

A PROPOSED DATA STANDARD FOR BREATH SAMPLE DATA AND METADATA

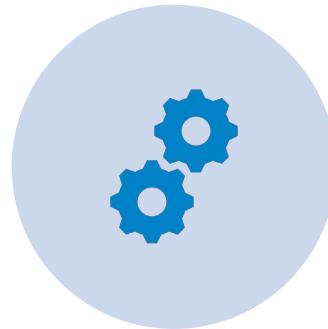
L

BO ZHAO

BACKGROUND



Metabolomics data
standards



Reproducibility



Quality assurance

POPULAR METABOLOMICS DATA STANDARDS



ISA Tools

ISA-Tab
ISA JSON



Metabolomics Workbench

mwTab



HUPO Proteomics Standards Initiative

mzTab
mzTab-M

A GOOD DATA STANDARD

Readability

- Human readable
- Machine processable

Usability

- Ease of use

Extendibility

- Configurable design

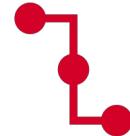


EXTENDING ISA



ISA data model

Diverse data support
Rich metadata



Enhanced relationship

Chain of processes



Quality control

Validated results

ENHANCED RELATIONSHIP



Ontology based

- Ontology annotations in ISA data model

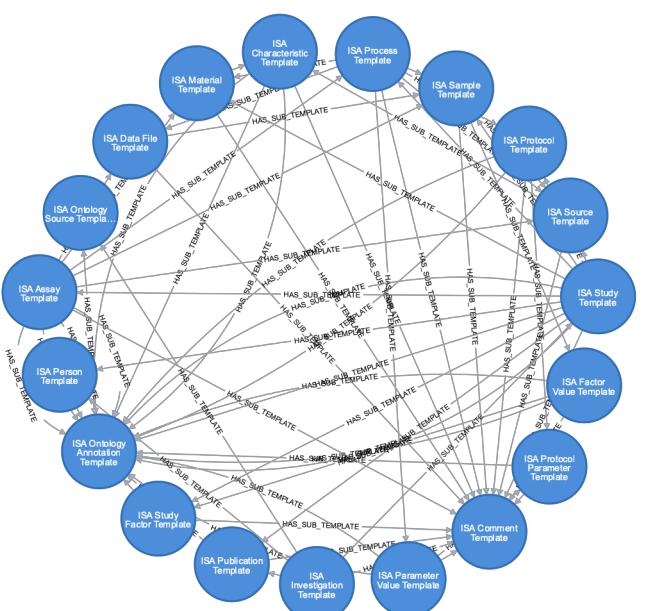
Pre-defined

- Model-based built-in relationship definitions

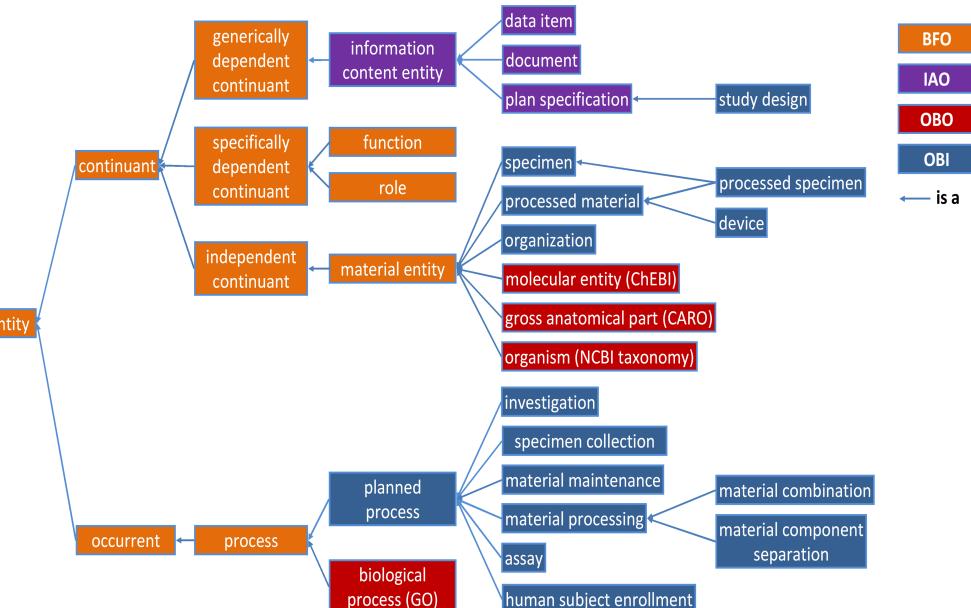
User defined

- Study-specific configurations

ENHANCED RELATIONSHIP

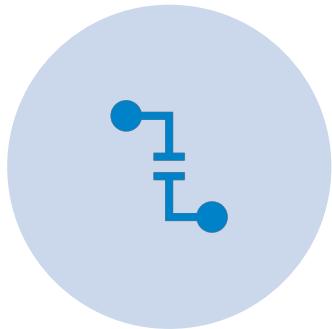


Low-level ISA components relationships



Partial high-level structure of OBI classes (Bandrowski et al. 2016)

ENHANCED QUALITY CONTROL



Extensively applied



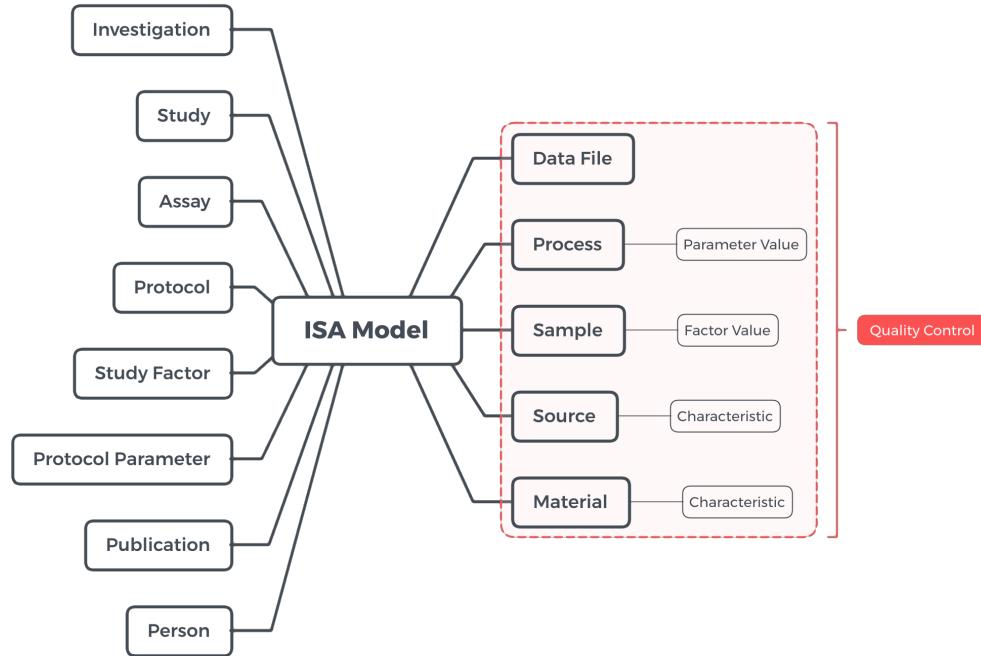
Customisable approaches
(in development)



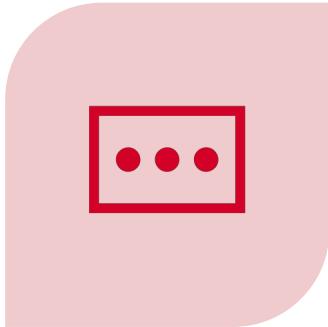
Built-in support for basic
statistical validations

ENHANCED QUALITY CONTROL

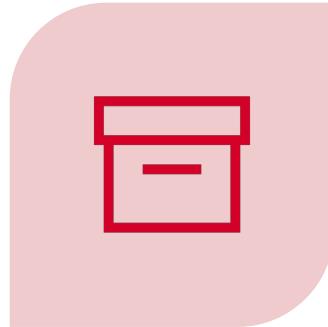
Component level verification
Configurable in study design
Quality control result file



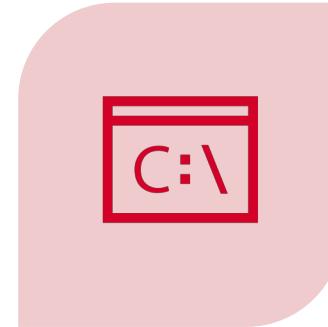
DATA MANAGEMENT TOOL



Local installation or on a
server

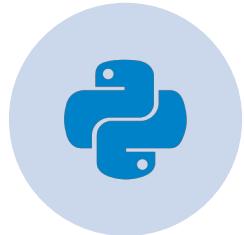


Compatible with original ISA
archives



Command line and user-
friendly interfaces

TOOL FEATURES



Written in python



Easy to install



Easy to embed



Guided access

USER INTERFACE – TEMPLATE BROWSER

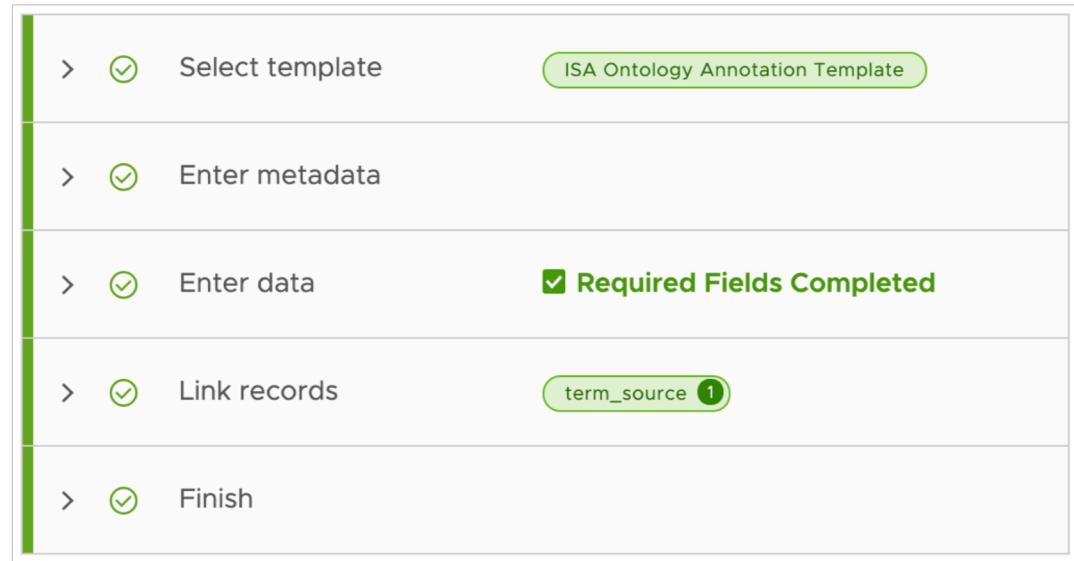
Type	Description	Used in Records	Child Templates	Actions
> ISA Factor Value Template		0	4	BROWSE RECORDS CREATE RECORD EXTEND TEMPLATE
< ISA Ontology Annotation Template		1	1	BROWSE RECORDS CREATE RECORD EXTEND TEMPLATE

Child templates

Merge field	Child template	Multiple instances
term_source	ISA Ontology Source Template	<input checked="" type="checkbox"/> No

USER INTERFACE – GUIDED RECORD CREATION

1. Template
2. Metadata
3. Data
4. Link
5. Review



USER INTERFACE – GUIDED RECORD CREATION

1. Template
2. Metadata
3. Data
4. Link
5. Review

▼ Select template ISA Ontology Annotation Template

Search for template

Minimum 3 characters required.

ISA compatible level

Exact Compatible Any

You have selected:

ISA Ontology Annotation Template

[NEXT](#)

USER INTERFACE – GUIDED RECORD CREATION

1. Template
2. Metadata
3. Data
4. Link
5. Review

▼ Enter metadata

Identifier

Required field. Must be unique.

Display

Description

NEXT

USER INTERFACE – GUIDED RECORD CREATION

1. Template
2. Metadata
3. Data
4. Link
5. Review

▼ Enter data Required Fields Completed

term
Lobophora rosacea

term_accession
http://purl.bioontology.org/ontology/BFO/BFO_0000001

NEXT

USER INTERFACE – GUIDED RECORD CREATION

1. Template
2. Metadata
3. Data
4. Link
5. Review

Link records term_source 1

term_source 1

LINK SELECTED LINK NONE

Record	Description	Value
<input type="radio"/> OntologySource Node 0	This is an example ontology source node. ID: [0]	{ "description": "OS description 0.", "file": "OS file 0.", "name": "OSName0", "version": "OS version 0." }
<input type="radio"/> OntologySource Node 9	This is an example ontology source node. ID: [9]	{ "description": "OS description 9.", "file": "OS file 9.", "name": "OSName9", "version": "OS version 9." }

items per page 5 1 - 2 of 2 items

NEXT

USER INTERFACE – GUIDED RECORD CREATION

1. Template
2. Metadata
3. Data
4. Link
5. Review

Finish

Please review your actions before continue.

Template

ISA Ontology Annotation Template

Metadata

term	Lobophora rosacea
term_accession	http://purl.bioontology.org/ontology/NCBITAXON/1638598

Records to link

term_source	1
-------------	---

FINISH

USER INTERFACE - REPORT

Study Information

Identifier	EMBER
Name	East Midlands Breathomics Pathology Node
Description	The East Midlands Breathomics Pathology Node (EMBER) is a Medical Research Council (MRC) & Engineering and Physical Sciences Research Council (EPSRC) funded Molecular Pathology Node hosted at the University of Leicester partnering with University Hospitals of Leicester NHS Trust, Loughborough University and Industry.
Publications	PMID 30852546 PMID 30755317
Contacts	Chris Brightling Salman Siddiqui
Submission	2015-09-01
Release	2019-09-09

Sample Stats

Sources	Specimen Type	Samples	Locations
400	Exhaled air specimen	1000	Glenfield Hospital Leicester Royal Infirmary



USER INTERFACE - REPORT

Protocol Information

Only showing protocols related to selected samples

Name	Version	Components	Parameters
ReCIVA Kits Preparation	2	Consumable TD Tube Consumable Mask	Seal Status Yes
ReCIVA Breath Collection	1	Instrument ReCIVA Consumable TD Tube Consumable Mask	Samples 1 Minimum air flow (L/min) 30 Minimum collection volume (%) 50 Maximum collection duration (second) 900

Sample Detail

EMBER-E0001-R-RP

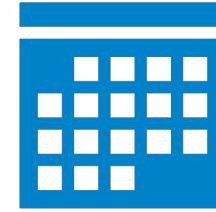
Processes	Performer	Date	Parameters	Quality Control
ReCIVA Kits Preparation	Bo Zhao	2019-09-09	Seal Status Yes	Seal Check Pass
ReCIVA Breath Collection	Jane Doe	2019-09-09	Samples 1 Air flow (L/min) 33 Collection volume (%) 100 Collection duration (second) 600	Sample Count Pass Minimum air flow Pass Minimum collection volume Pass Maximum collection duration Pass

AVAILABILITY



GitHub

<https://github.com/rcfgroup>



Later this year

OTHER WORKS

- LabPipe
- Ontology for Breath Sample Process

ACKNOWLEDGEMENT

- The research is funded by the EMBER project and NIHR Leicester Biomedical Research Centre – Respiratory.