pair	RAG_H3K4me3_Cy3.pair	H3K4me3 ChIP DNA from H3K9ac ChIP DNA from Rag2-/- pro-B cells	Mus musculus
Sample 3	RAG H3K9ac	RAG_H3K9ac_Cy5.pair	RAG_H3K9ac_Cy3.pair	H3K9ac ChIP DNA from Pax5-/- Rag2-/- pro-B cells	Mus musculus
Sample 6	RAG H3K9ac	RAG_H3K9ac_Cy5.pair	RAG_H3K9ac_Cy3.pair	H3K9ac ChIP DNA from Rag2-/- pro-B cells	Mus musculus

Search Sample name

RAG

Search raw data file 1

Search raw data file 2

Search ch1:

Search ch1: organism

User defined sample models

Sample type
defines a sharable
standard

Tied to assay processes

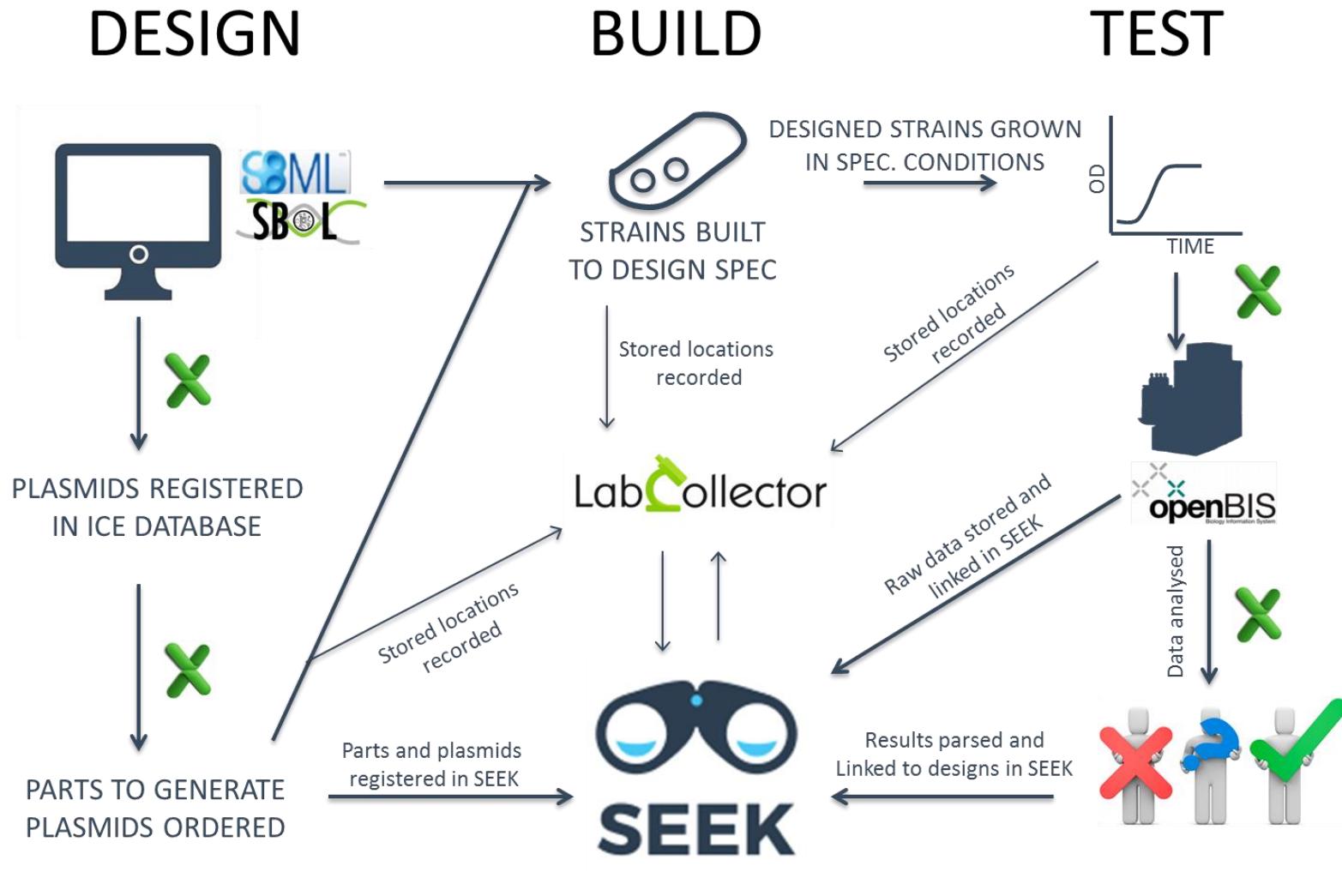
Template tooling
Auto extraction

Interlinking between sample types

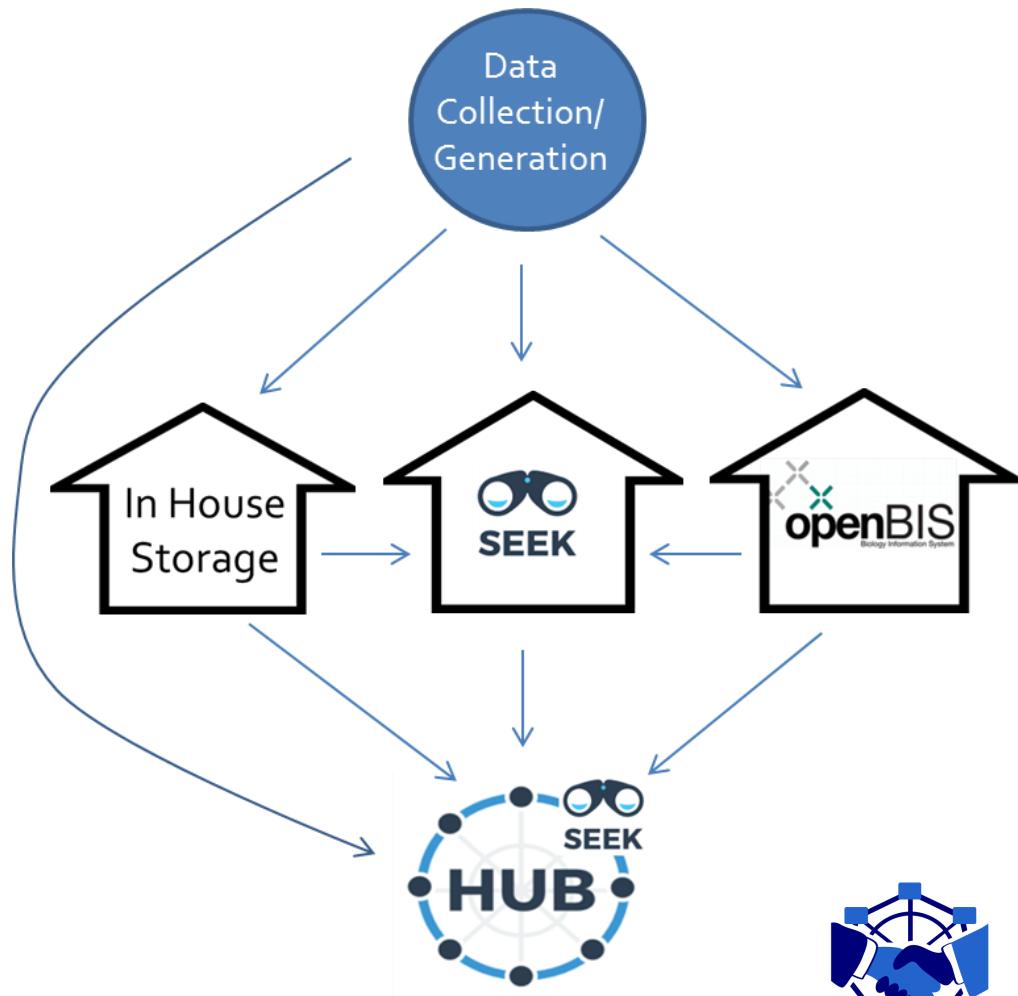
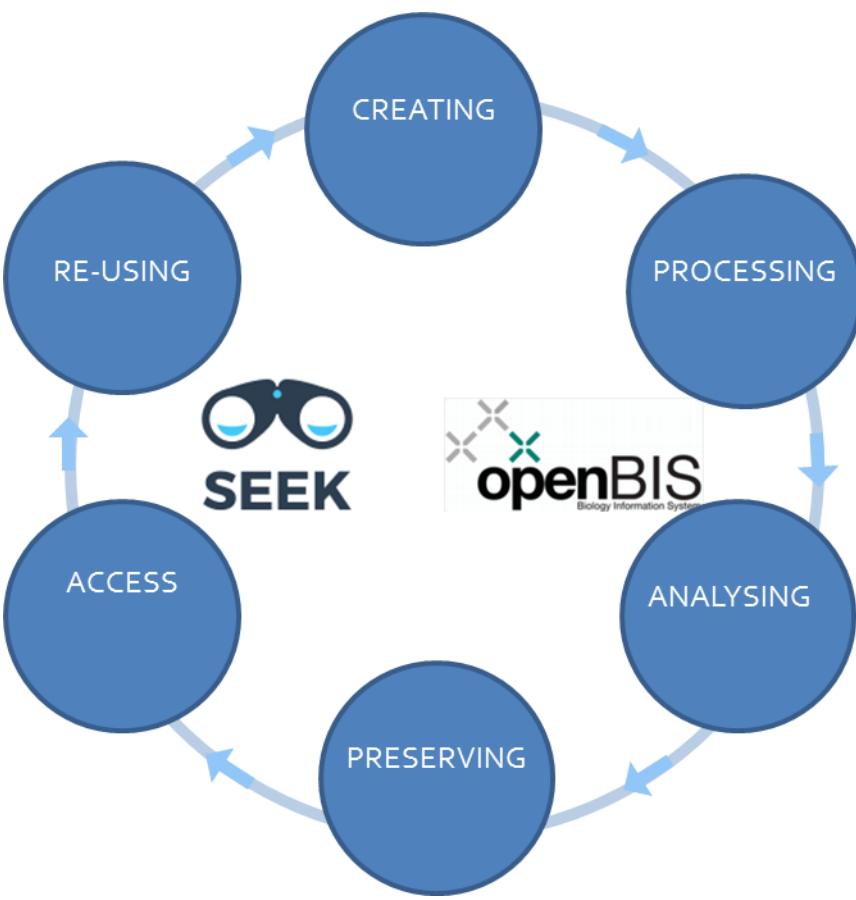


Spotlight: Synthetic Biology

SynBioChem Centre, UK



FAIRDOM Workflows and Pathways



Project Support – Understanding Process

Planning, Setups, Curation, Advice, Support



Community support



PALs project
ambassadors
co-design, tailoring,
communication,
requirements, review

model technical curation
with our JWS Online partners
standard, best practices

Special project support



More Projects and Centres

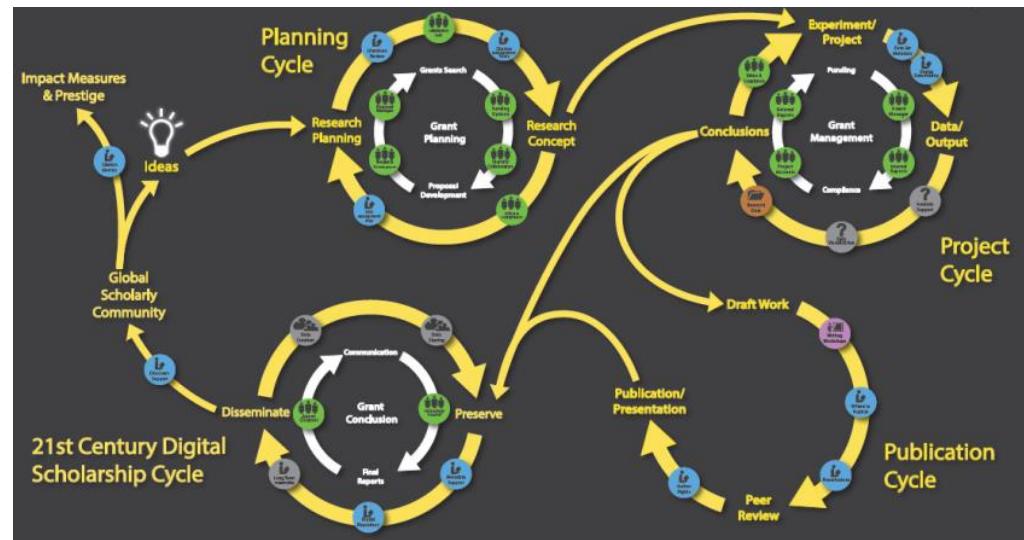


User Meeting, Barcelona@ICSB 2016



Processes and People

- 80% process, 20% tech
- Structuring the ISA
- PIs
- Sticking to conventions and policies
- Tension with standards take-up vs laissez faire
- Time and resource
- Local responsibility
- Recognition
- Institutional Repositories
- Automagic



The only fair
is laissez-faire.

FAIR Play

- Keep this Data file private (only visible to you)

Or share it with..

- Members of Projects associated with this Data file: View summary and get contents ▾

... and all other registered users: View summary only ▾

- All visitors (including anonymous visitors with no login): View summary and get contents ▾

- Licenses
- Negotiated access
- Embargos
- Permission controls
- Staged sharing
- Private walled gardens



Jurgen Hannstra
Vrije Universiteit Amsterdam,
Netherlands

Using FAIRDOM my own lab colleagues saw what I was doing and called to collaborate!

FAIR Play

- Drivers
 - External dominate
 - Personal productivity
- Trading behaviours
 - Tribal based
 - Modellers vs Experimentalists
- “enclave” sharing
 - Rather than public donation
- Reciprocity & credit
 - Citation

Select citing preferences:

Who can cite/reference this data-set now?

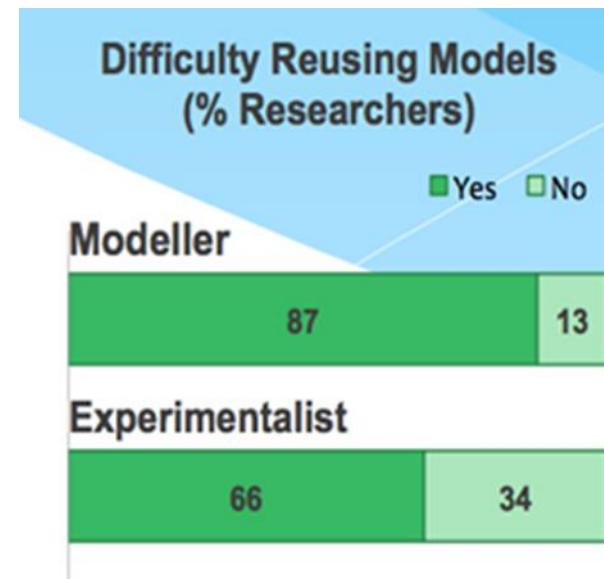
Remember that ... Publicly available data is significantly associated with a 69% increase in citations. (Piwowar, Day, and Fridsma 2007)	Public can cite	People involved in my projects can reference	Collaborators can reference	Nobody, Keep it Private
---	-----------------	--	-----------------------------	-------------------------

When should this data-set be accessible to the Public for citation?

You know
77% of your collaborators make their materials citable within 3 years.

in 6 Months (2016-07-12)	in 1 Year (2017-01-13)	in 3 Years (2019-01-13)	in 10 Years (2026-01-11)
-----------------------------	---------------------------	----------------------------	-----------------------------

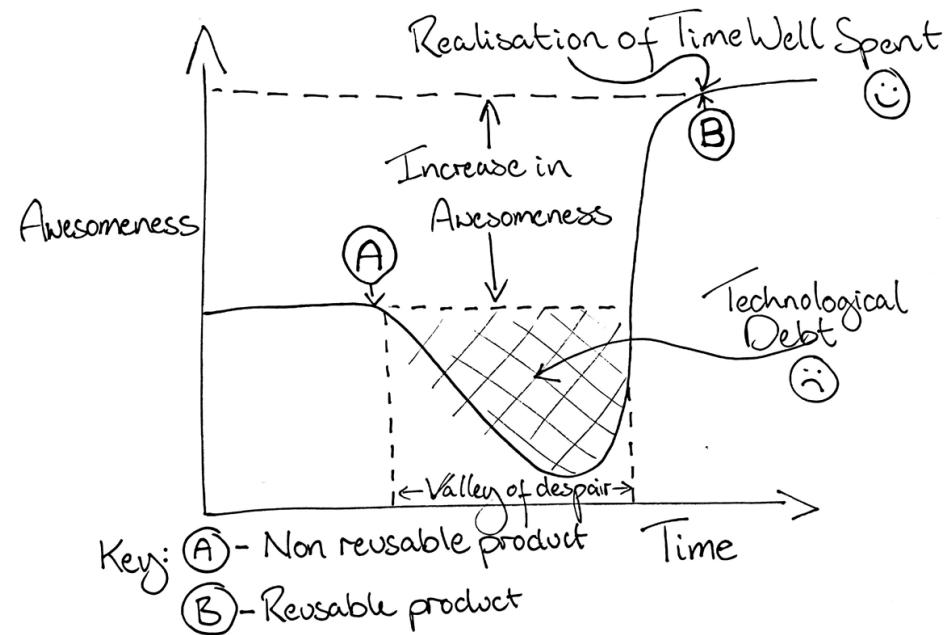
affecting behavioural change through libertarian paternalism*



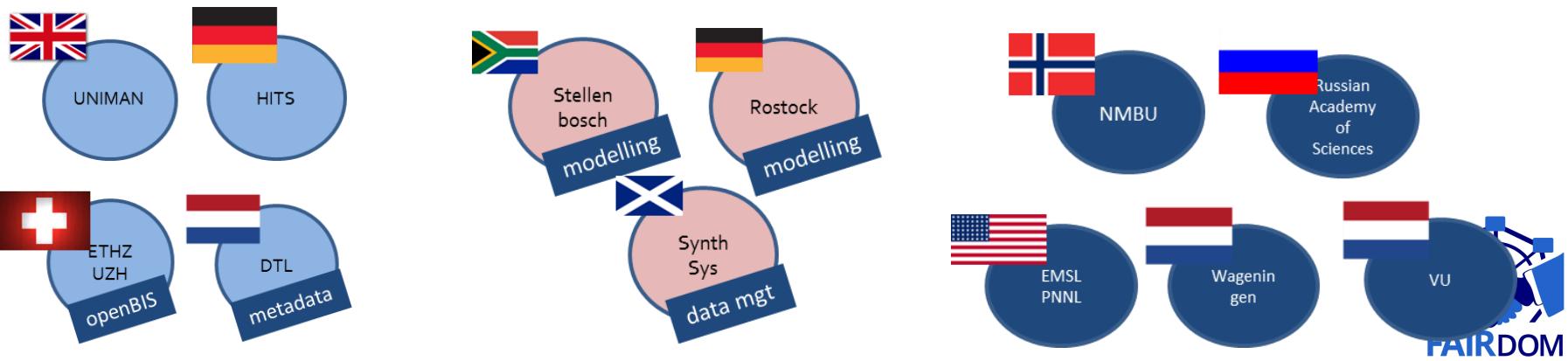
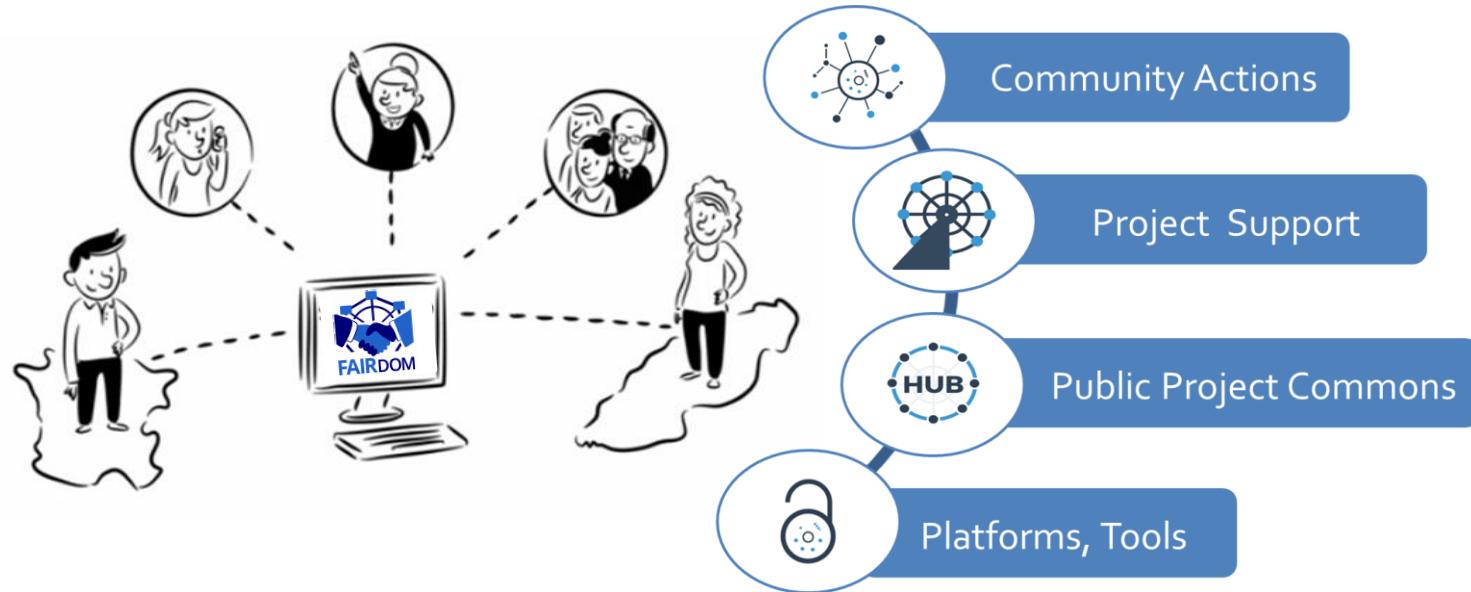
Stanford et al The evolution of standards and data management practices in systems biology, Molecular Systems Biology (2015) 11: 851 DOI 10.15252/msb.20156053



*Garza et al Framing the Community Data System Interface, <https://dx.doi.org/10.6084/m9.figshare.1300051.v5>



FAIRDOM Association e.V.





Carole Goble



Natalie Stanford



Stuart Owen



Finn Bacall



Jacky Snoep

Jon Olav Vik,
Norwegian University
of Life Science



Wolfgang
Mueller



Olga Krebs



Quyen Nguyen



Standards
Martin Golebiewski



Andrew Millar



Norman Morrison



Bernd Rinn



Lars
Malmstroem



Rostyslav
Kuzyakiv



Jakub
Straszewski



Caterina
Barillari



Alexey Kolodkin
Siberian Branch
Russian Academy of Sciences



Jacky Snoep



Dawie van Niekerk



Katy Wolstencroft



Dagmar Waltemath

Tomasz Zieliński,
SynthSys Centre
University Edinburgh



Adrian Pugh



Christian Ruckert



Petra Schulte



Daniel Von der Muehll