





MANCHESTER

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

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



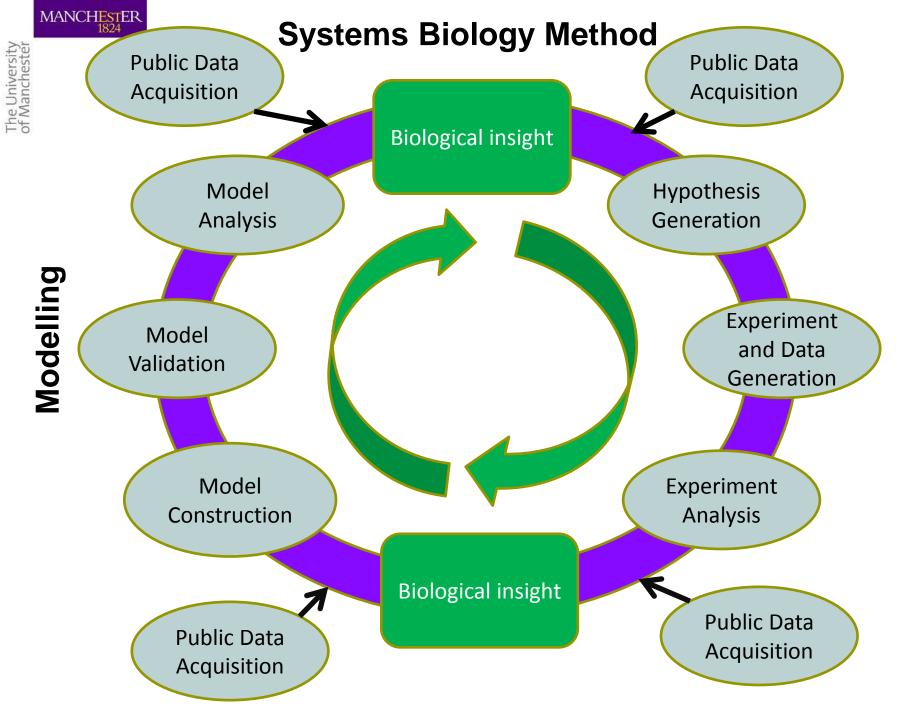




The University of Manchester



Federal Ministry of Education and Research



Experimental

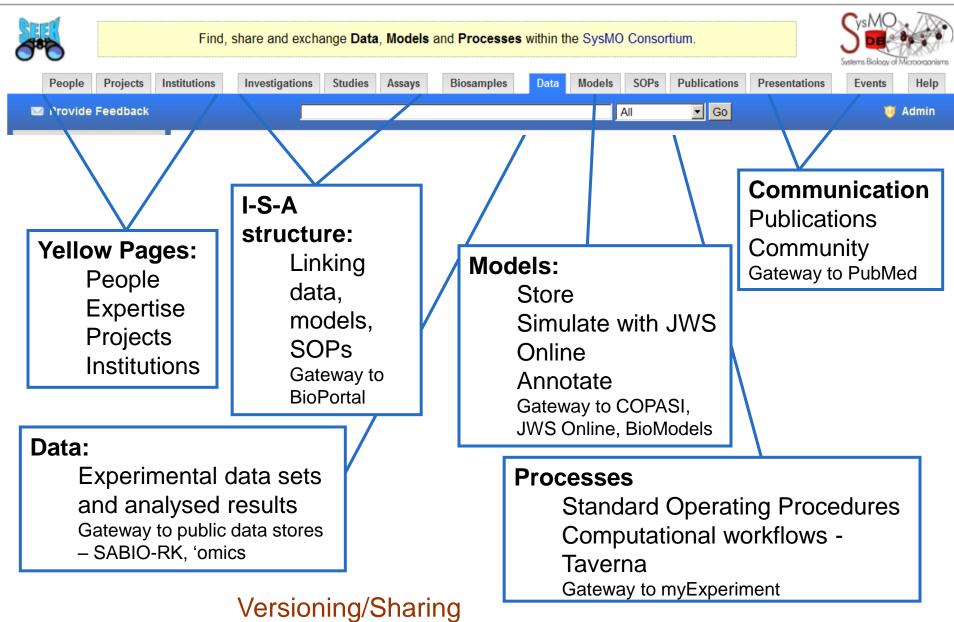


MANCHESTE

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.



Exploring and Analysing: The SEEK as a Gateway

Cytoscape

VmGLT KeqGLT

VmGLK KenGLK

KmGLKGLCi KmGLKG6P

KmGI KATP

KmGLKADI VmPGI

KeqPGI KmPGIG6 KmPGIF6

VmPEK

CIPFKATP

KPFKF26BP CPFKF26BP KPFKF16RP

gR L0 CPFKF6P KmPFKATF CPFKATP KPFKAMP CPFKAMP KIPFKATP

KmGLTGLCo KmGLTGLCi

Web Visualise Model with Cytoscape Web

CHEMBL101901 CHEMBL265374

MANCHESTER

coom)

CHEMBL5245

CHEMBLS65432

CHEMBL579219 CHEMBL52389 CHEMBL568574 CHEMBL373333

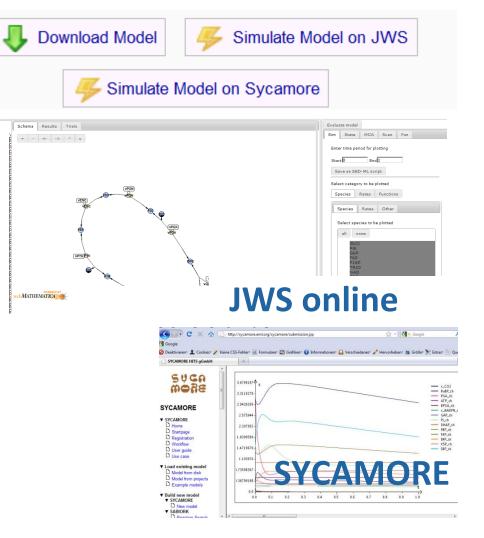
CHEMBL102074

CHEMBL567914

The University of Manchester

1	0	icose .					
81	Lini Annatabo	Att An	natidan 🐨 D	que Desa	Past Dea	Cew Selection	
510	sheet	2			1		
	A		c	D			
1	t	Giocone -	PEP	FBP			141
2	0.0	810	4.22	30.0			
	0.5	72.4	4.22	40.6	1		
4	1.0	71.2	4.22	51.2	.7		
53	1.5	75.0	4.22	52.2	1.0		
4	2.0	68.2	4.23	49.0		C	2
21	2.5	05.0	45	47.2			
R.	3.0	32.0	4.22	27.0	-10-1	e in ay in the second	1 49 10
8)	3.5	12.3	4.22	18.0		*****	
10	4.0	11.2	4.23	15.4			
10	45	30.9	4.22	14.1			
12	50	22.8	4.23	54			

Explore spreadsheets





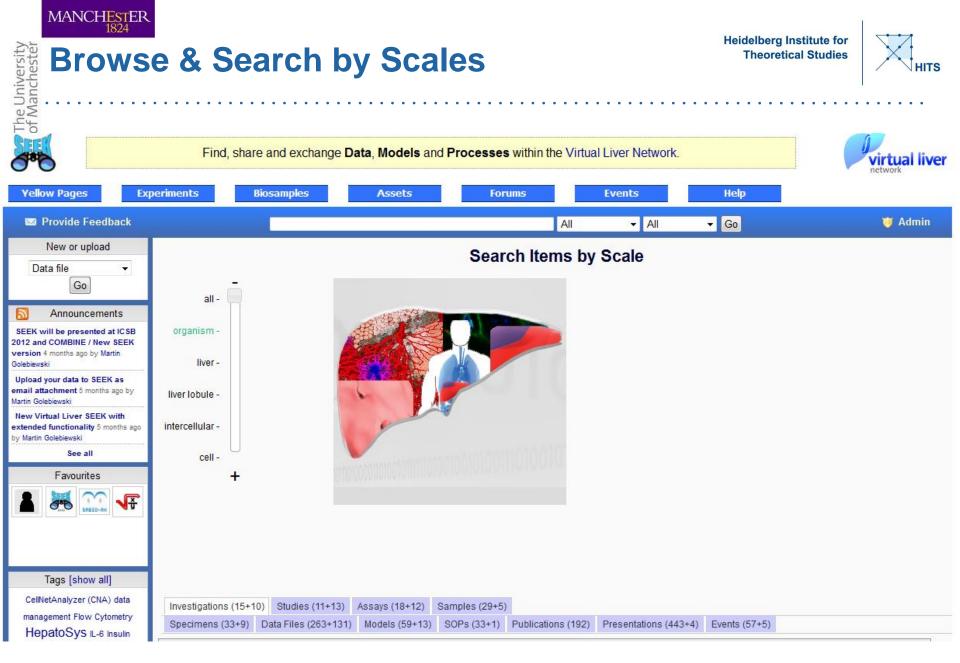


- Pan European collaboration
- 13 individual projects,
 >100 institutes
- Model dynamic molecular processes in microorganisms
- Pool research capacities and know-how
- Disseminate activities of the consortium

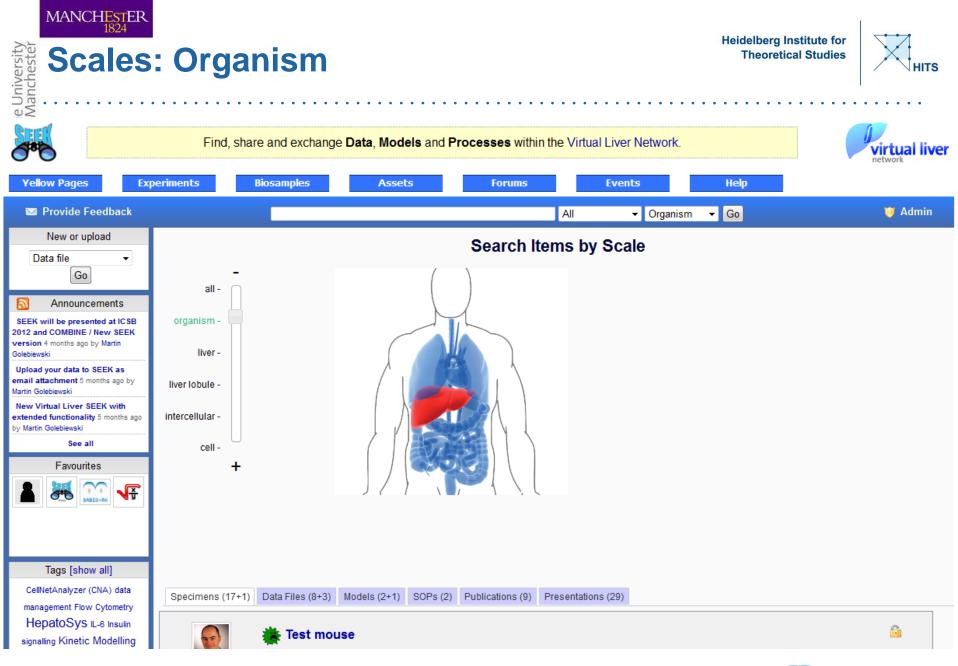
SEEK Communities



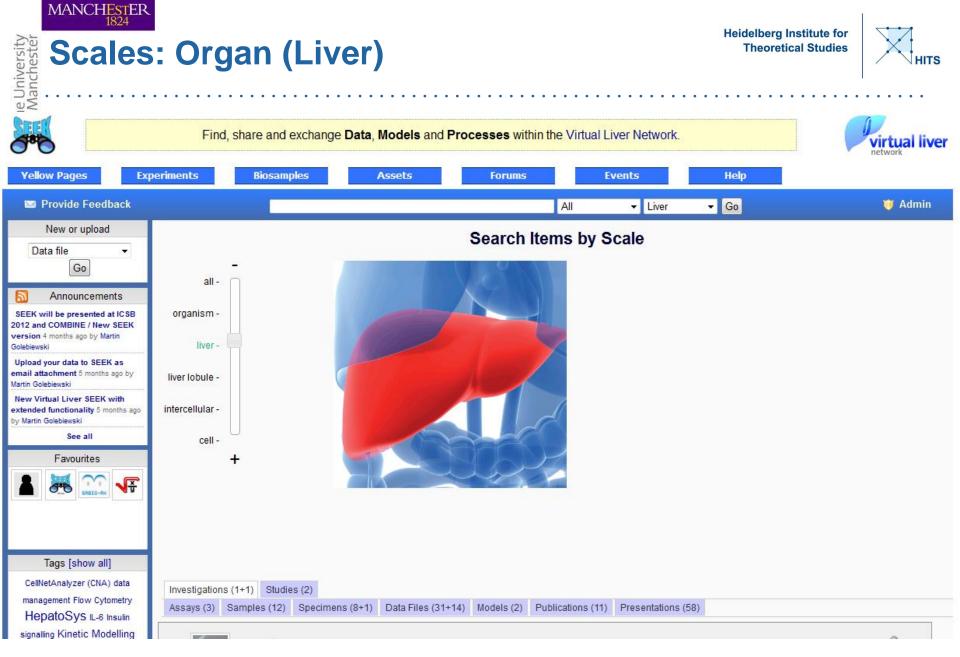
- German research network
- ~45 organisations, ~70 groups
- Multiscale representation of the liver
- Pool research capacities and know-how
- Disseminate activities of the consortium



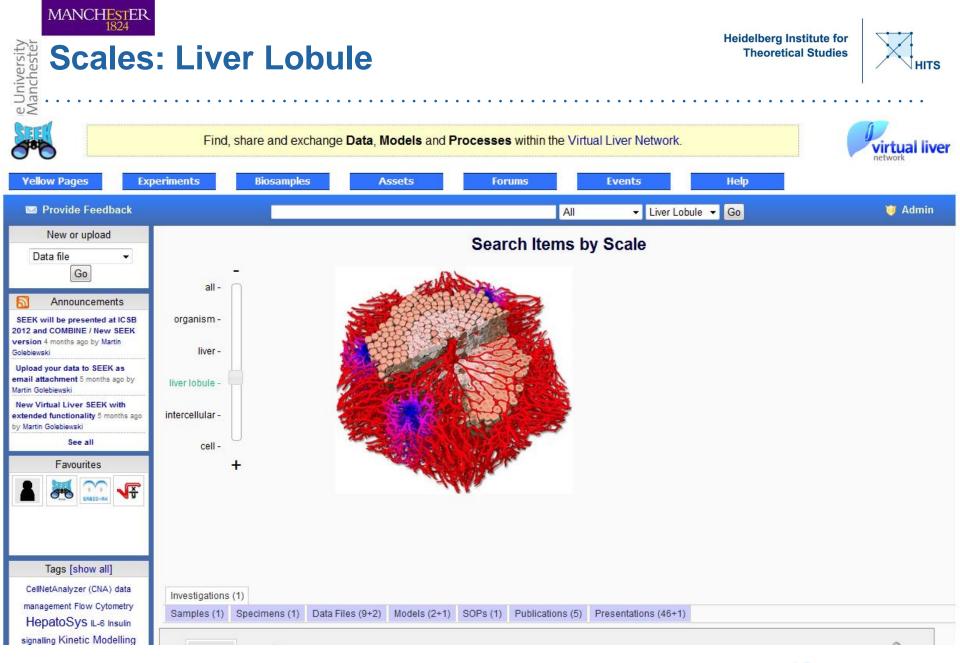




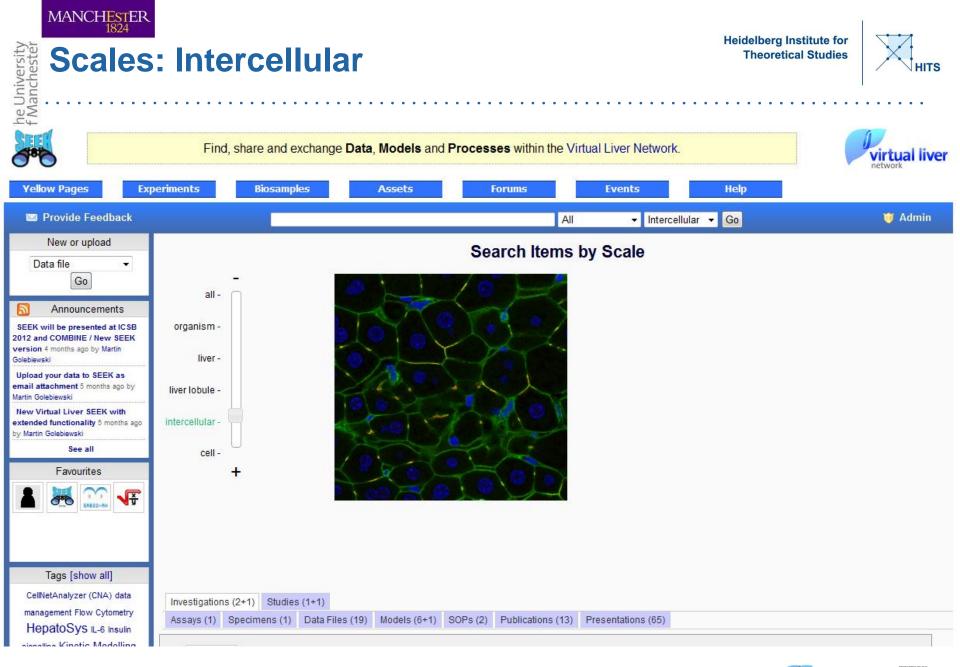




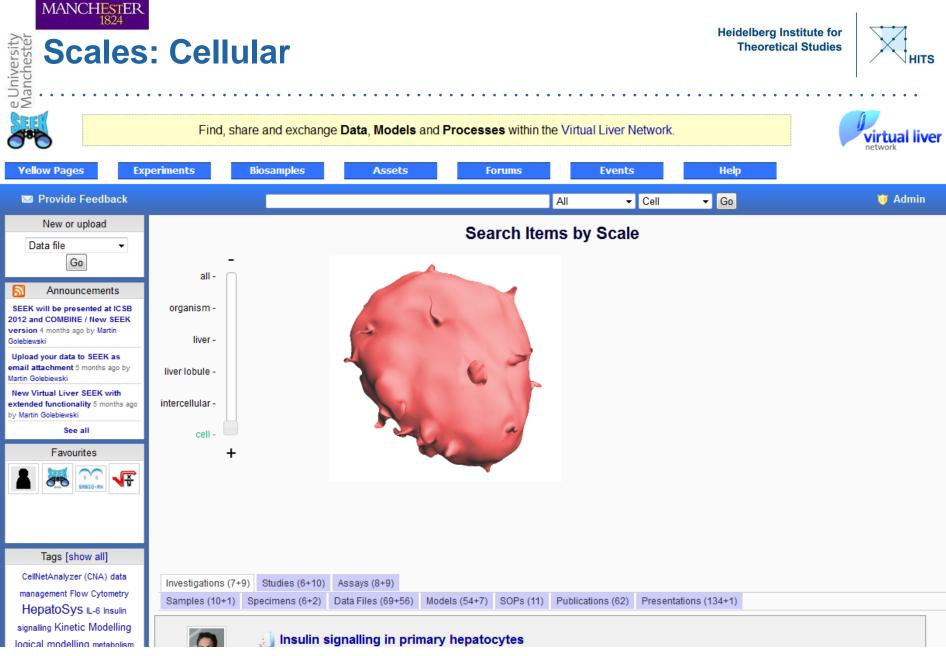














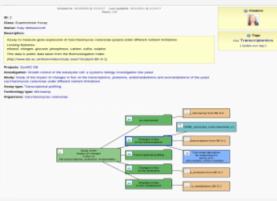


For finding, sharing and exchanging Data, Models and Processes in Systems Biology

Explore and annotate data

The of A

Organise your experiments and data

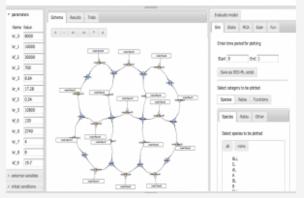


SEEK has adopted an ISATAB style structure for organising experiments and data.

1	G	ucose						
List Annotations		ns 👒 Add An	notation	🙀 Export Data	14 ⁴ Piet Data	X Cle	ar Selection	
She	et1 Shee	12						
_	A	В		C	0 *			E thee
1	t	Glucose	PEP	FBP			~	# PEP
2	0.0	80.0	4.22	30.0			1	
3	0.5	72.4	4.22	48.6	- /		7	
4	1.0	71.2	4.22	51.2	1 1			
5	1.5	70.0	4.22	52.2				
6	2.0	69.2	4.23	-49.0				~
7	25	66.0	4.5	47.2				
8	3.0	32.0	4.22	27.0	68.6	18 18	20 25 33 25 40	45 50
9	3.5	12.3	4.22	18.0			(section)	
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?

	C	
-	First share and nucleonge Data. Hodals and Presenance other its SysBC Consolars.	S= ++*
Course 1984	an Research Research States Surger States and States Area and	anamata Baama 1989
Log In Heginer	Olga Kreins	
	Term rem takan generali and an	Control of the second s
Bellegy search may schemes from the restored to the	New start is figured as the second start and starting the second start and starting the second start as th	NUM (N
Toportina Sector Action Sector Act	Entropy Payment III Main Company Magnetic IIII (Second Company) Main Company Magnetic IIIII (Second Company) Main Company Magnetic IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	

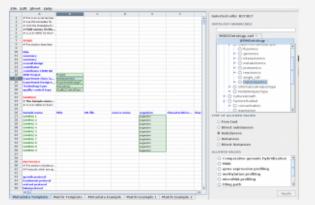
We recognise that people, and their knowledge, are important.

Flexible sharing controls

) Sharing	8
Here you can specify white can view the summary of, get access to the content st, and will the data file. However, 👷	
- Keep the lines for private (priv visible to pro)	
Or share it with.	
# - Members of projects associated with this Data file. View summary and get contants.	
and all other project members: Views summary anly *	
2 - All visitors (initialing anonymous visitors with no logity) (max contrary only	
	Advanced permission is
Additional fine-grained sharing permissions	
Se far you have selected its share this thata file with:	
Person Kaly Ministerandi (Here and edi surrany and activity) passe) Person Raat Dawn (An access) passes)	
Build up the link of people and progenic share with. Seried how the update before and sick "Mat" is apply your division and all	I collaboration to your surrent telecitors
Provide Temps	
O Projects and Institutions (0)	
() Individual People ()	
Please type names of people to share this Qubs like with into the lock below - suggestions will be displayed as you type. Share asket rights from disponser memory there will be applied by all choice people. Choice YAO' will add allocate people to the lock.	
Statum AM	

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.

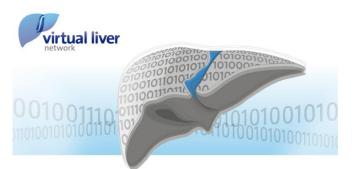


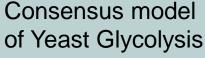


Developed for:



RASysBio Plus Host-pathogen interactions





JENA CENTRE FOR

SYSTEMS BIOLOGY OF AGEING

in Use

ElasticAP

eagle

Adopted by:







RosAge

Reactive oxygen species and the dynamics of ageing

nsus model





MANCHESTER

THE SEEK BASIC PRINCIPLE: USE WHAT IS AVAILABLE NO REINVENTION BUILD ON COMMUNITY ACTIVITIES



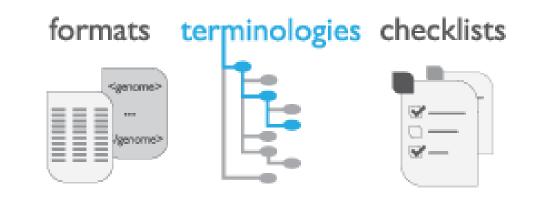
Minimum Information Models

Formats, Ontologies, Naming schemes and Controlled Vocabularies

What is the least amount of information required to: Find, Interpret, Understand, Reuse

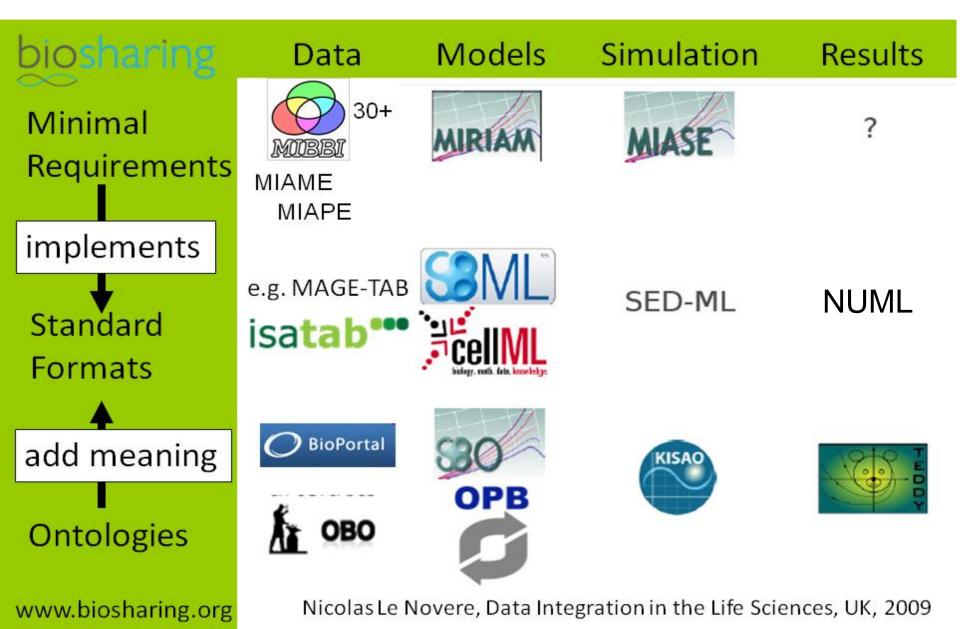
Realistic and pragmatic expectations for annotation Ask for too much = people overwhelmed and provide less Ask for too little = people will oblige

Transcriptomics Proteomics Metabolomics 30+ MIBBI





Matrix of Standards and Formats





JWS OneStop to Model Standards

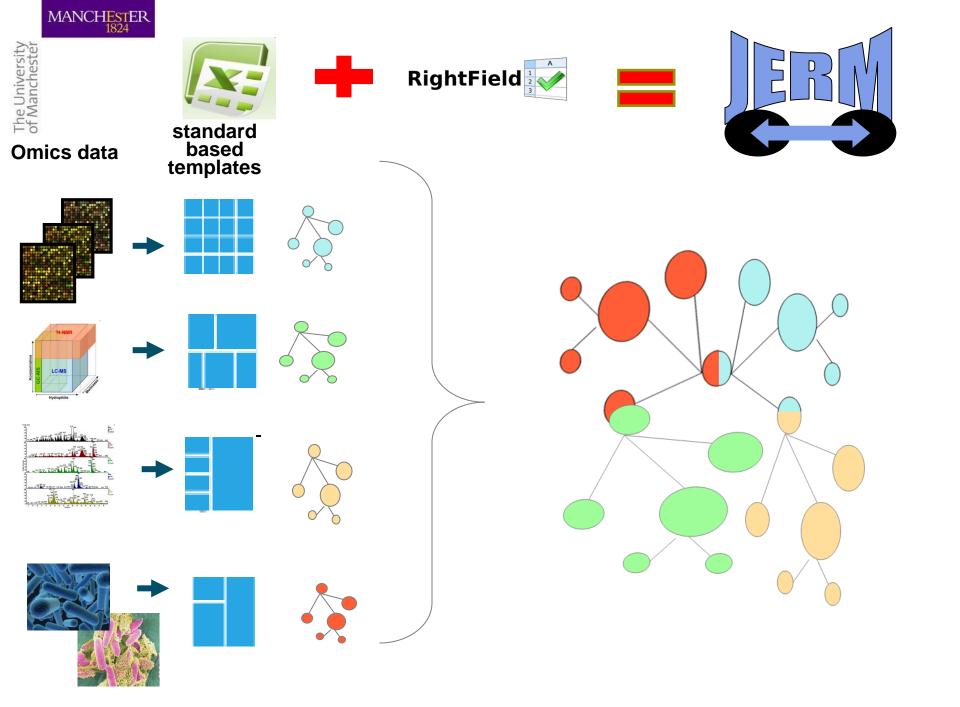
SBML model format, SBGN schema, MIRIAM annotation

Open/Close Panels	Model name *	Reactions	Annotat	ons		×
Load Example Model Select JWS Model Upload JWS format:	teusink	v[2] {1}CLCi + {1}Prb = {1}GGP v[5] {1}GGP = {1}FGP v[3] {1}GGP = {1}FGP v[3] {1}GGP + {1}Prb = {1}SCIyc v[4] {2}GGP + {1}Prb = {1}STrh v[6] {1}FGP + {1}Prb = {1}F1GP v[7] {1}GCP - {1}F1GP	GLC Prb G6F F6F	Anno	tations assigned to: GLCi glucose <u>Drop</u>	
Choose File) no file selected .dat: Go!	Model name help	(Reactions help) (Show schema) (Hide schema)	F16 TRIC	urn:miriam:kegg.compound:C00293 urn:miriam:pubchem.substance:3587		
Upload SBML format: Choose File no file selected	Rate equations	*	BPC <u>P3C</u> P2C	Find URNs for speci	es: GLCi Gol Search results	
.xml: (Go!)	(0.91*GLCo*GLCi[t])/(KmGL v[2] = (VmGLK*(-((GGP[t]*C 4*KeqAK*Prb[t]^2)^0.5))/((1 2*SUMAXP*Prb[t] + 8*KeqAI (KmGLKATP*KmGLKGLCi*(1	il(1/KeqGLT)/(KmGLTGLCo*(1 + GLCo/KmGLTGLCo + GLCi[t]/KmGLTGLCi + TGLCi*KmGLTGLCo))) UMAXP - (SUMAXP*2 - 2*SUMAXP*Prb[t] + 8*KeqAK*SUMAXP*Prb[t] + Prb[t]^2 - - 4*KeqAKY*KeqGLK() + (GLCi[t]*(-SUMAXP + Prb[t] - 4*KeqAK*Prb[t] + (SUMAXP^2 - *SUMAXP*Prb[t] + Prb[t]^2 - 4*KeqAK*Prb[t]^2/0.5))/(2 - 8*KeqAK)))/ + G6Pt[1/KmCLKG6P + GLCi[t]/*(MCKGLC)*(1 + (SUMAXP - (SUMAXP*Prb[t] - + G6Pt[1/KmCLKG6P + GLCi[t]/*(MCKGLC)*(1 + (SUMAXP - (SUMAXP*Prb[t] - + Prb[t]^2 - 4*KeqAK*Prb[t]^2/0.5)/ ((1 - 4*KeqAK)*KmGLKADP) + (-SUMAXP + Prb[t] - Rate equations help Show eqs graph Hide eqs graph	PEF PYF ACF NAD	urn:miriam:taxonomy:349519 urn:miriam:kegg.compound:C00293 urn:miriam:obo.chebi:CHEBI17234	Leuconostoc citreum KM20 glucose glucose glucose N-(p-hydroxyphenyl)glycine PSI associated light-harvesting complex I Glucose-1-phosphate adenylyltransferase glucosamine	
	Initial values * Parameter	er values * Assignment rules * Functions * Events *		urn:miriam:obo.chebi:CHEBI5417 urn:miriam:interpro:IPR004490 urn:miriam:obo.go:GO0043879	glucosamine Glycolate oxidase subunit GlcD glycolate transmembrane transporter activity	

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.





University

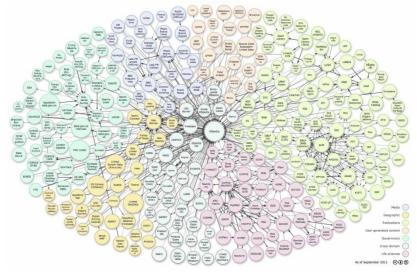
RightField: Managing Vocabularies

Right								
ile <u>E</u> d	lit <u>S</u> heet <u>H</u> elp			Selected parent term				
	A B		С	from the ontology				
1	[IDF] Investigation Description							
_	xcel workbook w harked-up cells	ith		ONTOLOGY HIERARCHIES				
7	# This section contains the top-I							
8	Asset Title	Title		O Thing				
9	Description	7.02		🕨 🔶 🔿 Asset				
10	Assay Title			← O Equipment				
11	Experiment Class (AssayType)	Transcriptomics						
12	Experiment Description			← O ExperimentalFactors				
13	Experimental Design	DesignType		 C ExperimentDescription 				
14	Technology Type	microarray		at				
15	# Please create as many Experi		Methods for spec	citying _n				
16	# describe the variables investig							
17	Experimental Factor Name		ontology terms					
18	Experimental Factor Type	FactorType						
19								
20	# Quality Control Type examples			VALUE TYPE AND PROPERTY				
21	Quality Control Type	Quality Control Type)	Subclasses				
22								
23	# Dates should be entered in the		Value Type and	✓ Include a property				
24	# it is recommended that you set							
25	# to help avoid any unwanted ch		Property					
26	Public Release Date	YYYY-MM-DD	. ,	http://www.mygrid.org.uk/ontology/JERMOntology#hasType				
27								
28	# Please list contact details in c			ALLOWED VALUES				
29	Person Last Name	name		Term lists for				
30	Person First Name			O gene expression profiling				
31	Person SEEK ID	SEEKID		— O methylation profiling Selected cells				
32	SEEK Project	Project	J					
33	Person Email			O microRNA profiling				
34	Person Phone							
35	35 Apply							

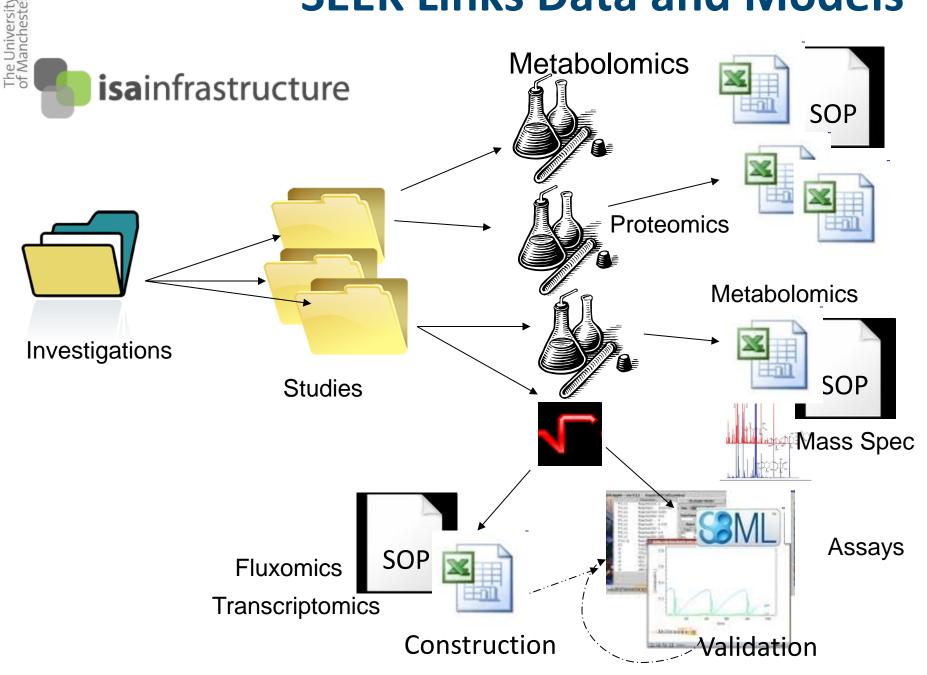


Semantic Linking and Querying Foring metadata in RDF (Resource

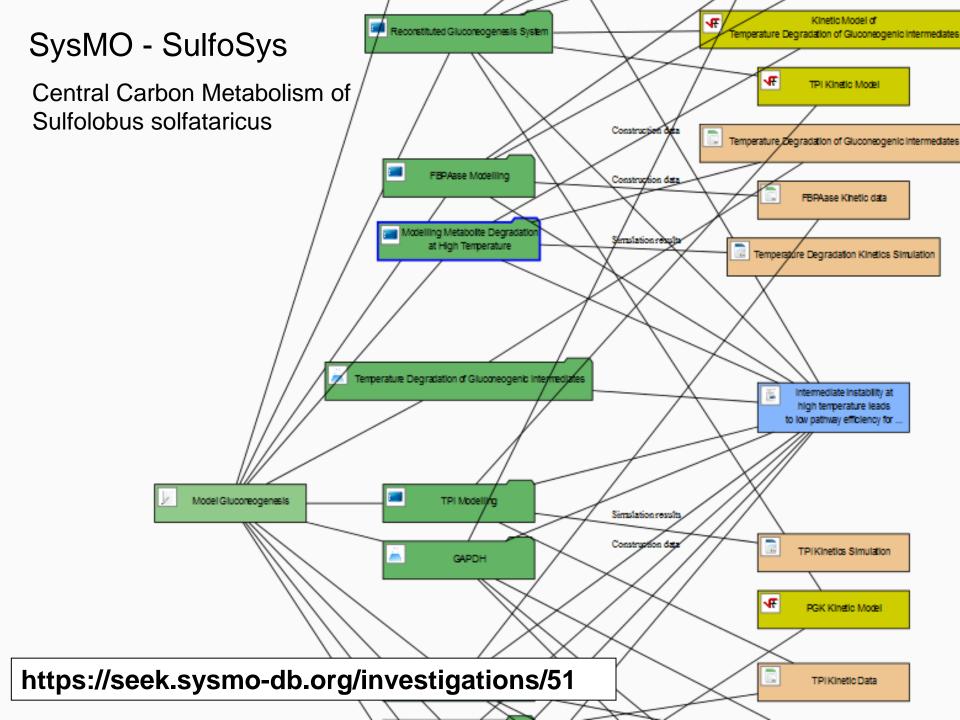
- Extracting and storing metadata in RDF (Resource Description Framework), via RightField
- Better searching and querying
- New representations and visualisation of relationships between SEEK assets
- Linking SEEK data to the web of Linked Data



SEEK Links Data and Models



MANCHESTER



Incentives for Using SEEK

00001

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

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

100 0 1000 10°

Incentives for using SEEK

MANCHEST

- 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



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

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

MANCHESTER 1824

The University of Manchester

Acknowledgements: SysMO-DB Team

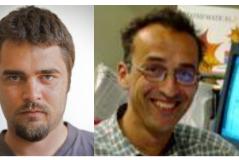


The University of Manchester



Carole Goble Stuart Owen Katy Wolstencroft





Franco Du Preez Jacky Snoep



Finn Bacall Matt Horridge

Sergejs Aleksejevs

Niall Beard





Wolfgang Mueller Olga Kreb: Quyen Nguyen







Lihua An David Shockley Martir

Rob Haines







Andreas Weidemann