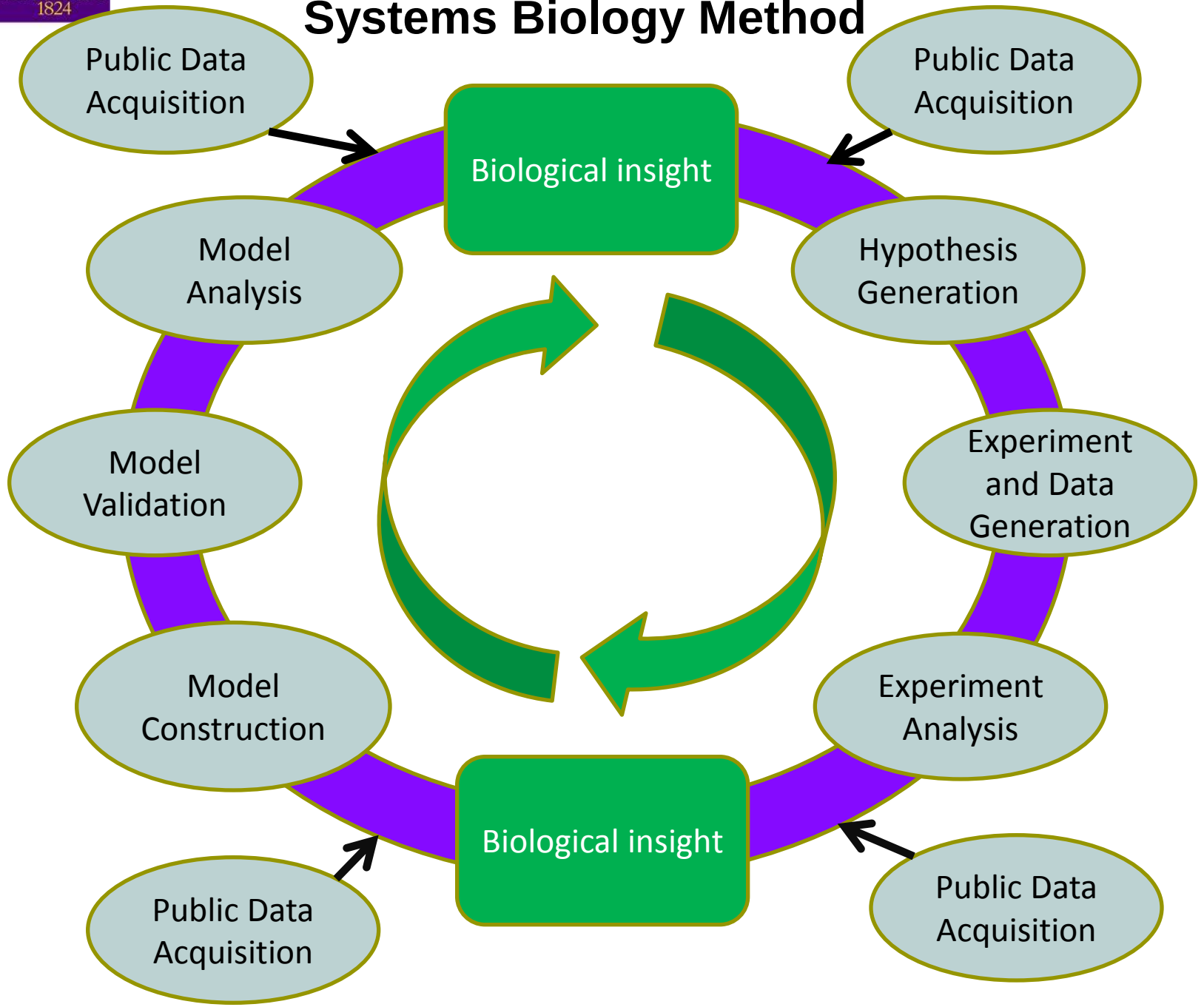
SysMO-DB and Virtual Liver SEEK: Sharing and Managing Systems Biology Experiments

Katy Wolstencroft
University of Manchester
Vrije Universiteit, Amsterdam
Martin Golebiewski
HITS

Systems Biology Method



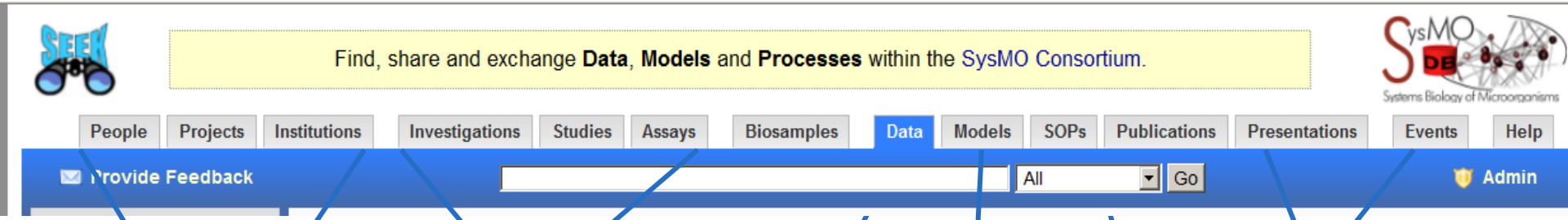
MANCHESTER 1824

The University of Manchester

SEEK: Systems Biology Data Sharing

- A platform for sharing Systems Biology data, models and protocols in the context of systems biology experiments
- Web based environment for sharing within a consortium, disseminating to the community, and exploring data and models (an e-Laboratory)
- Standards compliant and standards promoting
- Fitting in with laboratory practices

Assets Catalogue. Archive. Social Network. Sharing Space.



Yellow Pages:
- People
- Expertise
- Projects
- Institutions

I-S-A structure:
- Linking data, models, SOPs

Gateway to BioPortal

Models:
- Store
- Simulate with JWS Online
- Annotate

Gateway to COPASI, JWS Online, BioModels

Communication
- Publications
- Community

Gateway to PubMed

Data:
- Experimental data sets and analysed results

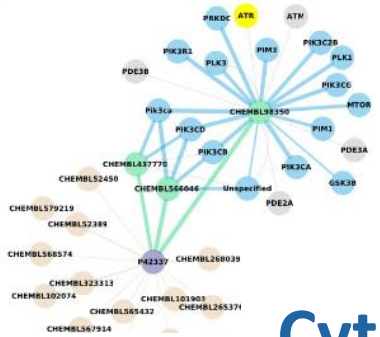
Gateway to public data stores – SABIO-RK, 'omics

Processes
- Standard Operating Procedures
- Computational workflows - Taverna

Gateway to myExperiment

Versioning/Sharing

Exploring and Analysing: The SEEK as a Gateway



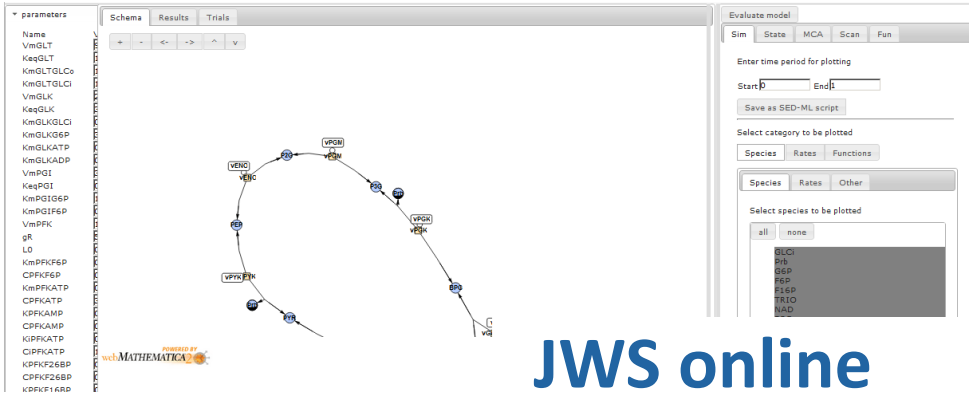
Cytoscape

Visualise Model with Cytoscape Web

Download Model

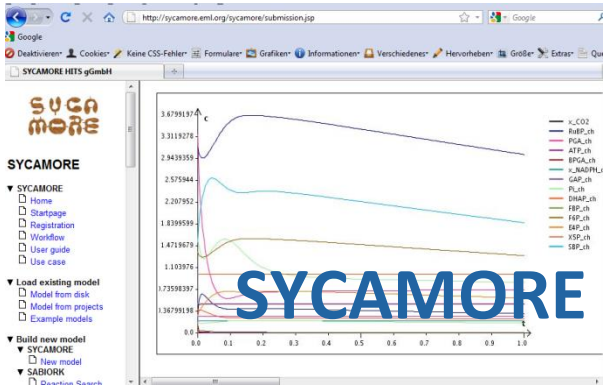
Simulate Model on JWS

Simulate Model on Sycamore



JWS online

Explore spreadsheets



SYCAMORE

SEEK Communities

- Pan European collaboration
- 13 individual projects, >100 institutes
- Model dynamic molecular processes in microorganisms
- Pool research capacities and know-how
- Disseminate activities of the consortium

- German research network
- ~45 organisations, ~70 groups
- Multiscale representation of the liver
- Pool research capacities and know-how
- Disseminate activities of the consortium

Browse & Search by Scales

Find, share and exchange **Data**, **Models** and **Processes** within the Virtual Liver Network.

Yellow Pages | Experiments | Biosamples | Assets | Forums | Events | Help

Provide Feedback | All | All | Go | Admin

New or upload

Data file

Go

Announcements

SEEK will be presented at ICSB 2012 and COMBINE / New SEEK version 4 months ago by Martin Golebiewski

Upload your data to SEEK as email attachment 5 months ago by Martin Golebiewski

New Virtual Liver SEEK with extended functionality 5 months ago by Martin Golebiewski

See all

Favourites

Tags [show all]

CellNetAnalyzer (CNA) data management Flow Cytometry HepatoSys IL-6 Insulin

Search Items by Scale

all -

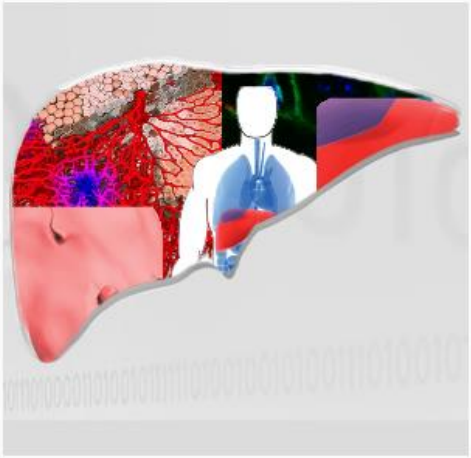
organism -

liver -

liver lobule -

intercellular -

cell -



Investigations (15+10) | Studies (11+13) | Assays (18+12) | Samples (29+5)

Specimens (33+9) | Data Files (263+131) | Models (59+13) | SOPs (33+1) | Publications (192) | Presentations (443+4) | Events (57+5)

Scales: Organism

Find, share and exchange **Data**, **Models** and **Processes** within the Virtual Liver Network.

Yellow Pages | Experiments | Biosamples | Assets | Forums | Events | Help

Provide Feedback | All | Organism | Go | Admin

New or upload

Data file

Go

Announcements

SEEK will be presented at ICSB 2012 and COMBINE / New SEEK version 4 months ago by Martin Golebiewski

Upload your data to SEEK as email attachment 5 months ago by Martin Golebiewski

New Virtual Liver SEEK with extended functionality 5 months ago by Martin Golebiewski

See all

Favourites

Tags [show all]

CellNetAnalyzer (CNA) data management Flow Cytometry HepatoSys IL-6 Insulin signalling Kinetic Modelling

Search Items by Scale

all -

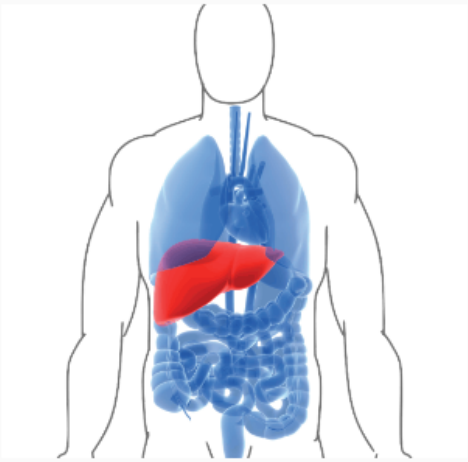
organism -

liver -

liver lobule -

intercellular -

cell -



Specimens (17+1) | Data Files (8+3) | Models (2+1) | SOPs (2) | Publications (9) | Presentations (29)

Test mouse

Scales: Organ (Liver)

Find, share and exchange **Data**, **Models** and **Processes** within the Virtual Liver Network.

Yellow Pages | Experiments | Biosamples | Assets | Forums | Events | Help

Provide Feedback | All | Liver | Go | Admin

New or upload

Data file

Go

Announcements

SEEK will be presented at ICSB 2012 and COMBINE / New SEEK version 4 months ago by Martin Golebiewski

Upload your data to SEEK as email attachment 5 months ago by Martin Golebiewski

New Virtual Liver SEEK with extended functionality 5 months ago by Martin Golebiewski

See all

Favourites

Tags [show all]

CellNetAnalyzer (CNA) data management Flow Cytometry HepatoSys IL-6 Insulin signalling Kinetic Modelling

Search Items by Scale

all -

organism -

liver -

liver lobule -

intercellular -

cell -

Investigations (1+1) | Studies (2)

Assays (3) | Samples (12) | Specimens (8+1) | Data Files (31+14) | Models (2) | Publications (11) | Presentations (58)

Scales: Liver Lobule

Find, share and exchange **Data**, **Models** and **Processes** within the Virtual Liver Network.

Yellow Pages | Experiments | Biosamples | Assets | Forums | Events | Help

Provide Feedback | All | Liver Lobule | Go | Admin

New or upload

Data file

Go

Announcements

SEEK will be presented at ICSB 2012 and COMBINE / New SEEK version 4 months ago by Martin Golebiewski

Upload your data to SEEK as email attachment 5 months ago by Martin Golebiewski

New Virtual Liver SEEK with extended functionality 5 months ago by Martin Golebiewski

See all

Favourites

Tags [show all]

CellNetAnalyzer (CNA) data management Flow Cytometry HepatoSys IL-6 Insulin signalling Kinetic Modelling

Search Items by Scale

all -

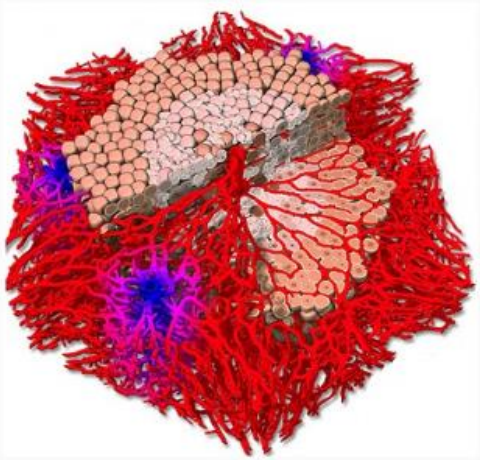
organism -

liver -

liver lobule -

intercellular -

cell -



Investigations (1)

Samples (1) | Specimens (1) | Data Files (9+2) | Models (2+1) | SOPs (1) | Publications (5) | Presentations (46+1)

Scales: Intercellular

Find, share and exchange **Data**, **Models** and **Processes** within the Virtual Liver Network.

Yellow Pages | Experiments | Biosamples | Assets | Forums | Events | Help

Provide Feedback | All | Intercellular | Go | Admin

New or upload

Data file

Go

Announcements

SEEK will be presented at ICSB 2012 and COMBINE / New SEEK version 4 months ago by Martin Golebiewski

Upload your data to SEEK as email attachment 5 months ago by Martin Golebiewski

New Virtual Liver SEEK with extended functionality 5 months ago by Martin Golebiewski

See all

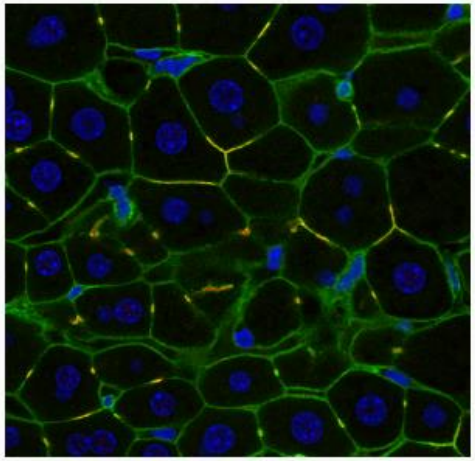
Favourites

Tags [show all]

CellNetAnalyzer (CNA) data management Flow Cytometry HepatoSys IL-6 Insulin

Search Items by Scale

all -
organism -
liver -
liver lobule -
intercellular -
cell -



Investigations (2+1) | Studies (1+1)

Assays (1) | Specimens (1) | Data Files (19) | Models (6+1) | SOPs (2) | Publications (13) | Presentations (65)

Scales: Cellular

Find, share and exchange **Data**, **Models** and **Processes** within the Virtual Liver Network.

Yellow Pages | Experiments | Biosamples | Assets | Forums | Events | Help

Provide Feedback | All | Cell | Go | Admin

New or upload

Data file

Go

Announcements

SEEK will be presented at ICSB 2012 and COMBINE / New SEEK version 4 months ago by Martin Golebiewski

Upload your data to SEEK as email attachment 5 months ago by Martin Golebiewski

New Virtual Liver SEEK with extended functionality 5 months ago by Martin Golebiewski

See all

Favourites

Tags [show all]

CellNetAnalyzer (CNA) data management Flow Cytometry HepatoSys IL-6 Insulin signalling Kinetic Modelling logical modelling metabolism

Search Items by Scale

all -

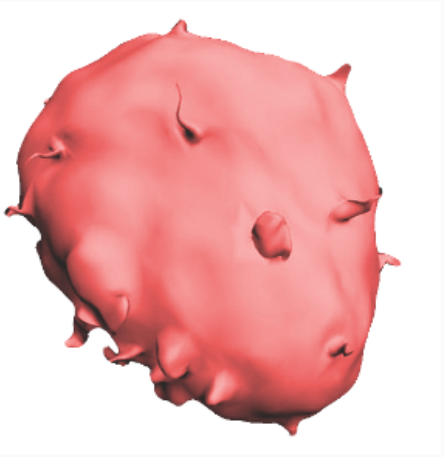
organism -

liver -

liver lobule -

intercellular -

cell -



Investigations (7+9) | Studies (6+10) | Assays (8+9)

Samples (10+1) | Specimens (6+2) | Data Files (69+56) | Models (54+7) | SOPs (11) | Publications (62) | Presentations (134+1)

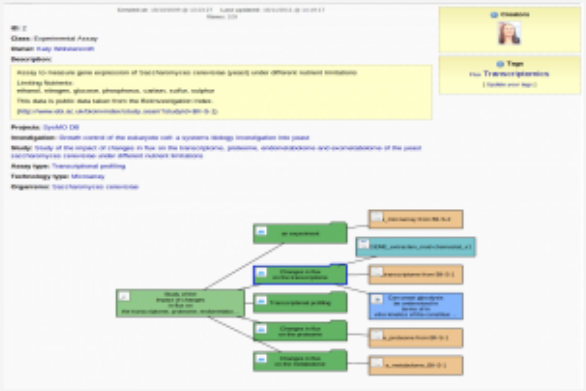
Insulin signalling in primary hepatocytes

SEEK for Science

http://www.seek4science.org

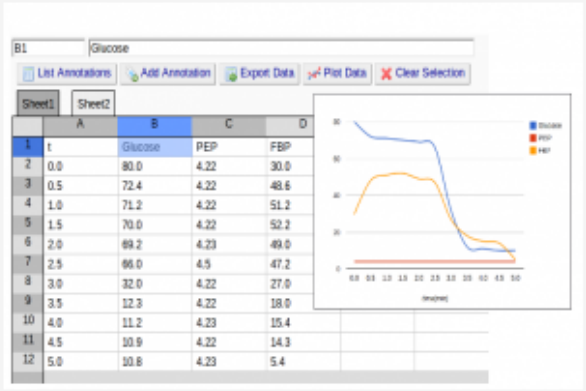
For finding, sharing and exchanging Data, Models and Processes in Systems Biology.

Organise your experiments and data



SEEK has adopted an ISATAB style structure for organising experiments and data.

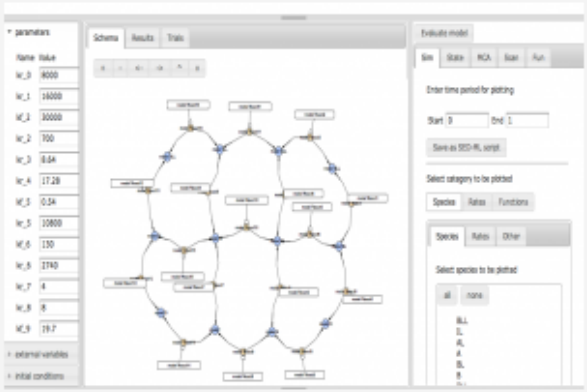
Explore and annotate data



	A	B	C	D
1	t	Glucose	PEP	FBP
2	0.0	80.0	4.22	30.0
3	0.5	72.4	4.22	48.6
4	1.0	71.2	4.22	51.2
5	1.5	70.0	4.22	52.2
6	2.0	69.2	4.23	49.0
7	2.5	66.0	4.5	47.2
8	3.0	32.0	4.22	27.0
9	3.5	12.3	4.22	18.0
10	4.0	11.2	4.23	15.4
11	4.5	10.9	4.22	14.3
12	5.0	10.8	4.23	5.4

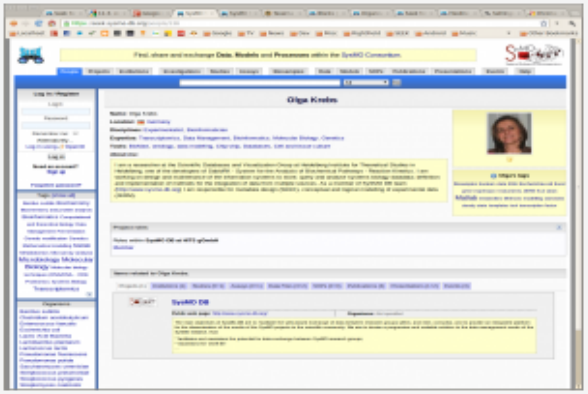
Excel spreadsheets can be explored and annotated without the need to download.

Simulate SBML models



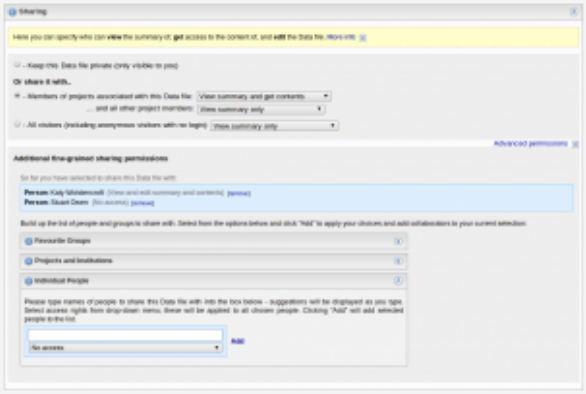
Most models that conforms to the SBML format can be simulated within SEEK.

Who's doing what, where?



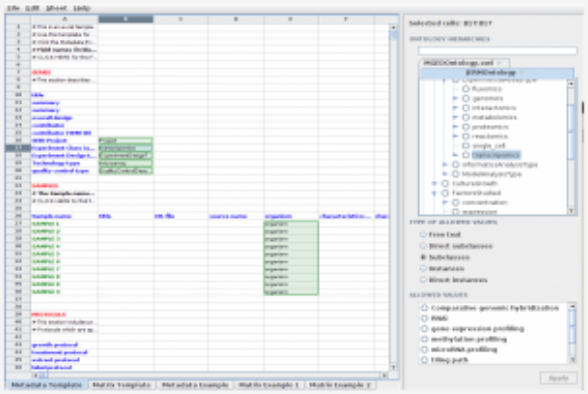
We recognise that people, and their knowledge, are important.

Flexible sharing controls



There is a lot of flexibility and control over who can see, download or edit your items.

Semantic spreadsheet templates



Using RightField we are producing a wide collection of template files.

MANCHESTER 1824

The University of Manchester

SEEK in Use

Developed for:

SysMO
Systems Biology of Microorganisms

ERASysBio Plus
Host–pathogen interactions

virtual liver network

Adopted by:

ElasticAP

eagle genomics

unicellsys

EVIMalaR

SBCancer.

JENAGE
JENA CENTRE FOR SYSTEMS BIOLOGY OF AGEING

RosAge
Reactive oxygen species and the dynamics of ageing

Consensus model of Yeast Glycolysis

CISBIC

sybacol

THE SEEK BASIC PRINCIPLE: USE WHAT IS AVAILABLE

NO REINVENTION

BUILD ON COMMUNITY ACTIVITIES

Minimum Information Models

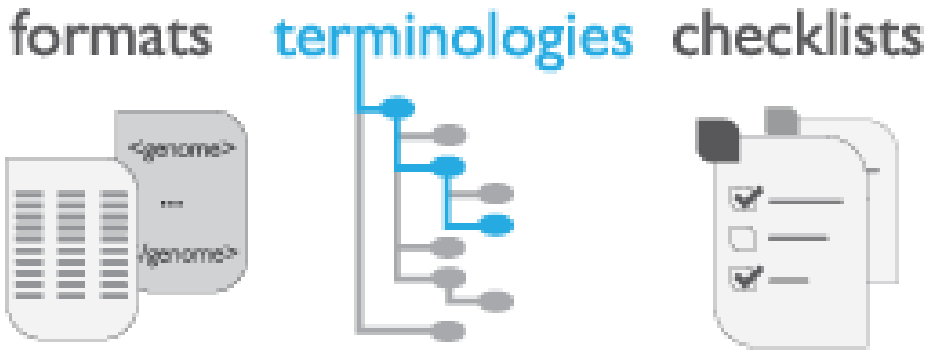
Formats, Ontologies, Naming schemes and Controlled Vocabularies

What is the least amount of information required to:

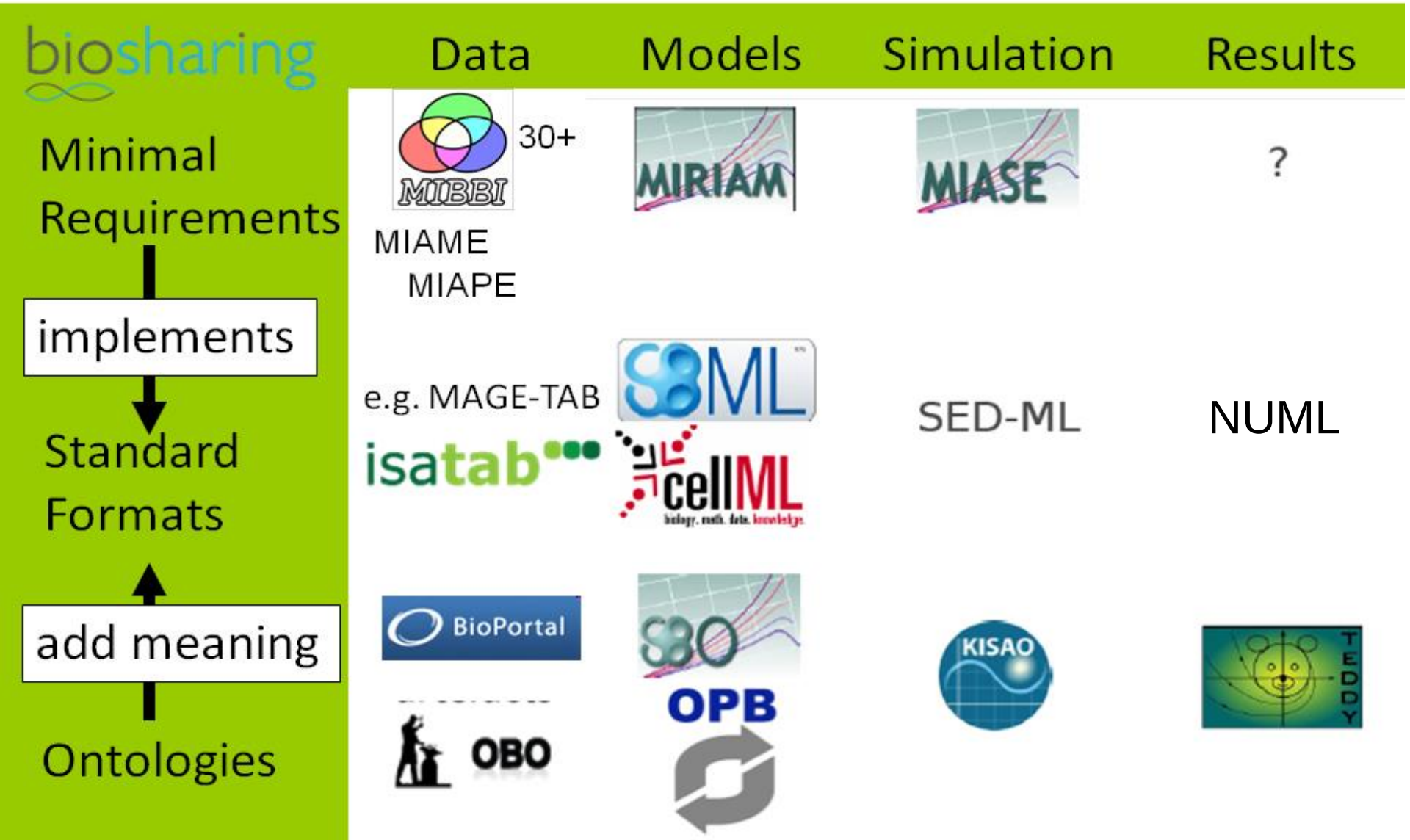
Find, Interpret, Understand, Reuse

Realistic and pragmatic expectations for annotation
Ask for too much = people overwhelmed and provide less
Ask for too little = people will oblige

Transcriptomics
Proteomics
Metabolomics
30+ MIBBI



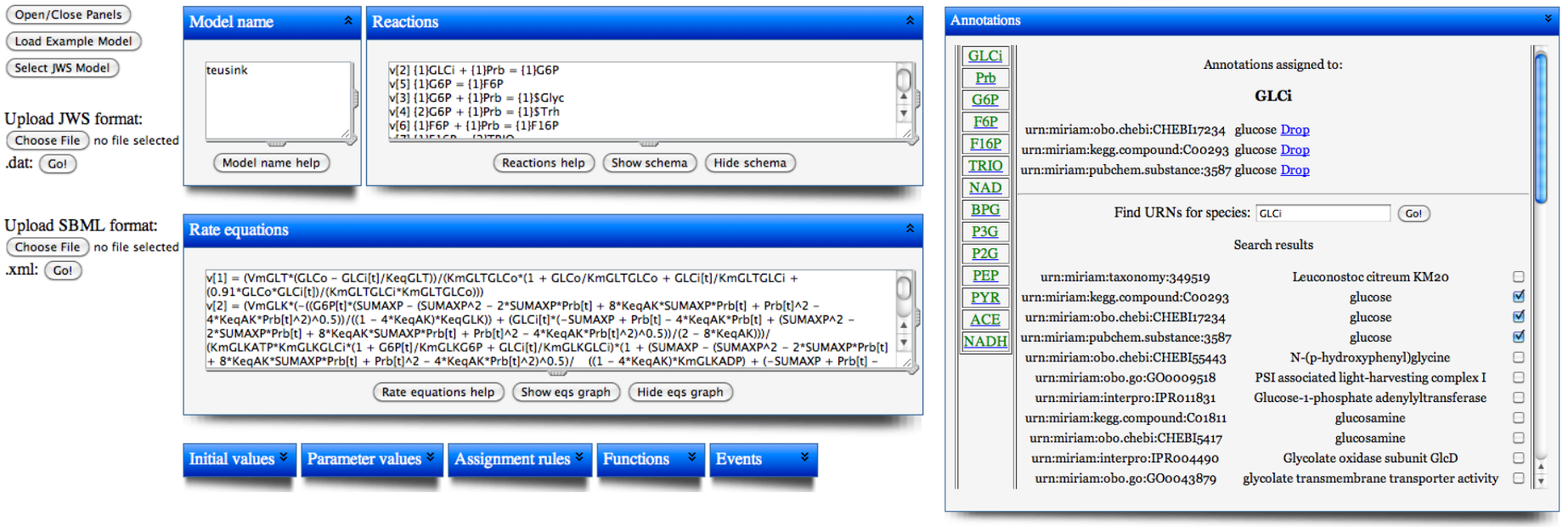
Matrix of Standards and Formats



Nicolas Le Novere, Data Integration in the Life Sciences, UK, 2009

JWS OneStop to Model Standards

SBML model format, SBGN schema, MIRIAM annotation



SEEK integrated tool to:

- construct or modify models, save in SBML format
- get automated generation of SBGN schema
- annotate model (MIRIAM), implementing web services from semanticSBML.

semantic SBML

MANCHESTER 1824

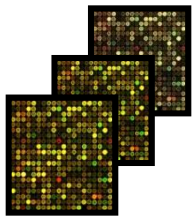
The University of Manchester

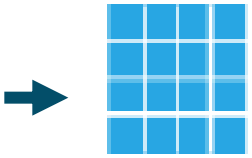
Omics data

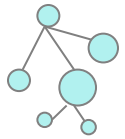
standard based templates

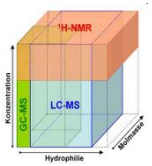
RightField

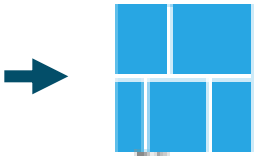
JERM

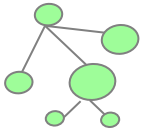


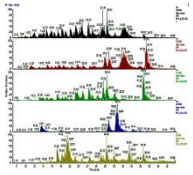


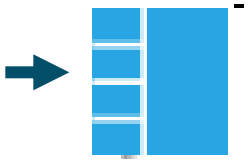


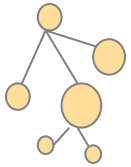


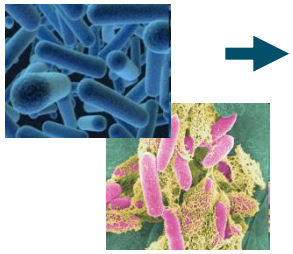


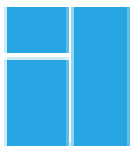


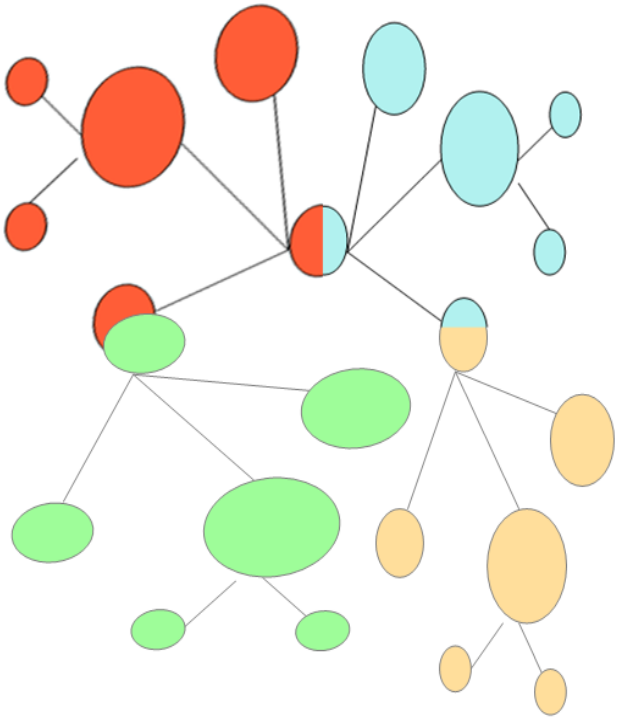




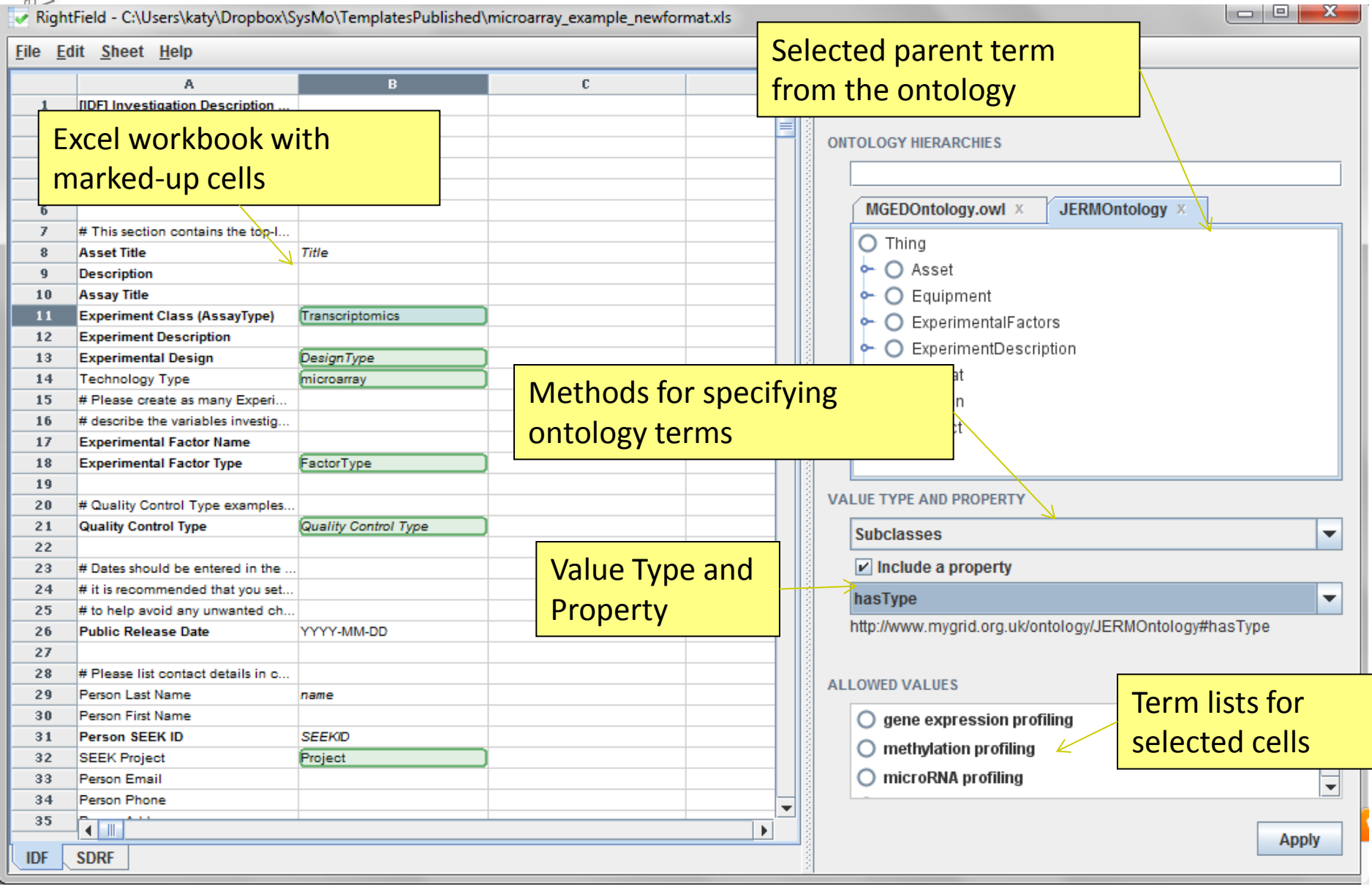








RightField: Managing Vocabularies



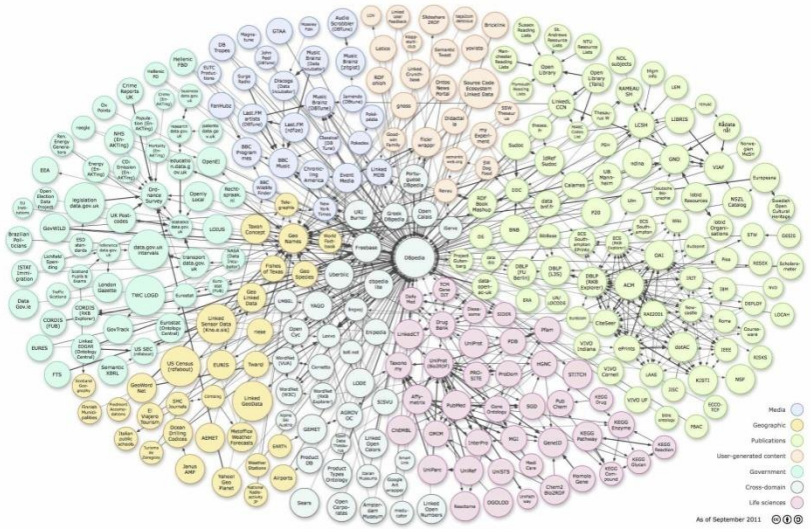
MANCHESTER 1824

The University of Manchester

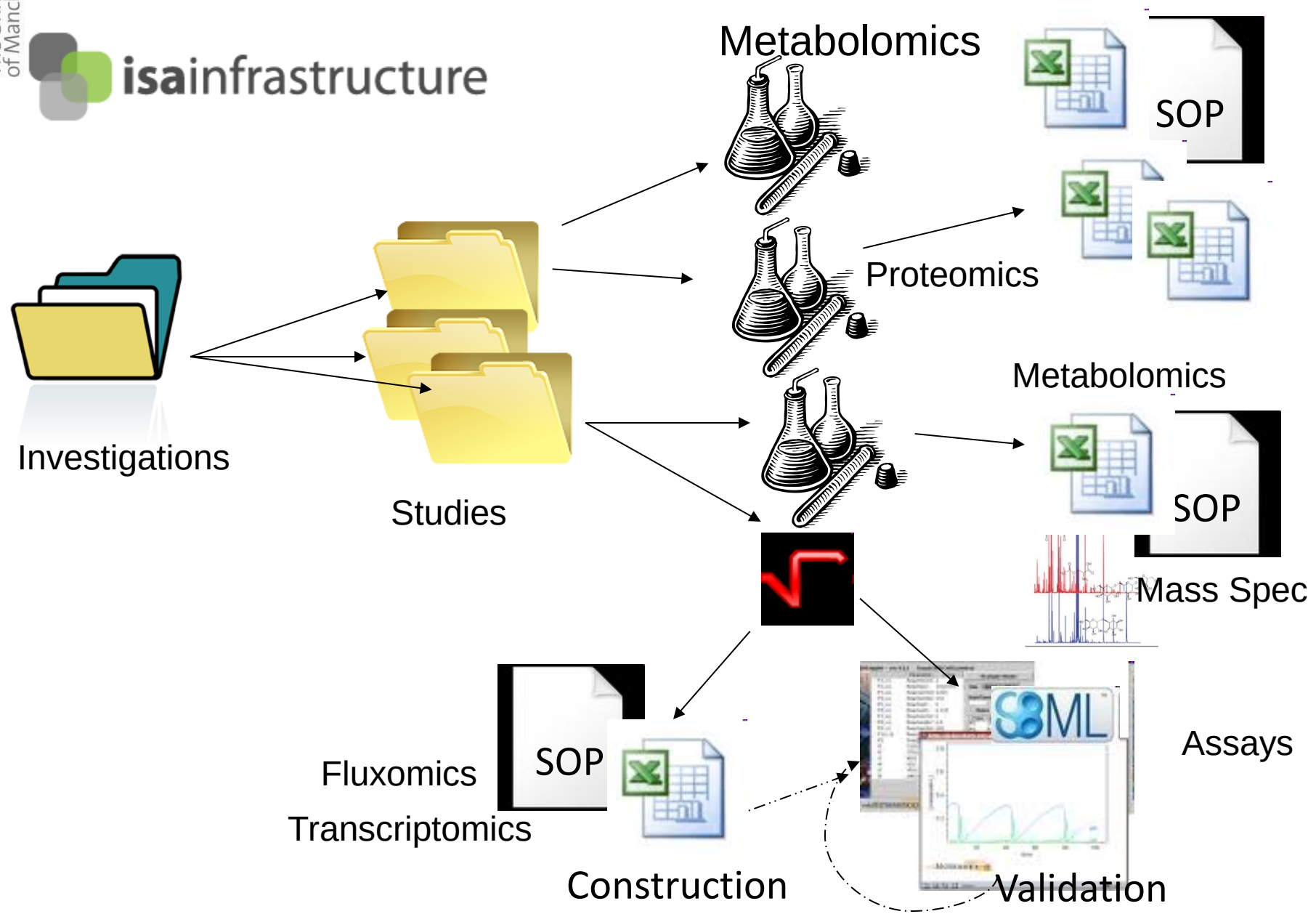
Semantic Linking and Querying

Extracting and storing metadata in RDF (Resource Description Framework), via RightField

- Better searching and querying
- New representations and visualisation of relationships between SEEK assets
- Linking SEEK data to the web of Linked Data

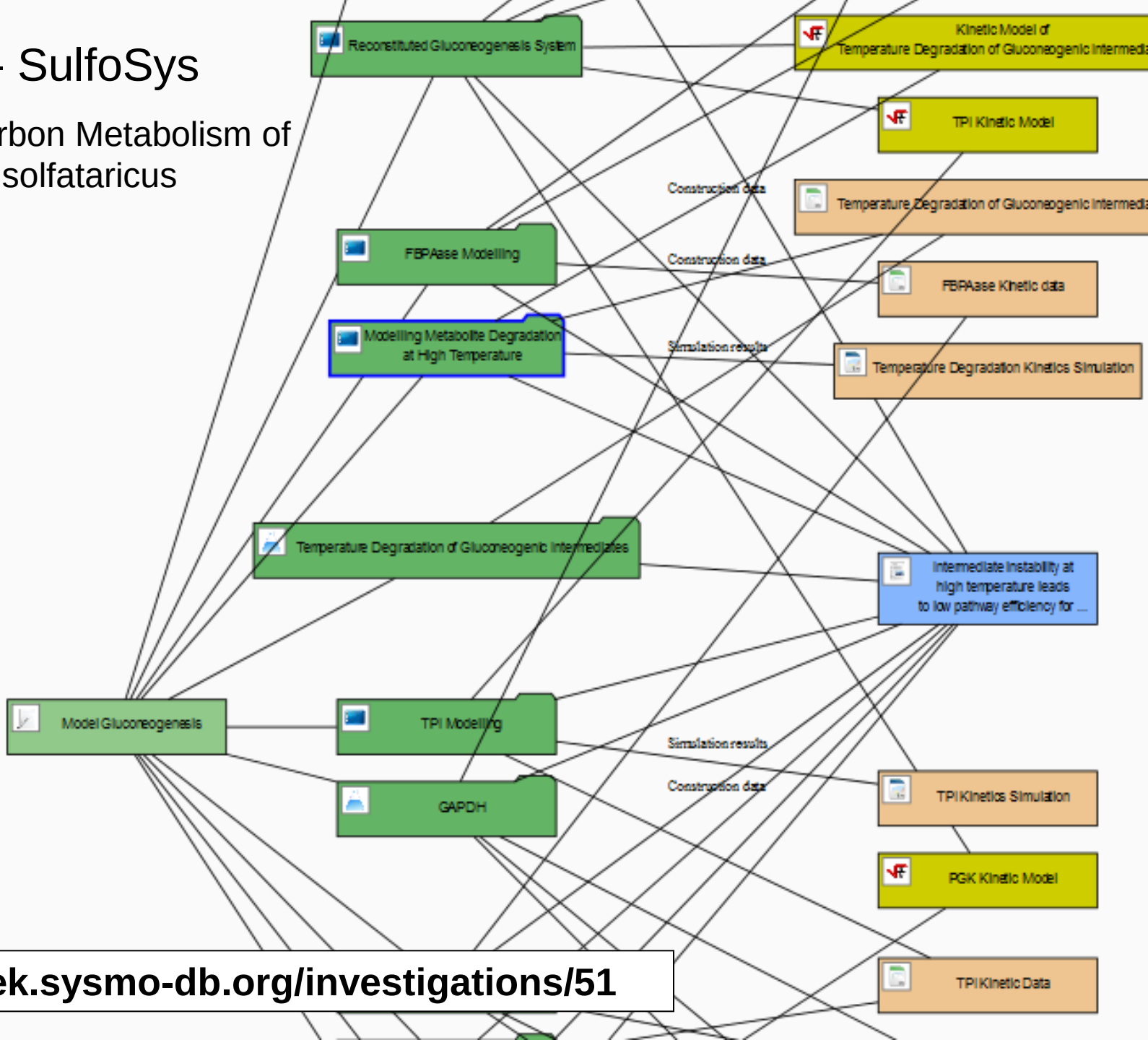


SEEK Links Data and Models



SysMO - SulfoSys

Central Carbon Metabolism of Sulfolobus solfataricus



https://seek.sysmo-db.org/investigations/51

MANCHESTER 1824

The University of Manchester

Incentives for Using SEEK

- Safe haven for Systems Biology assets
- Credit and attribution
  - Linking people to the data/models/sops etc
  - Linking assets to publications

DataCite
International Data Citation

- Publicity
  - SEEK records when data was uploaded/shared
  - Dissemination as well as internal project sharing

Incentives for using SEEK

- Exporting and linking
  - Supplementary materials store – persistent URIs
    - Linking publications and data
    - E.g. FEBS Journal
  - Submission to public repositories
    - Silos not ideal for Systems Biology, but useful for meta -analysis

MANCHESTER 1824

The University of Manchester

Why it works for us

- Off the shelf Systems Biology sharing environment
- Fits in with existing data and model management practices
- Incremental production with rapid prototypes and feedback from SEEK focus group (aka, the PALS)
- Publish and share within the consortium and beyond
- Scientists stay in control

Briefings in Bioinformatics Advance Access published October 9, 2012
BRIEFINGS IN BIOINFORMATICS. page 1 of 14
doi:10.1093/bib/bbs064

Data management strategies for multinational large-scale systems biology projects

Wasco Wruck, Martin Peuker and Christian R.A. Regenbrecht

Submitted: 13th July 2012; Received (in revised form): 4th September 2012

Key Points

- Incentives to share data can be given by data citation credits (datacite).
- Open access to research data can be advanced via making sharing a condition of funding.
- Data management systems might be made attractive via alleviating and improving researchers' work, e.g. support for standard formats and publications.
- The reviewed systems proved useful for systems biology projects at least in dedicated environments, SysMO-SEEK out-of-the-box provides most useful features for large-scale systems biology projects.

http://bib.oxfordjournals.org/content/early/2012/10/09/bib.bbs064.full.pdf

Acknowledgements: SysMO-DB Team


Carole Goble


Stuart Owen


Katy Wolstencroft


Rob Haines


Niall Beard


Franco Du Preez Jacky Snoep


Finn Bacall Matt Horridge


Sergejs Aleksejevs


Wolfgang Mueller Olga Krebs Quyen Nguyen


Lihua An


David Shockley


Martin Golebiewski


Meik Bittkowski


Andreas Weidemann