Introducing the "Fostering the uptake of RDA indicators in Systems Biomedicine as a measure for model quality and FAIRness within the COMBINE community" project

Dr. Irina Balaur[1], Dr. Venkata Satagopam[1] and Prof. Dagmar Waltemath[2]

1. Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Luxembourg

2. Medical Informatics Laboratory, University Medicine Greifswald, Germany

EOSC SYMPOSIUM 2022, Prague, November 16, 2022

Brief introduction

Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg:
https://wwwen.uni.lu/lcsb/

Medical Informatics Laboratory, University Medicine Greifswald:
https://www.medizin.uni-greifswald.de/medizininformatik

Common research interests
- FAIRification
- Research data management
- Health Standards, Systems Biology Standards
- Reproducible research

What is systems biomedicine?

Systems Biomedicine is the science that studies how biological function emerges from the interactions between the components of living systems, in particular in healthy systems and during disease progression ...

… and how these emergent properties enable or constrain the behavior of the components under study.

Example of a curated biosimulation Novak and Tyson (1997). DOI: 10.1073/pnas.94.17.9147. BIOMD0000000007

New **models** of atherosclerosis and multi-drug therapeutic interventions
A Parton, V McGilligan, M Chemaly, M O'Kane... - ..., 2018 - academic.oup.com
... Such treatments are **clinically** effective in managing patient risk ... This enabled us to compile a series of **clinical** observations ... (Stein, 2008) 1.51 2.29 0.6 2.1 S3.2 (continued) New **models** of atherosclerosis 3 Page 4. T able 1. Continued Quantitative comparison From literature ...

[HTML] A systems pharmacology **model** for inflammatory bowel disease
V Balbas-Martinez, L Ruiz-Cerdá, I Irurzun-Arana... - PloS one, 2018 - journals.plos.org
... More recently, Dwivendi et al., [21], based on the results of a **clinical** trial with the anti–IL6R antibody ... **models** enclosing temporal operators) to a standard Markup language in Systems Biology for qualitative **models** (SBML qual [35]) which promotes **model** interoperability, and ...

[HTML] Personalization of Logical **Models** With Multi-Omics Data Allows **Clinical** Stratification of Patients
J Béal, A Montagud, P Traynard, E Barillot... - Frontiers in ..., 2018 - ncbi.nlm.nih.gov
... The merging of the generic **model** with the patient profiles creates a personalized MaBoSS **model** with an unchanged BND file and a CFG file per patient. Then, **clinical** relevance of these patient-specific **models** can be assessed before providing original and ...

COVID-19 DiseaseMap: Application of models on disease exploration
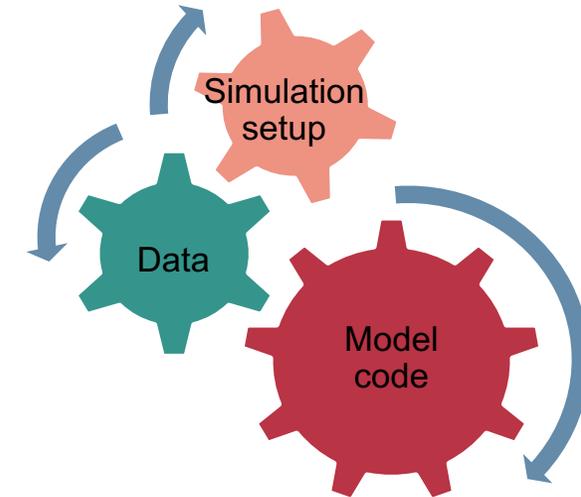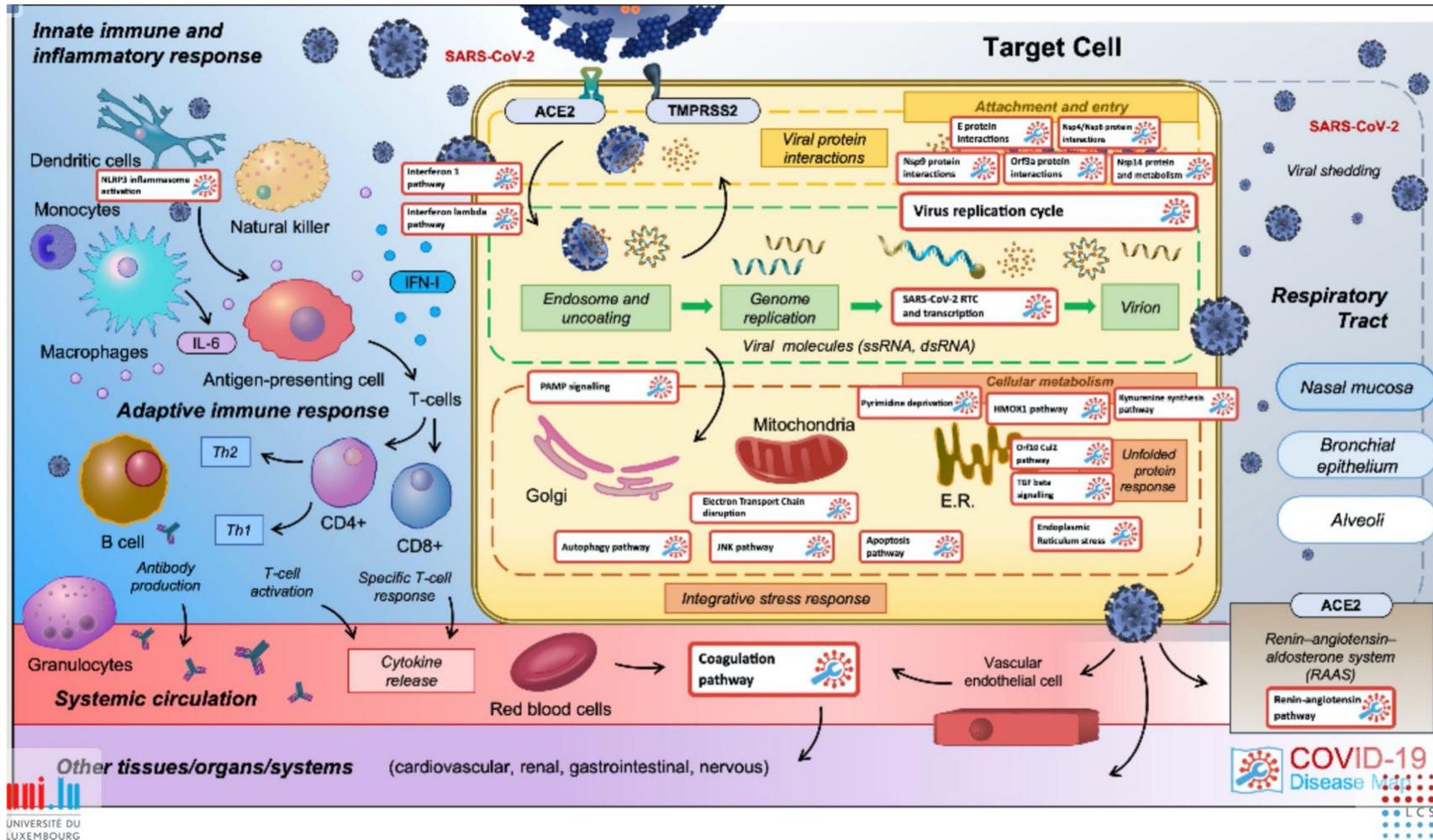
COmputational MOdeling in BIology NEtwork

combine-org.github.io

Resources
- https://combine-org.github.io/
- Specifications, guidelines, documentation
- Tool evaluation, show cases
- links to model- and code repositories

Recognised voice
- Journals
- Legal entities and organisations
- Funders
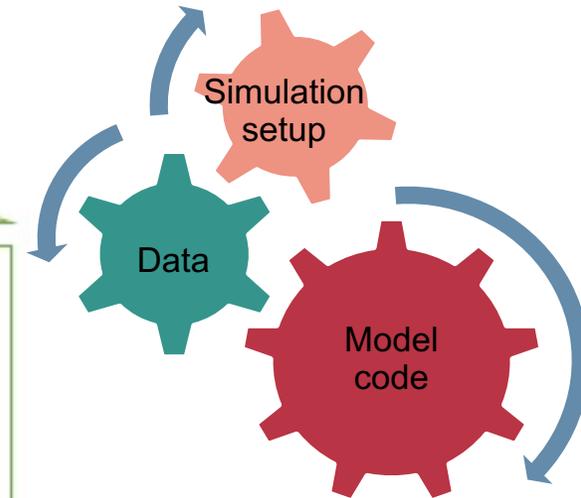- The public

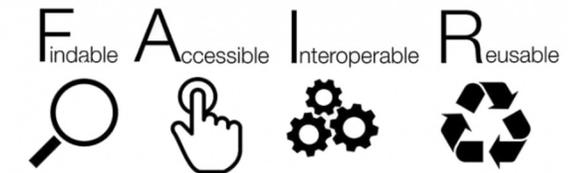**Community meetings**
- HARMONY (spring)
- COMBINE (autumn)

Coordination
- Cross-standard publications
- Communication and outreach
- Cross-standard developments of libraries, tools, and guidelines

Simulation setup

Data

Model code

Adherance to FAIR principles fosters reuse and interoperability

F indable   A ccessible   I nteroperable   R eusable
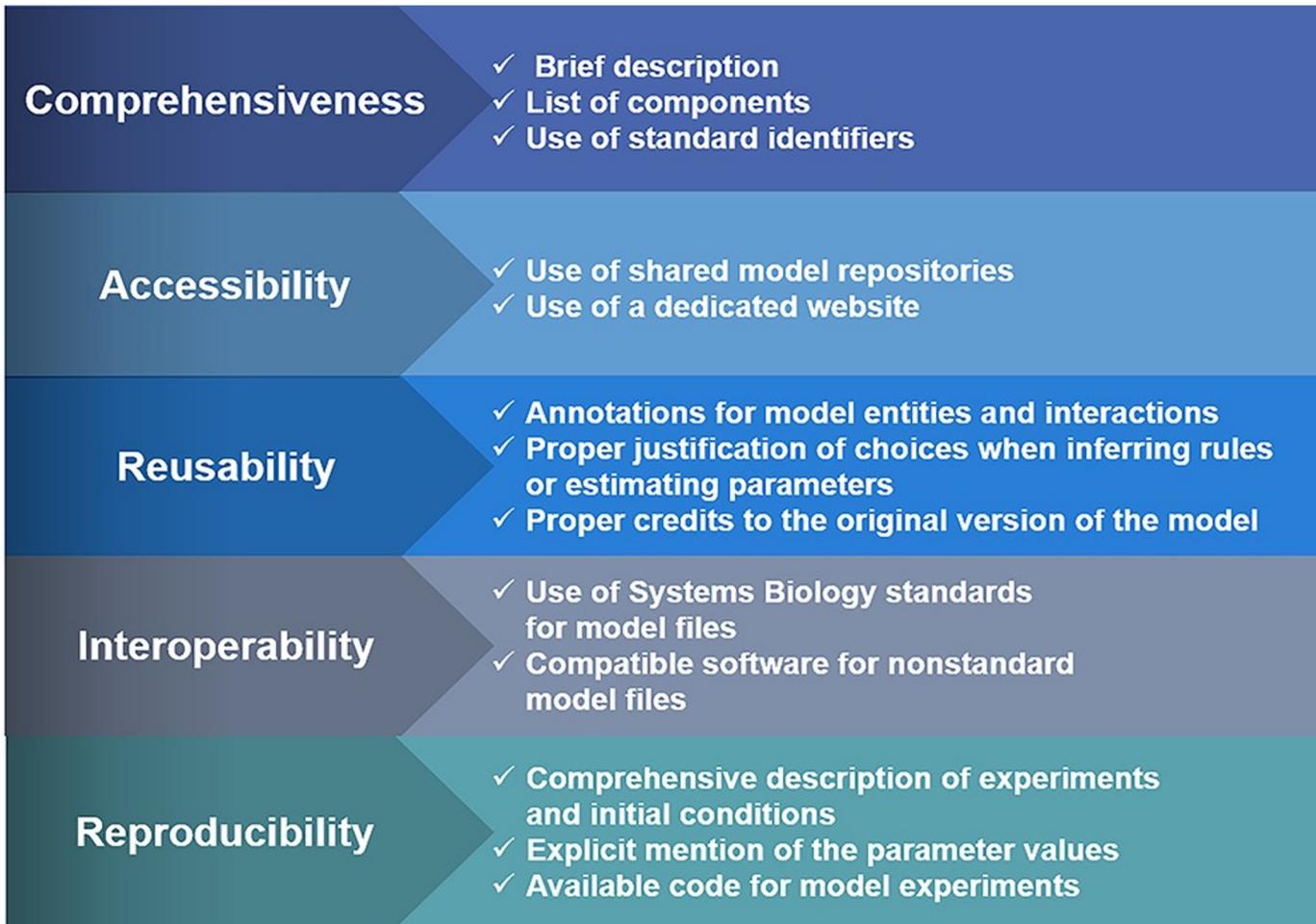
1. during modeling
2. during model curation
3. during publication
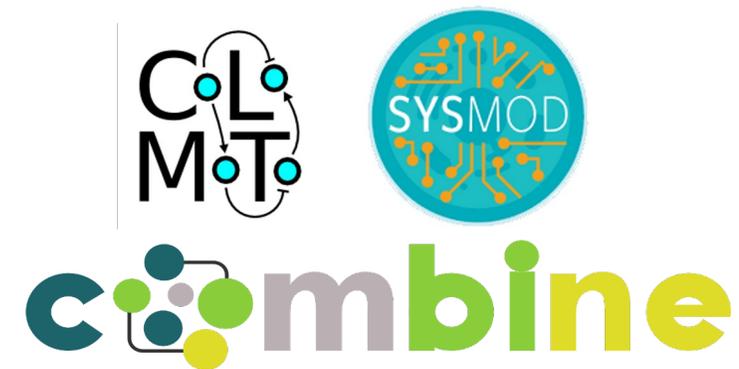
Main challenges to increase the impact of computational models:
Niarakis et al. (2022). DOI: 10.1093/bib/bbac212

Fostering the uptake of RDA indicators

Project goals:

1. To implement FAIR evaluation as a standard procedure during model curation/ development
2. To provide, together with the COMBINE core partners, FAIR model indicators for the Systems Biomedicine community
3. Semi-automatic FAIR evaluation tool for use by the community

	Sub-principle	ID	Indicator	Priority	Assessment overall	Assessment Essential	Assessment non-essential
Findable	F1	RDA-F1-01M	Metadata is identified by a persistent identifier	Essential	1	1	-
Findable	F1	RDA-F1-01D	Data is identified by a persistent identifier	Essential	1	1	-
Findable	F1	RDA-F1-02M	Metadata is identified by a globally unique identifier	Essential	1	1	-
Findable	F1	RDA-F1-02D	Data is identified by a globally unique identifier	Essential	1	1	-
Findable	F2	RDA-F2-01M	Rich metadata is provided to allow discovery	Essential	1	1	-
Findable	F3	RDA-F3-01M	Metadata includes the identifier for the data	Essential	1	1	-
Findable	F4	RDA-F4-01M	Metadata is offered in such a way that it can be harvested and indexed	Essential	1	1	-
Accessible	A1	RDA-A1-01M	Metadata contains information to enable the user to get access to the data	Important	1	-	1
Accessible	A1	RDA-A1-02M	Metadata can be accessed manually (i.e. with human intervention)	Essential	1	1	-
Accessible	A1	RDA-A1-02D	Data can be accessed manually (i.e. with human intervention)	Essential	1	1	-
Accessible	A1	RDA-A1-03M	Metadata identifier resolves to a metadata record	Essential	1	1	-
Accessible	A1	RDA-A1-03D	Data identifier resolves to a digital object	Essential	1	1	-
Accessible	A1	RDA-A1-04M	Metadata is accessed through standardised protocol	Essential	1	1	-
Accessible	A1	RDA-A1-04D	Data is accessible through standardised protocol	Essential	1	1	-
Accessible	A1	RDA-A1-05D	Data can be accessed automatically (i.e. by a computer program)	Important	1	-	1
Accessible	A1.1	RDA-A1.1-01M	Metadata is accessible through a free access protocol	Essential	1	1	-
Accessible	A1.1	RDA-A1.1-01D	Data is accessible through a free access protocol	Important	1	-	1
Accessible	A1.2	RDA-A1.2-01D	Data is accessible through an access protocol that supports authentication and authorisation	Useful	1	-	1
Accessible	A2	RDA-A2-01M	Metadata is guaranteed to remain available after data is no longer available	Essential	1	1	-
Interoperable	I1	RDA-I1-01M	Metadata uses knowledge representation expressed in standardised format	Important	0	-	0
Interoperable	I1	RDA-I1-01D	Data uses knowledge representation expressed in standardised format	Important	1	-	1
Interoperable	I1	RDA-I1-02M	Metadata uses machine-understandable knowledge representation	Important	1	-	1
Interoperable	I1	RDA-I1-02D	Data uses machine-understandable knowledge representation	Important	1	-	1
Interoperable	I2	RDA-I2-01M	Metadata uses FAIR-compliant vocabularies	Important	0	-	0
Interoperable	I2	RDA-I2-01D	Data uses FAIR-compliant vocabularies	Useful	0	-	0
Interoperable	I3	RDA-I3-01M	Metadata includes references to other metadata	Important	1	-	1
Interoperable	I3	RDA-I3-01D	Data includes references to other data	Useful	0	-	0
Interoperable	I3	RDA-I3-02M	Metadata includes references to other data	Useful	1	-	1
Interoperable	I3	RDA-I3-02D	Data includes qualified references to other data	Useful	0	-	0
Interoperable	I3	RDA-I3-03M	Metadata includes qualified references to other				
Interoperable	I3	RDA-I3-04M	Metadata include qualified references to other				
Reusable	R1	RDA-R1-01M	Plurality of accurate and relevant attributes are				
Reusable	R1.1	RDA-R1.1-01M	Metadata includes information about the licenc				
Reusable	R1.1	RDA-R1.1-02M	Metadata refers to a standard reuse licence				
Reusable	R1.1	RDA-R1.1-03M	Metadata refers to a machine-understandable				
Reusable	R1.2	RDA-R1.2-01M	Metadata includes provenance information acc				
Reusable	R1.2	RDA-R1.2-02M	Metadata includes provenance information acc				
Reusable	R1.3	RDA-R1.3-01M	Metadata complies with a community standard				
Reusable	R1.3	RDA-R1.3-01D	Data complies with a community standard				
Reusable	R1.3	RDA-R1.3-02M	Metadata is expressed in compliance with a m				
Reusable	R1.3	RDA-R1.3-02D	Data is expressed in compliance with a machin				

Legend – indicator priority:
Essential
Important
Useful

	Priority	Assessment overall	Assessment Essential	Assessment non-essential
	Sum	30	18	12
Score total = # fullfilled / # total (in relevant category)	Score total	73.10%	90%	57.10%
Score applicable = # fullfilled / (# total - # NA) (in relevant category)	Score applicable	73.10%	90%	57.10%

Adapted from RDA indicators (DOI: 10.15497/rda00045)

# A method for semi-automated FAIR assessment

Project goals:
1. To implement FAIR evaluation as a standard procedure during model curation/ development
2. To provide, together with the COMBINE core partners, FAIR model indicators for the Systems Biomedicine community
3. Semi-automatic FAIR evaluation tool for use by the community

Workshop on "FAIRness assessment of COMBINE archives", COMBINE 2022, Berlin, Oct. 7, 2022

FAIR principle	COMBINE indicator	Description	Score
F1	CA-F1.2	Each version of the COMBINE archive has a persistent and globally unique identifier/DOI.	✗
I1	CA-I1.1	The COMBINE archive is valid.	✓

Interested? Join us!
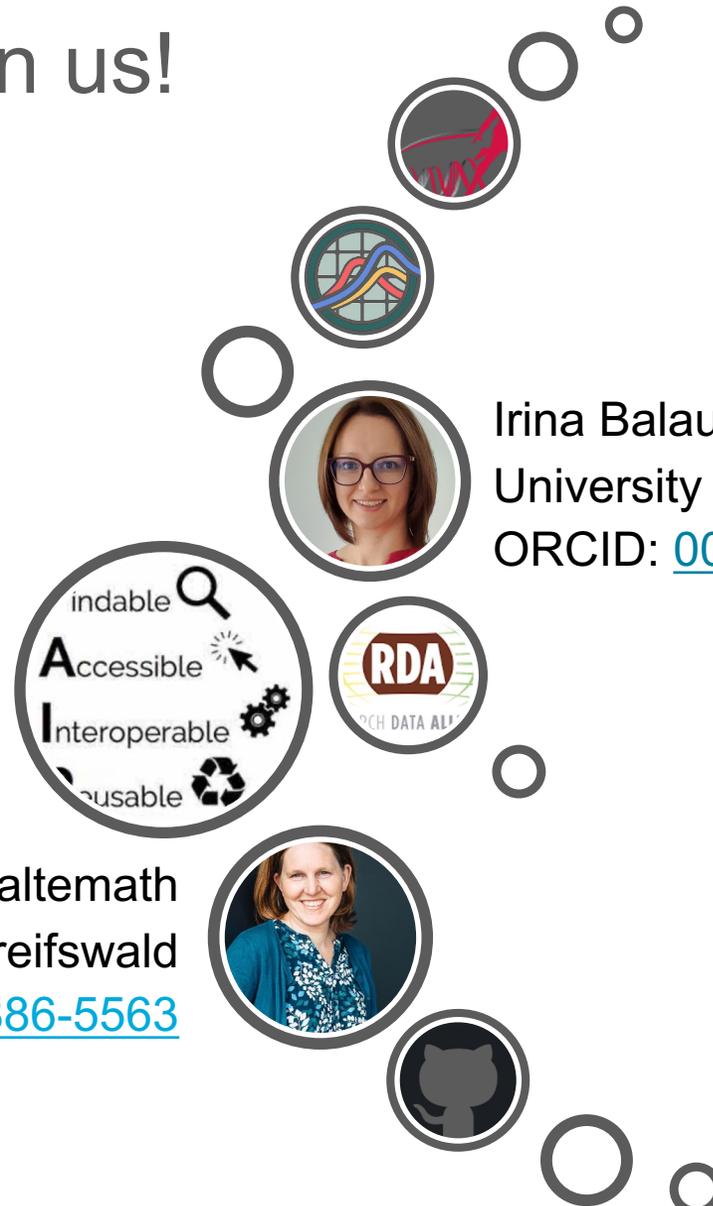
Irina Balaur
University of Luxembourg
ORCID: 0000-0002-3671-895X

Dagmar Waltemath
University Medicine Greifswald
ORCID: 0000-0002-5886-5563

Project website:
https://fair-ca-indicators.github.io/

9