Data Management for Systems Biology Projects

Maja Rey

Tools for systems biology modeling and data exchange, Heidelberg, March 20th, 2019
Why Data Management?
Systems Biology Cycle

1. Hypothesis
2. Experiments
3. Data Generation
4. Data Analysis
5. Biological Insight
6. Model Construction
7. Model Validation
8. Model Analysis
Purpose of project data management

- Organisation
  - Add standardized metadata
  - Find and reuse later
  - Enables team to reuse
- Communication
  - Helps others find out
  - Reuse with new partners
- Dissemination
  - Tell more and take credit

of data associated with a project
How to Make More Published Research True

John P. A. Ioannidis

1 Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, United States of America, 2 Department of Prevention Research Center, Stanford, California, United States of America, 3 Department of Health Research and Policy, Stanford University School of California, United States of America, 4 Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California, United States of America

HEALTH SCIENCE

Accessible Reproducible Research

As use of computation in research grows, new tools are needed to expand recording, reporting, and reproduction of methods and data.

SCIENTIFIC AMERICAN OCTOBER 2018

Demand of open, reproducible, long-term maintained „FAIR“ data

Science under Scrutiny: The Problem of Reproducibility

Better incentives could reduce the alarming number of studies that turn out to be wrong

By Shannon Palus

H2020 Programme

Guidelines on FAIR Data Management in Horizon 2020

Mandatory Data Management Plans

Katie Corker wondered what temperature the coffee was supposed to be. In a psychology experiment—well, redoing an experiment—findings, suggesting that holding something warm can make you feel warmly, had been published in 2008 in the prestigious journal Science to a flurry of media coverage. Yet as Corker tried to document the findings and then to embed them directly into
SCIENTIFIC DATA

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

Mark D. Wilkinson et al.

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 measured 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 exemplary implementations in the community.

Box 2 | The FAIR Guiding Principles

To be Findable:
F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:
A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

To be Interoperable:
I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

To be Reusable:
R1. (meta)data are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards
FAIR DOM

- Findable
- Accessible
- Interoperable
- Reusable

- Data
- Operations
- Models

fair-dom.org
FAIRDOM Platform Installations

https://fairdomhub.org/

https://seek4science.org/

LiSyM - Liver Systems Medicine

NTM - New Medicines for Trypanosomatid Infections

FungiNet

ERASysAPP - ERA-NCP for Applied Systems Biology

Effect Network in Water Research

de.NBI - German Network for Bioinformatics Infrastructure

Heidelberg Institute for Theoretical Studies

HITS
https://fairdomhub.org
Status 03/2019: 133 Projects, 1215 People, 2167 Data files, 451 Models
ISA provides a directory structure for experiments and models.
Investigation: 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: 55

Projects: Whole body modeling of glucose metabolism in malaria patients

Related Items:
- People (1)
- Projects (1)
- Studies (3)
- Assays (24)
- Data files (16)
- Models (19)
- SOPs (13)
- Publications (1)

David Van Nielen

Discipline: Modeller
Role: not specified
Expertise: not specified
Tools: not specified

Investigation: Glucose metabolism in Plasmodium falciparum trophozoites

Study:
- Study Model construction
- Study Model validation
- Study Model analysis

Assay:
- Modelling Analysis: FFK

Model:
- Model: FFK Kinetic model

SOP:
- SOP: FFK Kinetic characterisation

PFK Kinetic model

Specific activity of the glycokinase enzymes was determined in falciparum GS40BL strain expressing vector pGEX-1 with recombinant bacterial protein in E. coli BL21 (DE3). The values were 172 μmol/min/mg (184 μmol/min/mg) of PFK activity for Plasmodium falciparum strain H3/GR, respectively. The results showed a significant change in the enzyme activity at different stages of the parasites in the presence of PFK (2 μM) from PFK activity in control conditions. The enzyme activity was measured by a spectrophotometer at 340 nm (ε = 22,600 M^(-1) cm^(-1)).

Selected Item: Model: FFK Kinetic model

https://fairdomhub.org/investigations/56
ISA graph versus ISA tree representation

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.

SEEK ID: https://sandbox4.fairdomhub.org/investigations/15
Projects: de.NBI Training course for data management
FAIRDOM Catalogue

Standards

External Databases

Data tools

Experimental Data

Local Stores

Publications

SOPs

Models

Modelling tools

Publishing
Data Sharing in SEEK

PGI Kinetic model Version 2

Mathematica notebook for the parameterisation of the PGI rate equation based on SEEK linked ex

SEEK ID: https://sandbox4.fairdomhub.org/models/10?version=2

1 item (and an image) are associated with this Model:
- PGI-SEEK.nb (Mathematica Notebook - 84.8 KB)

Organism: Plasmodium falciparum

Here you can specify who can view the summary of, get access to the content of, and edit the Model.

- Public: No Access
- de.NBI Training course for data management: View, Download
- de.NBI Training course for data management @ Heidelberg Institute for Theoretical Studies: View, Download

Other creators:
Dawie Van Niekerk, Jacky Snoep
Making an Investigation, Study or Assay citable

Here you can specify who can view the summary of and edit the Study.

<table>
<thead>
<tr>
<th></th>
<th>No Access</th>
<th>View</th>
<th>Edit</th>
<th>Manage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Example use cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Share with a person  Share with a project/institution  Share with a programme

FAIRDOM TRAINING

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.
Make a snapshot
Assigning a DOI

Oscillation studies in Saccharomyces cerevisiae glycolysis

No description specified

DOI: None

Created at: 23rd Apr 2018 at 10:39

Contents

Snapshots

Snapshot 1 (23rd Apr 2018)

Activity

Views: 169 Downloads: 4
Created: 23rd Apr 2018 at 10:39
(Since September 2018)

Fingerprints

These checksums allow you to check a Snapshot you have downloaded hasn’t been modified. For details on how to use these please visit this guide

Confirmation

You are about to mint a DOI for the following Study snapshot. Please review the contents of the snapshot below and then click the “Mint DOI” button to continue, or “Cancel” to go back.

Please note:
Once a snapshot has been given a DOI, it can no longer be deleted.

Contents

Mint DOI  Cancel
FAIRDOM helps with

Standards compliance
- Systems and Synthetic Biology standards support
- Support in finding adequate standards

Consistent reporting
- Structured using ISA
- Our specially made Just Enough Results Model (due to JERM ontology)

Metadata curation
- Spreadsheet templates for omics data and samples
- Data and model annotation tools

Integration with existing systems
- Integrated tools for modelling, ELNs, LIMS
- REST API to plugin to your systems

Export
- Package and export into other repositories
- Export into other FAIRDOM installations
- COMBINE Archive export
de.NBI-SysBio serving the Systems Biology Cycle

Modelling
- Model Construction
- Model Validation
- Model Analysis

Data Generation
- Data Analysis

Experiments
- Experiments
- RightField

Hypotheses

Biological Insight

FAIRDOM Hub

Seek

SABIO-RK

COPASI

SedML

O-Net Analyzer

SeMS

Heidelberg Institute for Theoretical Studies
Registering in FAIRDOMHub

Signup for FAIRDOMHub

Login name

Login should contain a minimum of 3 characters.

Email address

Password

Password should contain a minimum of 4 characters.

Confirm Password

By registering, you are indicating that you have read and agree that: You are responsible for any content that you upload. We receive and store certain types of information whenever you interact with us which we will use only for operational purposes. We will not make personally identifiable data available to third parties without consent. Please contact us if you have any queries about our terms or our privacy policy.

If you have previously provided a login and password but didn’t select or create a profile, then your details have already been stored and you should Login to continue the registration process.

Register
Registering in SEEK

In the next few minutes you should receive an email address containing a link which you must click on to activate your account. This is to ensure that you are really who you say you are.

Please contact one of the SEEK administrators if you do not receive the email, or you do receive it and it contains incorrect information.
HITS SDBV group @ Heidelberg, Germany

Thanks for your attention!