FAIRDOM approach for semantic interoperability of systems biology data and models

Olga Krebs[1*], Katy Wolstencroft [3], Natalie J Stanford[2] , Martin Golebiewski[1], Stuart Owen [2], Quyen Nguyen[1], Peter Kunszt[5], Bernd Rinn[6], Jacky L. Snoep [2,4], Wolfgang Müller[1] and Carole Goble [2]

* presenting authors

Findable

Accessible

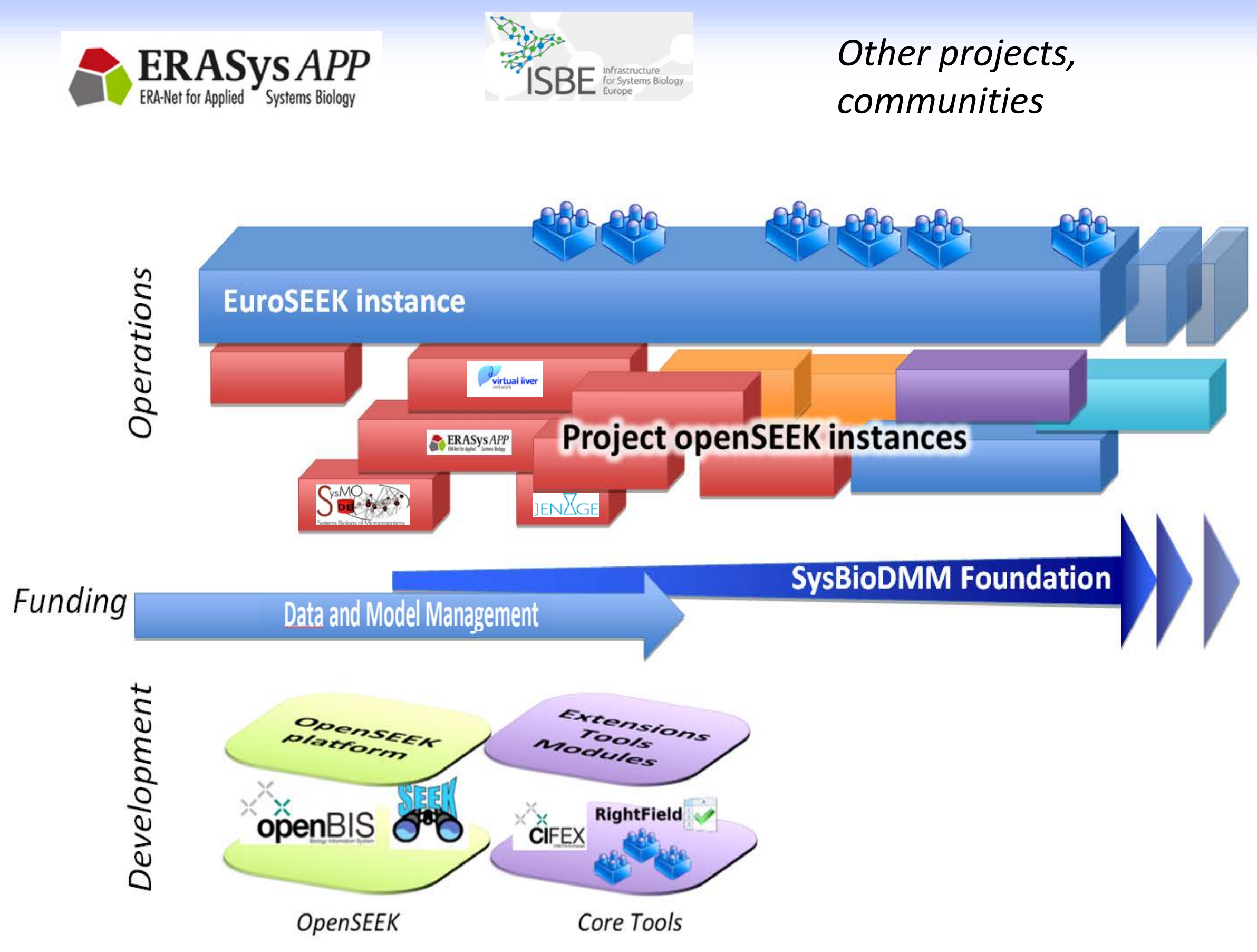
Interoperable

Reusable

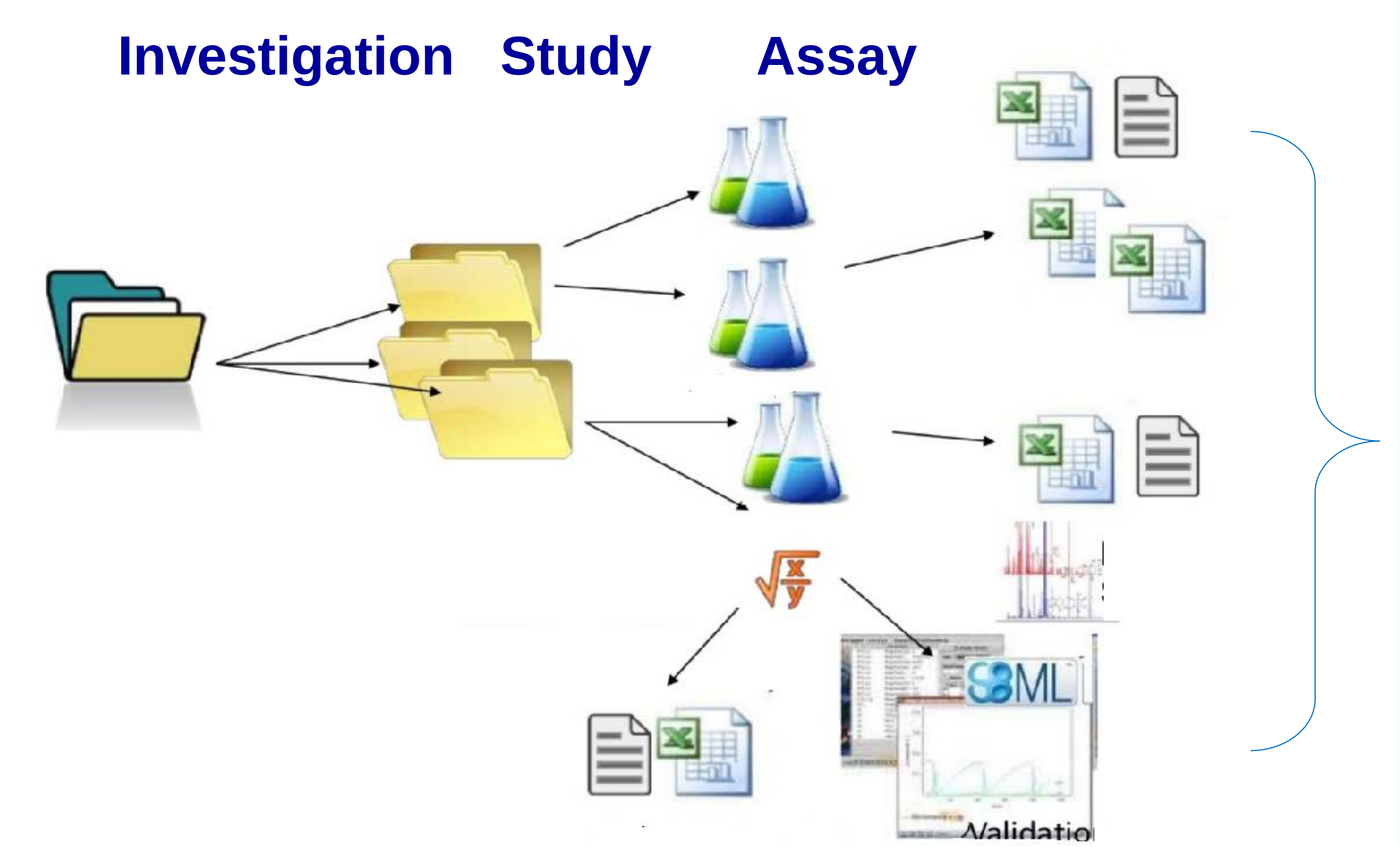
Data

Operations

Models

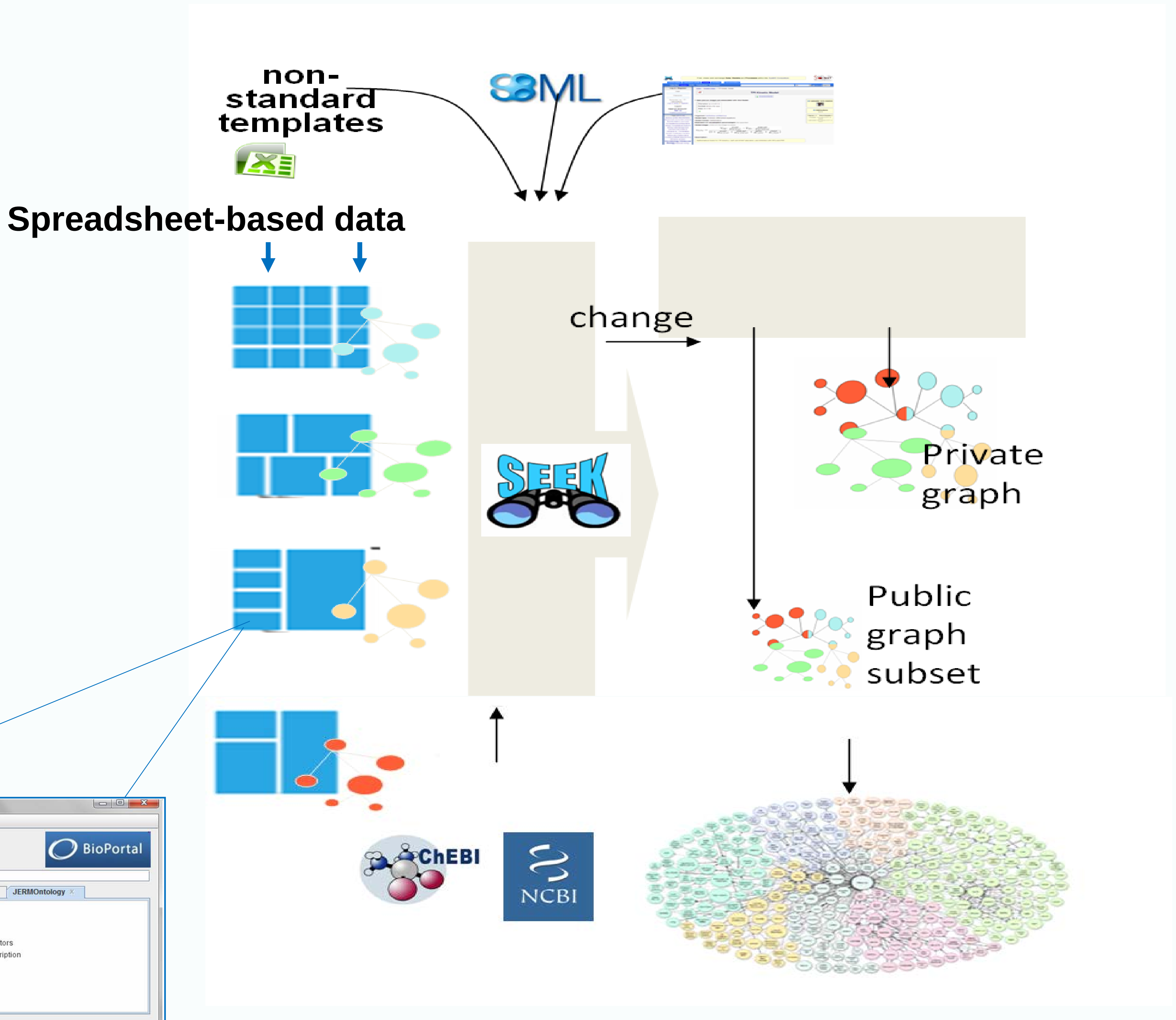


FAIRDOM will establish a support and service network for European Systems Biology Projects. This joint action aims at establishing an internationally sustained Data and Model Management service, initially supporting research projects funded by ERA-Net for Systems Biology Applications (ERASysAPP), and systematically extending the services to the European systems biology community.
This is not a new realization, in FAIRDOM we are building on the experience and expertise of two leading systems biology management platforms: SEEK, mainly developed in the context of SysMO-DB project and openBIS, mainly developed in the context of the SystemsX.ch SyBIT project.



The Investigation, Studies, Assay (ISA) framework is used so that registered assets can be interlinked to provide descriptions, scientific context, relationships between assets

Just Enough Results Model (JERM)

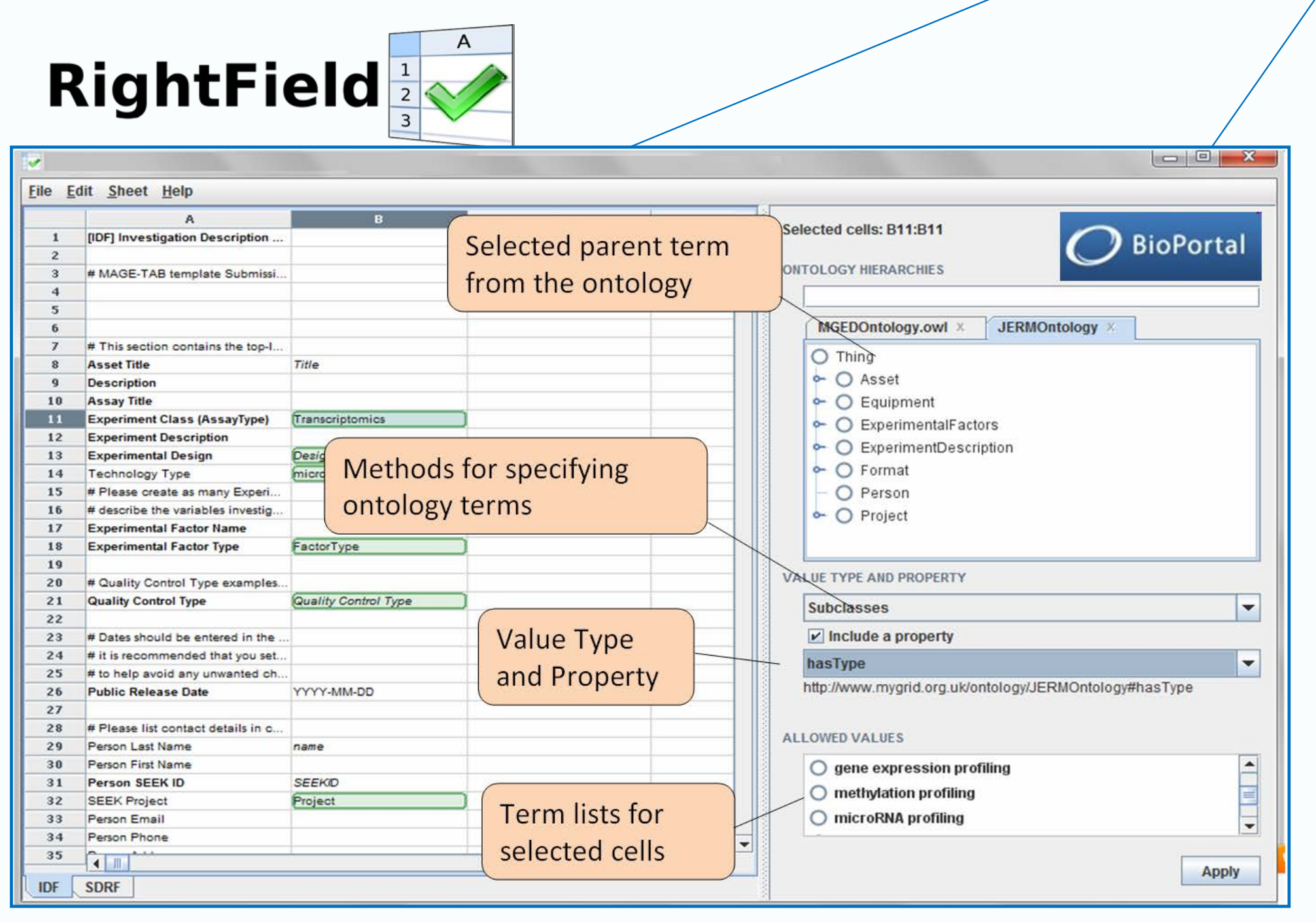


JERM – based templates

Each spreadsheet has multiple sheets, defining experimental metadata, organims/samples, data and optional sheets for instrument descriptions and results.

RightField is a tool for embedding ranges of ontology terms into spreadsheets to allow the users of those spreadsheets to semantically annotate their data from simple drop-down lists.

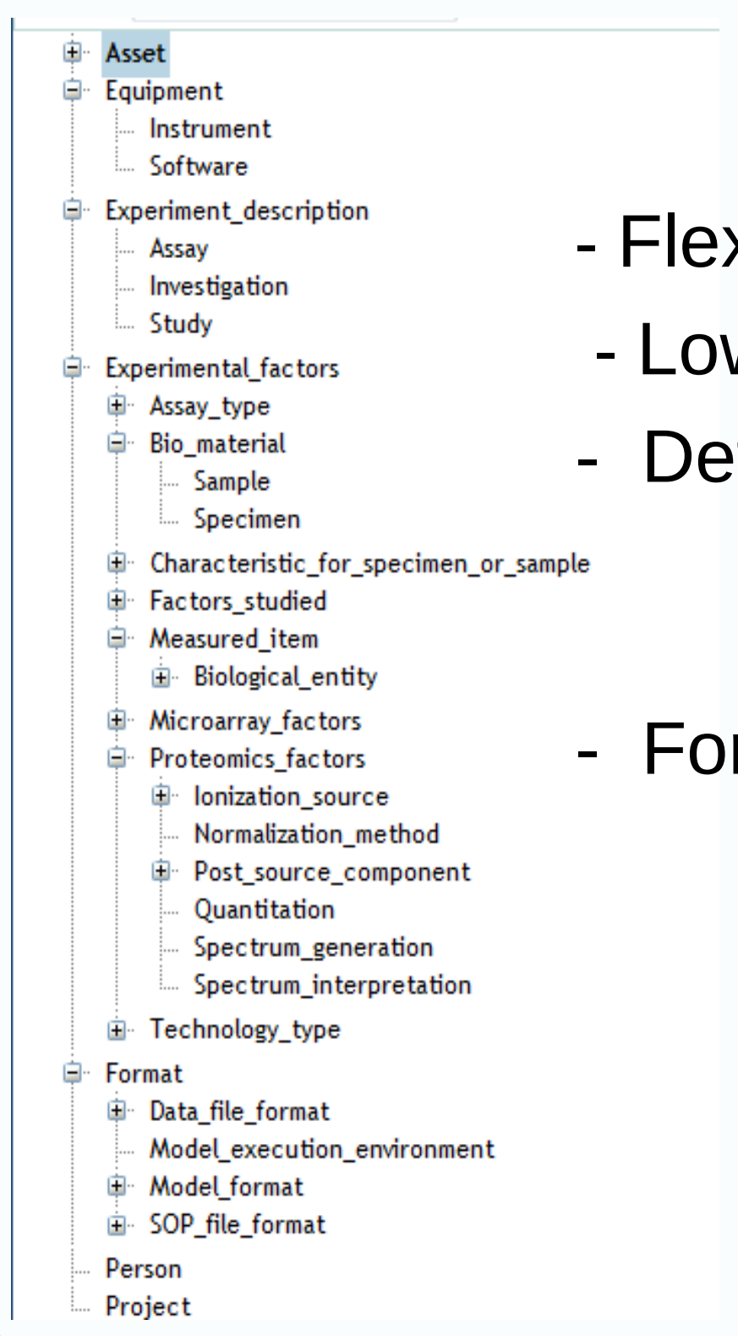
RightField



RDF Generation

- RDF generated from relational db, file content and from RightField marked-up cells
- Community ontologies: SOIC, FOAF, DC, MGED…
- Common identifiers: CheBI ids for substances, NCBI ids for organisms, etc
- Sparql endpoint, canned queries and API
- Authorization of the individual triples generated from SEEK
- RDF generated split into two graphs:
  - *- all information*
  - *- public information*
- Object links in the public graph pointing to a private resource blocked as unauthorized when attempt to access the resource

JERM – Ontology



- Asset
- Equipment
  - Instrument
  - Software
- Experiment_description
  - Assay
  - Investigation
  - Study
- Experimental_factors
  - Assay_type
  - Bio_material
    - Sample
    - Specimen
  - Characteristic_for_specimen_or_sample
  - Factors_studied
  - Measured_item
    - Biological_entity
  - Microarray_factors
  - Proteomics_factors
    - Ionization_source
    - Normalization_method
    - Post_source_component
    - Quantitation
    - Spectrum_generation
    - Spectrum_interpretation
  - Technology_type
- Format
  - Data_file_format
  - Model_execution_environment
  - Model_format
  - SOP_file_format
- Person
- Project

- Flexibly represent different types of data
- Lower barrier of semantic annotation
- Define relationships, cross-link, aggregate
  - *--data, models, results, samples, protocols, standard operating procedures, publications*
- For the experiments
  - *--reuse community ontologies, markups, mim, identifiers*
  - *-- MGED/EFO, PSI-MS, SBO, CheBI ids*

http://bioportal.bioontology.org/ontologies/JERM expressed in OWL

References and links:

FAIIRDOM	http://fair-dom.org/
EraSysApp	https://www.erasysapp.eu/
SEEK	http://www.seek4science.org/
OpenBis	http://www.cisd.ethz.ch/software/openBIS/
RightField	http://www.rightfield.org.uk/
ISA	http://isacommons.org/
SyBIT	http://www.sybit.net/

Funding:

Affiliations: [1] Heidelberg Institute for Theoretical Studies, Germany
[2] School of Computer Science, University of Manchester, UK
[3] Leiden Insitute of Advanced Computer Science, NL
[4] University of Stellenbosch, SA
[5] University of Zurich, Swiss
[6] ETH Zurich, Swiss