

This presentation: 10.5281/zenodo.3258123

Standardization efforts around deposition of isotopic tracer data

ELIXIR Implementation Study on Standardizing fluxomics workflows

Philippe Rocca-Serra (<http://orcid.org/0000-0001-9853-5668>)

University of Oxford

June 27th-28th, 2019 The Hague,

Twitter: @Phil_at_OeRC

Marta Cascante, Vitaly Selivanov, Pedro de Atauri, Kenneth Haug, Reza Salek, Claire o'Donovan, Chris Steinbeck, Hunter Moseley, Teresa Fan,
Andrew Lane

Outline

- Bringing awareness to existing resources
 - Request for extending IUPAC InChi specification to support isotopologue reporting
 - (Hunter Moseley)
 - A controlled Terminology: MSIO
 - A syntax for describing experimental setup: ISA
 - Used by EMBL-EBI Metabolights Repository
 - Fluxomics use case require specific handling
 - Dedicated ISA-Configurations reflecting annotation needs formulated by the community
 - Dedicated Data Matrix Formats:
 - Option 1: derived from MAF file for compatibility with ISAcreator plugin
 - Validation
 - Layout: wide table + fixed height (i.e. fixed number of molecular dimensions)
 - Option 2: Exploratory JSON Tabular Data packages
 - Layout: long table + fixed width (i.e. fixed number of measurement dimensions)

SIRM & flux studies != Metabolite profiling studies

- A lot in common for acquisition techniques (MS, NMR)
- BUT
 - labeling experiments require specific descriptors
 - analytical tools differ
 - Reporting results require specific descriptors

Analysis: ESI/QTOF negative ion mode

Show values for a selected metabolite or ratios for 2 selected metabolites

Select	Metabolite Name	Refmet Name	WorkBench Metabolite_ID	PubChem Compound_ID	Kegg Id	Retention time/index	m/z ratio	Other Id	Other Id Type
<input type="checkbox"/>	2 or 3-phosphoglycerate_13C_1	3-Phosphoglyceric acid	ME087081	724 ↗	C00597 ↗	-	-	2PG/3PG_13C_1	UM_Target_ID
<input type="checkbox"/>	2 or 3-phosphoglycerate_13C_2	3-Phosphoglyceric acid	ME087082	724 ↗	C00597 ↗	-	-	2PG/3PG_13C_2	UM_Target_ID
<input type="checkbox"/>	2 or 3-phosphoglycerate_13C_3	3-Phosphoglyceric acid	ME087083	724 ↗	C00597 ↗	-	-	2PG/3PG_13C_3	UM_Target_ID
<input type="checkbox"/>	2 or 3-phosphoglycerate_2PG/3PG	3-Phosphoglyceric acid	ME087080	724 ↗	C00597 ↗	-	-	2PG/3PG	UM_Target_ID
<input type="checkbox"/>	3-Glycerophosphate	3-Phosphoglyceric acid	ME087146	754 ↗	-	-	-	GI_OH3P	UM_Target_ID
<input type="checkbox"/>	3-Glycerophosphate_13C_1	3-Phosphoglyceric acid	ME087147	754 ↗	-	-	-	GI_OH3P_13C_1	UM_Target_ID
<input type="checkbox"/>	3-Glycerophosphate_13C_2	3-Phosphoglyceric acid	ME087148	754 ↗	-	-	-	GI_OH3P_13C_2	UM_Target_ID
<input type="checkbox"/>	3-Glycerophosphate_13C_3	3-Phosphoglyceric acid	ME087149	754 ↗	-	-	-	GI_OH3P_13C_3	UM_Target_ID
<input type="checkbox"/>	3-phosphoglyceraldehyde	Glyceraldehyde-3-phosphate	ME087135	439168 ↗	C00118 ↗	-	-	G3P	UM_Target_ID
<input type="checkbox"/>	3-phosphoglyceraldehyde_13C_1	Glyceraldehyde-3-phosphate	ME087136	439168 ↗	C00118 ↗	-	-	G3P_13C_1	UM_Target_ID
<input type="checkbox"/>	3-phosphoglyceraldehyde_13C_2	Glyceraldehyde-3-phosphate	ME087137	439168 ↗	C00118 ↗	-	-	G3P_13C_2	UM_Target_ID
<input type="checkbox"/>	3-phosphoglyceraldehyde_13C_3	Glyceraldehyde-3-phosphate	ME087138	439168 ↗	C00118 ↗	-	-	G3P_13C_3	UM_Target_ID
<input type="checkbox"/>	6-Phosphogluconate	6-Phosphogluconic acid	ME087084	91493 ↗	C00345 ↗	-	-	6PG	UM_Target_ID
<input type="checkbox"/>	6-Phosphogluconate_13C_1	6-Phosphogluconic acid	ME087085	91493 ↗	C00345 ↗	-	-	6PG_13C_1	UM_Target_ID
<input type="checkbox"/>	6-Phosphogluconate_13C_2	6-Phosphogluconic acid	ME087086	91493 ↗	C00345 ↗	-	-	6PG_13C_2	UM_Target_ID
<input type="checkbox"/>	6-Phosphogluconate_13C_3	6-Phosphogluconic acid	ME087087	91493 ↗	C00345 ↗	-	-	6PG_13C_3	UM_Target_ID
<input type="checkbox"/>	6-Phosphogluconate_13C_4	6-Phosphogluconic acid	ME087088	91493 ↗	C00345 ↗	-	-	6PG_13C_4	UM_Target_ID
<input type="checkbox"/>	6-Phosphogluconate_13C_5	6-Phosphogluconic acid	ME087089	91493 ↗	C00345 ↗	-	-	6PG_13C_5	UM_Target_ID
<input type="checkbox"/>	6-Phosphogluconate_13C_6	6-Phosphogluconic acid	ME087090	91493 ↗	C00345 ↗	-	-	6PG_13C_6	UM_Target_ID
<input type="checkbox"/>	Acetyl-CoA	Acetyl-CoA	ME087091	444493 ↗	C00024 ↗	-	-	aCoA	UM_Target_ID
<input type="checkbox"/>	Acetyl-CoA_13C_1	Acetyl-CoA	ME087092	444493 ↗	C00024 ↗	-	-	aCoA_13C_1	UM_Target_ID
<input type="checkbox"/>	Acetyl-CoA_13C_2	Acetyl-CoA	ME087093	444493 ↗	C00024 ↗	-	-	aCoA_13C_2	UM_Target_ID
<input type="checkbox"/>	Adenosine diphosphate_13C_1	ADP	ME087095	6022 ↗	C00008 ↗	-	-	ADP_13C_1	UM_Target_ID
<input type="checkbox"/>	Adenosine diphosphate_13C_2	ADP	ME087096	6022 ↗	C00008 ↗	-	-	ADP_13C_2	UM_Target_ID
<input type="checkbox"/>	Adenosine diphosphate_13C_3	ADP	ME087097	6022 ↗	C00008 ↗	-	-	ADP_13C_3	UM_Target_ID
<input type="checkbox"/>	Adenosine diphosphate_13C_4	ADP	ME087098	6022 ↗	C00008 ↗	-	-	ADP_13C_4	UM_Target_ID
<input type="checkbox"/>	Adenosine diphosphate_13C_5	ADP	ME087099	6022 ↗	C00008 ↗	-	-	ADP_13C_5	UM_Target_ID
<input type="checkbox"/>	Adenosine monophosphate_13C_1	AMP	ME087101	6083 ↗	C00020 ↗	-	-	AMP_13C_1	UM_Target_ID
<input type="checkbox"/>	Adenosine monophosphate_13C_2	AMP	ME087102	6083 ↗	C00020 ↗	-	-	AMP_13C_2	UM_Target_ID
<input type="checkbox"/>	Adenosine monophosphate_13C_3	AMP	ME087103	6083 ↗	C00020 ↗	-	-	AMP_13C_3	UM_Target_ID
<input type="checkbox"/>	Adenosine monophosphate_13C_4	AMP	ME087104	6083 ↗	C00020 ↗	-	-	AMP_13C_4	UM_Target_ID
<input type="checkbox"/>	Adenosine monophosphate_13C_5	AMP	ME087105	6083 ↗	C00020 ↗	-	-	AMP_13C_5	UM_Target_ID

Issue 1: Molecular Identity

- **Inchi Isotopologue and Isotopomer Proposal**

<https://github.com/MSI-Metabolomics-Standards-Initiative/inchi-isotopologue-extension>

Molecular Identification: Extension to Inchi Proposal

MSI-Metabolomics-Standards-Initiative / inchi-isotopologue-extension

Unwatch 3 Star 0 Fork 1

Code Issues 0 Pull requests 2 Projects 0 Wiki Security Insights

specification extension to InChi to better support isotopologue reporting

inchi inchi-standard inchi-string Manage topics

10 commits 1 branch 0 releases

Branch: master New pull request

Create new file Upload files

hunter-moseley Fixed formatting.

.gitignore further fixes

README.md Fixed formatting.

README.md

inchi-isotopologue-extension

specification extension to InChi to better support isotopologue reporting

Purpose:

Develop enhanced specifications within the regular InChI standard for representing isotopologues and isotopomers. More specifically, augment the isotopic layer specifications of the regular InChI standard so that specific isotopologues, isotopomers, partial isotopomers, and isotopologue fragments can be represented by a single InChI string and used to identify isotope-informative analytical features.

Moseley, Hunter; Salek, Reza; Arita, Masanori; Schymanski, Emma; Rocca-Serra, Philippe (2018): InChI Isotopologue and Isotopomer Proposal. figshare. Dataset.

<https://doi.org/10.6084/m9.figshare.7150964.v1>

InChI Isotopologue and Isotopomer Proposal

Dataset posted on 30.09.2018, 21:20 by [Hunter Moseley](#), [Reza Salek](#), [Masanori Arita](#), [Emma Schymanski](#), [Philippe Rocca-Serra](#)

Proposal for an extension to the IUPAC International Chemical Identifier (InChI) that supports the representation of isotopologues, isotopomers, partial isotopomers, and isotopologue fragments.

- <https://doi.org/10.6084/m9.figshare.7150964.v1>

Isotopically-Resolved Isotopologue Specification:

In the following proposed specifications, isotopologue designations are indicated by parentheses with the *element* identified first, while isotopomer designations lack parentheses with the *atom position* identified first: e.g. /a(C2+1) for isotopologue and /i7+1 for isotopomer. The lack of parentheses for isotopomers ensures backwards compatibility with previous InChI specifications. See individual specifications below for details.

Isotopically-Resolved Isotopologue Specification:

Rationale: to enable the use of InChI to specify an isotopologue or isotopologue fragment.

Simple Definition: /a(Ee#<+|->#[, #]...)

Complete Definition:

/a(<element><isotope_count><isotope_designation>[, <atom_number>])

<element> - one or two letter Element code (Ee).

<isotope_count> - number of atoms with the designated isotope (#).

<isotope_designation> - isotope designation indicated by a sign (+ or -) and number indicating the unit mass difference from the rounded average atomic mass of the element. For example, the average atomic mass of Sn (118.710) is rounded to 119. We specify two ¹¹⁸Sn atoms as "/a(Sn2-1)".

<atom_number> - specific atom(s) that may have the specified isotope. If empty, then all atoms of the given element may have the specified isotope (# in [, #]).

If multiple atoms need to be specified, these are comma separated.

Regular Expression Definition: [/]a([A-Z][a-z]?[d+[-]\d+([,]\d+)*])

Examples:

1. ¹³C₂ isotopologue of alpha-D-glucopyranose:

InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/a(C2+1)

2. ¹³C₂²H₃ isotopologue of alpha-D-glucopyranose:

InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/a(C2+1),(H3+1)

3. ¹⁷O₂¹⁶O₁ isotopologue of alpha-D-glucopyranose:

InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/a(O2+1),(O1+2)

Augmented Range Specification:

Augmented Range Specification:

Rationale: to enable a more concise string representation of isotopomer and isotopologues. Note in the isotopologue definition, the first sign and following number indicates the isotope, whereas in the isotopomer definition the last sign and following number indicates the isotope.

Simple Isotopologue Definition: /a(Ee#<+|->#[,-#|#]...)

Complete Isotopologue Definition:

/a(<element><isotope_count><isotope_designation>[,<start_atom_number>-<end_atom_number>])

Simple Isotopomer Definition: /i#-#<+|->#

The augmented isotopomer range specification is limited to a single atom_range and isotope_designation, due to parsing ambiguities that would arise if mixed atom_range and single atom designations were allowed.

Complete Isotopomer Definition:

/i<start_atom_number>-<end_atom_number><isotope_designation>

Examples:

- $^{13}\text{C}_2$ limited to atoms 4,5,6 isotopologue fragment of alpha-D-glucopyranose:
InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/a(C2+1,4-6)
- $^{13}\text{C}_8$ isotopomer of alpha-D-glucopyranose:
InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/i1-6+1

Nominal-Mass Isotopologue Specification:

Nominal-Mass Isotopologue Specification:

Rationale: to enable an InChI representation for mass spectral features that are not isotope-resolved.

Simple Nominal-Mass Isotopologue Definition: /a(#n[,#]...)

Complete Nominal-Mass Isotopologue Definition:

/a(<neutron_count>n[,<atom_number>])

<neutron_count> - number of neutrons of extra mass.

Notice the reversal of neutron_count relative to the "n" representing neutron. This is done so that no ambiguity will exist with the isotopically-resolved isotopologue specification, where the element is indicated first and the isotope_count second.

Examples:

1. M+3 isotopologue of alpha-D-glucopyranose:

InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/a(3n)

The lack of a specified list of atoms indicates the whole molecule.

2. M+4 isotopologue limited to carbon atoms of alpha-D-glucopyranose:

InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/a(4n,1,2,3,4,5,6)

With augmented range specifications:

InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/i(4n,1-6)

This situation could arise from use of tandem mass spectrometry features that localize the extra mass to certain elements.

Partial Isotopomer Specification:

Partial Isotopomer Specification:

Rationale: to enable an InChI representation of a partial isotopomer, which is directly useful for describing most individual nuclear magnetic resonance spectroscopy spectral features.

Simple Definition: same as the full isotopomer specification but with the unknown isotope specifications removed (Example 2). The designation of the isotope for all atoms of a given element will represent a full isotopomer (Example 1).

Examples:

1. ^{13}C at 4th carbon for full isotopomer (with respect to carbon) of alpha-D-glucopyranose:

InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/i1+0,2+0,3+0,4+1,5+0,6+0

With augmented range specification:

InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/i1-3+0,4+1,5-6+0

2. ^{13}C at 4th carbon for partial isotopomer of alpha-D-glucopyranose:

InChI=1/C6H12O6/c7-1-2-3(8)4(9)5(10)6(11)12-2/h2-11H,1H2/t2-,3-,4+,5-,6+/m1/s1/i4+1

Only the atom with the known isotopic composition is indicated.

Cross-Constitutional Isomer Isotopologue Specification:

Cross-Constitutional Isomer Isotopologue Specification:

Rationale: to enable a nonstandard InChI representation of isotopologues that span multiple constitutional isomers, i.e. either a fully or partially isotopically-resolved molecular formula. This specification is directly useful for describing individual features in mass spectrometry when the specific constitutional isomer is unknown.

Simple Definition: 1/Ee#[Ee#].../a(Ee#<+|->#) OR 1/Ee#[Ee#].../a(#n)

Examples:

1. $^{13}\text{C}_2$ isotopologue of molecules with molecular formula $\text{C}_6\text{H}_{12}\text{O}_6$:

InChI=1/C6H12O6//a(C2+1)

2. $^{13}\text{C}_1^{18}\text{O}_1$ isotopologue of molecules with molecular formula $\text{C}_6\text{H}_{12}\text{O}_6$:

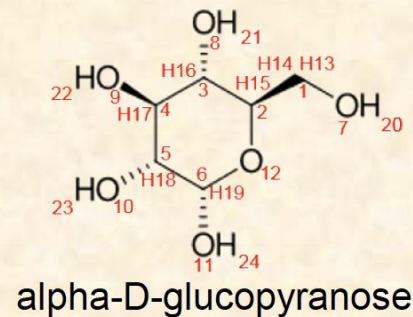
InChI=1/C6H12O6//a(C2+1),(O1+2)

3. M+3 isotopologue of molecules with molecular formula $\text{C}_6\text{H}_{12}\text{O}_6$:

InChI=1/C6H12O6//a(3n)

This last example enables a major use-case in mass spectrometry spectral feature assignment.

The Proposal



- **Full isotopomer with respect to carbon for alpha-D-glucopyranose with ^{13}C at the 4th carbon:**
 - $\text{InChI}=1/\text{C}_6\text{H}_{12}\text{O}_6/\text{c}7-1-2-3(8)4(9)5(10)6(11)12-2/\text{h}2-11\text{H},1\text{H}2/\text{t}2-,3-,4+,5-,6+/\text{m}1/\text{s}1/\text{i}1+0,2+0,3+0,4+1,5+0,6+0$
- **Partial isotopomer of alpha-D-glucopyranose with ^{13}C at the 1st and 2nd carbons:**
 - $\text{InChI}=1\text{S}/\text{C}_6\text{H}_{12}\text{O}_6/\text{c}7-1-2-3(8)4(9)5(10)6(11)12-2/\text{h}2-11\text{H},1\text{H}2/\text{t}2-,3-,4+,5-,6+/\text{m}1/\text{s}1/\text{i}1+1,2+1$
- **$^{13}\text{C}_2^2\text{H}_3$ isotopologue of alpha-D-glucopyranose:**
 - $\text{InChI}=1/\text{C}_6\text{H}_{12}\text{O}_6/\text{c}7-1-2-3(8)4(9)5(10)6(11)12-2/\text{h}2-11\text{H},1\text{H}2/\text{t}2-,3-,4+,5-,6+/\text{m}1/\text{s}1/\text{a}(\text{C}2+1)$
- **$^{13}\text{C}_2$ limited to atoms 4,5,6 isotopologue fragment of alpha-D-glucopyranose:**
 - $\text{InChI}=1/\text{C}_6\text{H}_{12}\text{O}_6/\text{c}7-1-2-3(8)4(9)5(10)6(11)12-2/\text{h}2-11\text{H},1\text{H}2/\text{t}2-,3-,4+,5-,6+/\text{m}1/\text{s}1/\text{a}(\text{C}2+1,4,5,6)$

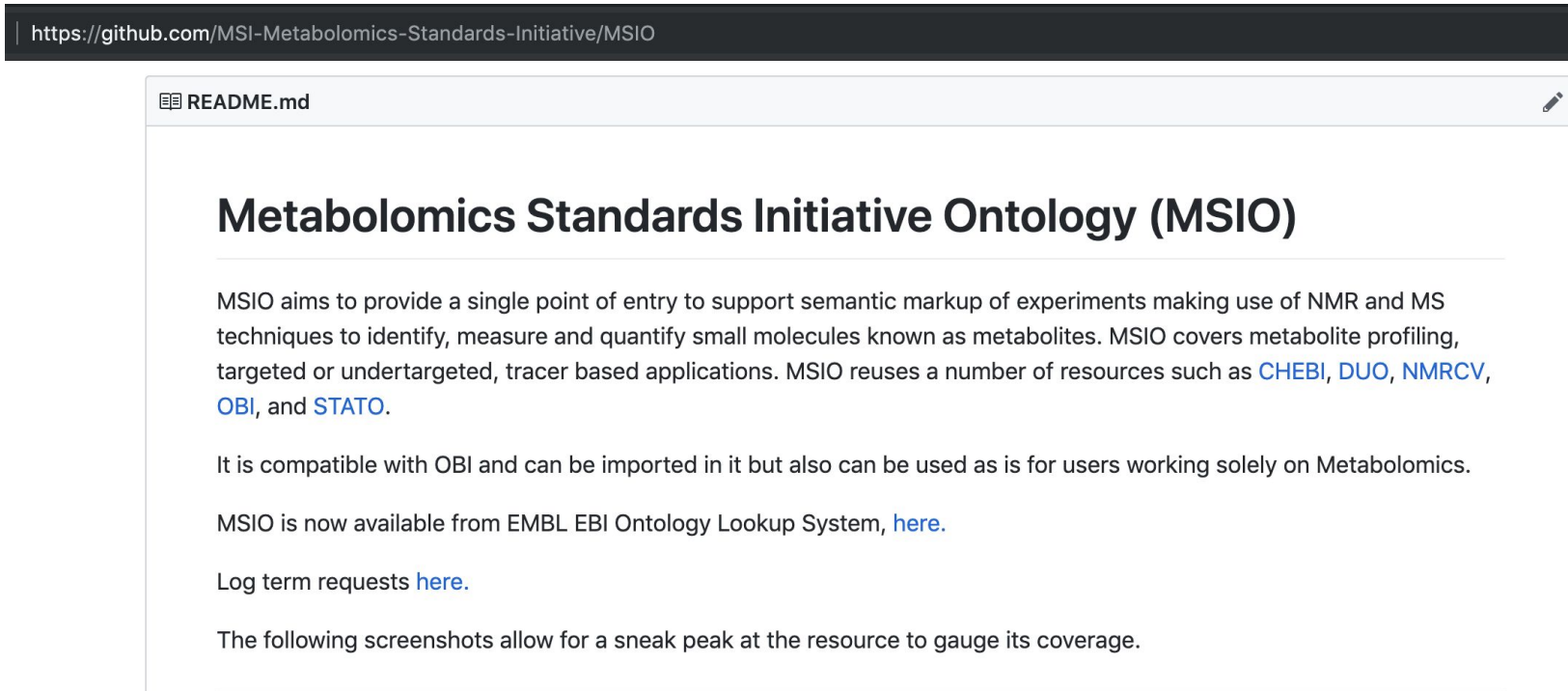
• Credit: Hunter Moseley, University of Kentucky

Issue 2: describing and depositing Stable Isotope Resolved Metabolomic Studies to Metabolights @EMBL-EBI

- ISA the syntax**
- MSIO the terminology**

Quick Overview of MSIO:

- <https://github.com/MSI-Metabolomics-Standards-Initiative/MSIO>



The screenshot shows the GitHub repository page for MSIO. The browser address bar displays the URL: <https://github.com/MSI-Metabolomics-Standards-Initiative/MSIO>. The file name in the editor is README.md. The main heading is "Metabolomics Standards Initiative Ontology (MSIO)". The text describes the project's goals, compatibility with OBI, and availability from EMBL EBI.

Metabolomics Standards Initiative Ontology (MSIO)

MSIO aims to provide a single point of entry to support semantic markup of experiments making use of NMR and MS techniques to identify, measure and quantify small molecules known as metabolites. MSIO covers metabolite profiling, targeted or undertargeted, tracer based applications. MSIO reuses a number of resources such as [CHEBI](#), [DUO](#), [NMRCV](#), [OBI](#), and [STATO](#).

It is compatible with OBI and can be imported in it but also can be used as is for users working solely on Metabolomics.

MSIO is now available from EMBL EBI Ontology Lookup System, [here](#).

Log term requests [here](#).

The following screenshots allow for a sneak peak at the resource to gauge its coverage.

Application ontology:
Built by reusing ontology
modules extracted from
existing resources

Enriched and augmented
by the term additions
collected during the
COSMOS and
PhenoMenal Work

Quick Overview of MSIO:

- OBI
- CHEBI
- STATO
- PSI-MS
- NMR-CV
- DUO

MSIO v1.0.1

Edit

 proccaserra released this on 5 Oct 2018 · [2 commits](#) to master since this release

MSIO v1.0.1:

minor service release from MSIO v1.0 to address ontology metadata for serving by EMBL-EBI Ontology Lookup Service

- *support for Stable Isotope Resolved Metabolomics
- *support for MS and NMR based metabolomics studies
- *support for QA/QC reporting (both description of elements and type of quality controls)
- *support for derivatisation processes and reagents (import of chemical compounds from [CHEBI](#))
- *support for description of data matrices (MzTab-M, MAF or [Metabolomics specific JSON data packages](#))
- *support for data use and consent information reporting (import of terms from [DUO](#))

- *compatible with the Ontology of BioMedical Investigations [OBI](#)
- *compatible and optional import of [STATO](#), the ontology for statistics

Acknowledgments:

*produced with support from [H2020 PhenoMeNal](#) (H2020-EU.1.4.1.3, 654241)

In production, available from EMBL-EBI OLS

EMBL-EBI Services Research

Ontology Lookup Service

Home **Ontologies** Documentation About

OLS > Metabolomics Standards Initiative Ontology (MSIO) MSIO

Metabolomics Standards Initiative Ontology (MSIO)

an application ontology for supporting description and annotation of mass-spectrometry and nmr-spectroscopy based metabolomics experiments and fluxomics studies.

[Terms](#) [Download](#) [Ontology Homepage](#) [Contact](#)

[Browse Terms](#) [Browse Properties](#) [Ontology history](#)

- entity
 - continuant
 - generically dependent continuant
 - independent continuant
 - immaterial entity
 - material entity
 - AbsoluteIDQ Stero17 Kit
 - AbsoluteIDQ p150 Kit
 - AbsoluteIDQ p180 Kit
 - AbsoluteIDQ p400 Kit

Ontology info

Ontology IRI: <http://purl.obolibrary.org/obo/msio.owl>
Ontology id: msio
Version: 1.0.1
Number of terms: 1293
Last loaded: Mon Nov 05 14:54:01 GMT 2018

bug-database
<https://github.com/MSI-Metabolomics-Standards-Initiative/MSIO/issues>

Format
OWL

- What is posted is a reasoned and merged version of the ontology
- It is updated with every release automatically

MSIO has 3 main branches

- Process branch:
 - covering actions such as:
 - Material processing

Branch: master | MSIO / docs / img / msio-material-processing.png

proccaserra more edits to owl reflecting discussions with Hunter Moseley, additio... b966

1 contributor

82.9 KB | Download

- 'material processing'
 - 'establishing cell culture'
 - 'material combination'
 - derivatization
 - acetylation
 - trifluoroacetylation
 - alkylation
 - methylation
 - oximation
 - silylation
 - 'isotopic labeling'
 - 'isotopic labeling with positional labeled tracer'
 - 'isotopic labeling with uniformly labeled tracer'
 - 'material component separation'
 - 'cell layer enzymatic digestion'
 - 'cell scraping'
 - centrifugation
 - extraction
 - 'metabolite extraction'
 - 'lipid extraction'
 - 'polar metabolite extraction'
 - 'gas chromatography'
 - 'high performance liquid chromatography'
 - 'metabolism quenching'
 - 'direct metabolism quenching by solvent'
 - 'metabolism quenching by snap freezing in liquid nitrogen'
 - 'metabolism quenching using precooled 60 percent methanol ammonium bicarbonate buffer'
 - 'metabolism quenching using precooled 60 percent methanol HEPES buffer'
 - 'metabolism quenching using precooled PBS buffer'

MSIO has 3 main branches

- Process branch:
 - covering actions such as:
 - Assay/ Data Acquisition



MSIO has 3 main branches

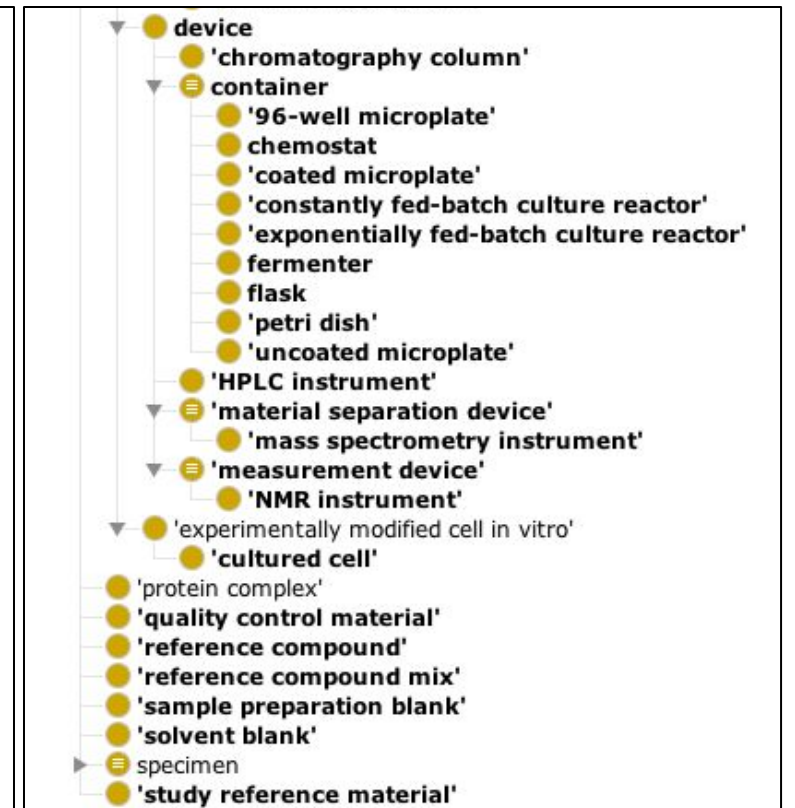
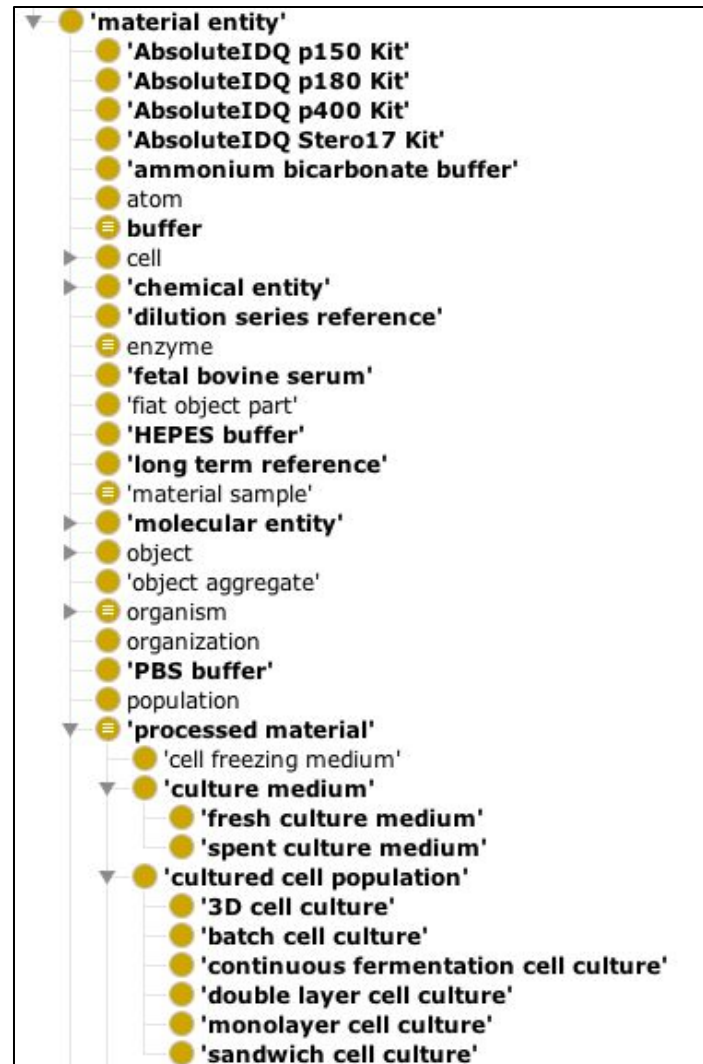
- Process branch:
 - covering actions such as:
 - Data Transformation

Data Transformation:



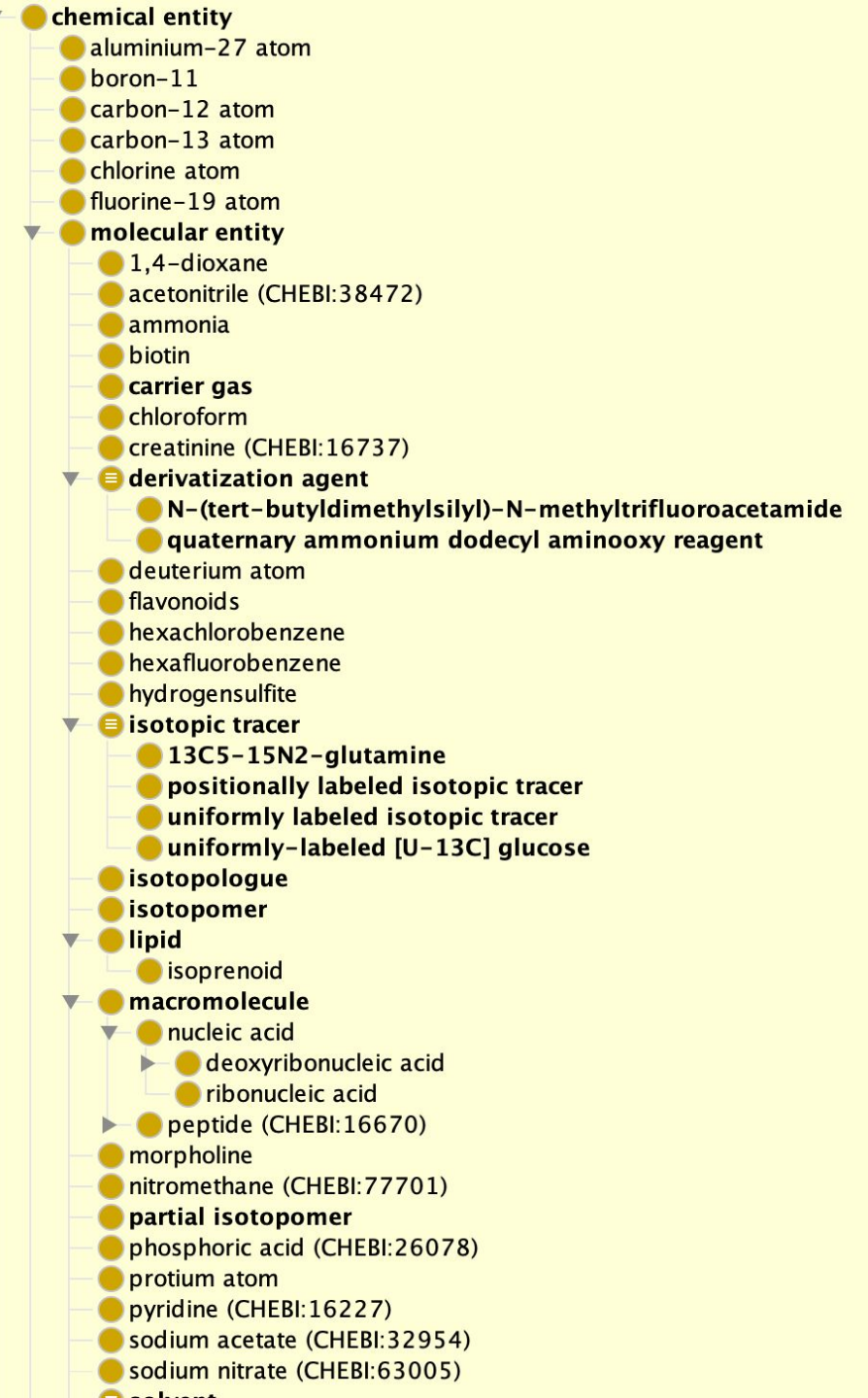
MSIO has 3 main branches

- Process branch:
 - covering actions such as:
 - Material Entity
 - MSIO includes Terms supporting reporting Quality Control Elements



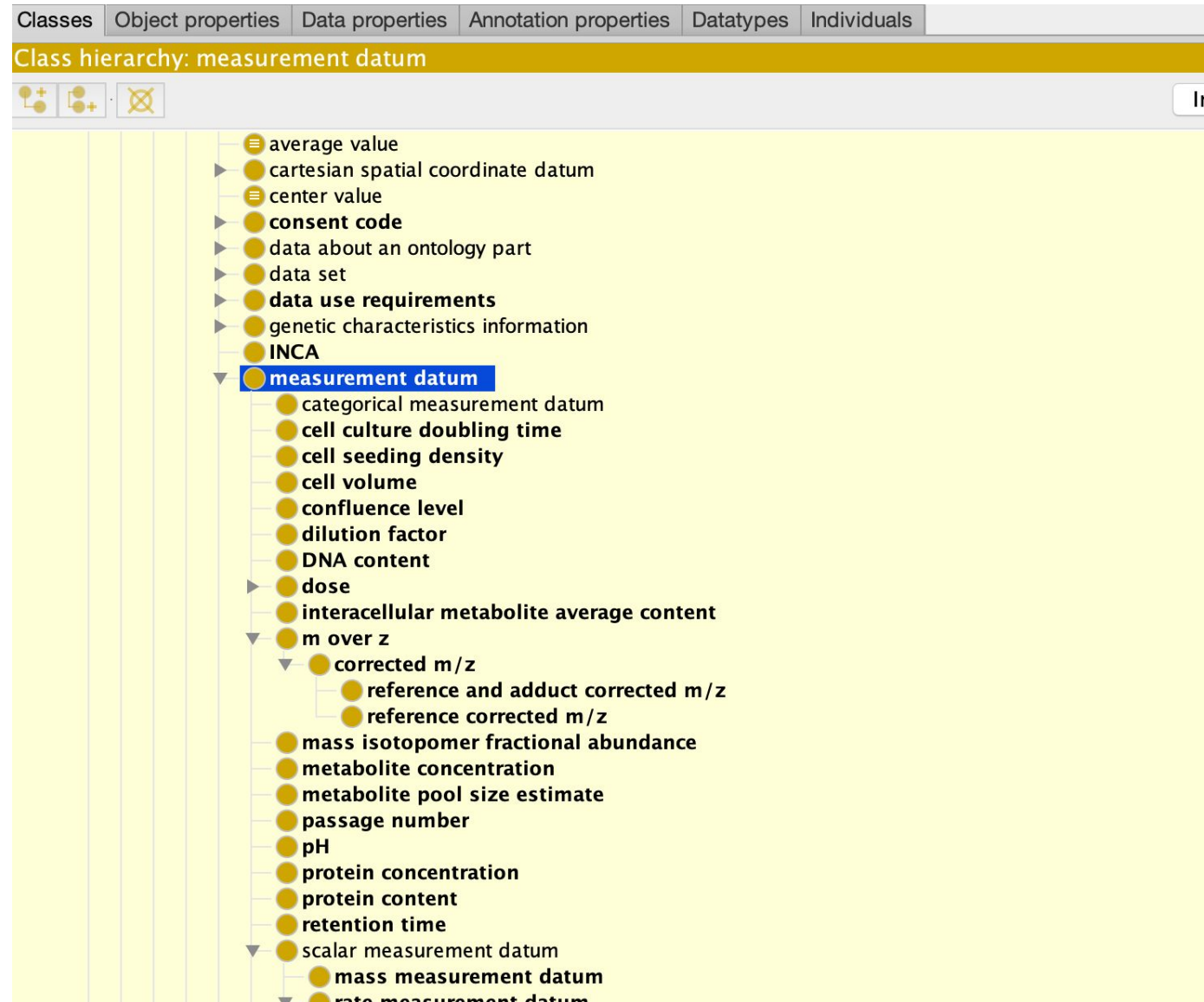
MSIO has 3 main branches

- Process branch:
 - covering actions such as:
 - Molecular Entity
 - Mostly imports from CHEBI



MSIO has 3 main branches

- Information Entity
 - covering actions such as:
 - Measurement types/datum



MSIO support for Fluxomics specific software

The screenshot displays the MSIO web interface for the class 'isodyn' (MSIO:0000049). The interface is divided into two main sections: a class hierarchy on the left and a detailed view of the class on the right.

Class Hierarchy (Left Panel):

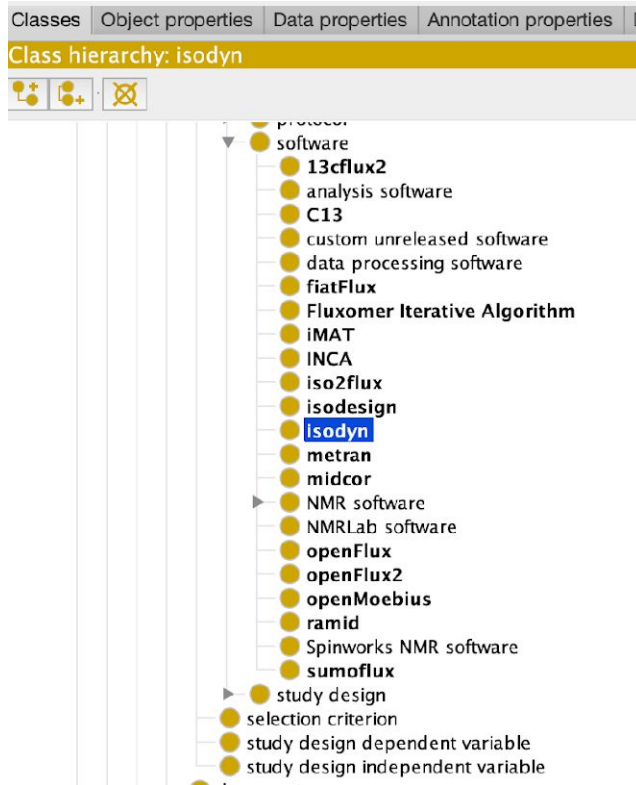
- process
 - software
 - isodyn** (highlighted)
 - 13cflux2
 - analysis software
 - C13
 - custom unreleased software
 - data processing software
 - fiatFlux
 - Fluxomer Iterative Algorithm
 - iMAT
 - INCA
 - iso2flux
 - isodesign
 - metran
 - midcor
 - NMR software
 - NMRlab software
 - openFlux
 - openFlux2
 - openMoebius
 - ramid
 - Spinworks NMR software
 - sumoflux
 - study design
 - selection criterion
 - study design dependent variable
 - study design independent variable

Class Details (Right Panel):

Annotations: isodyn

- label** [language: en]
isodyn
- definition** [language: en]
"Isodyn" is a C++-program that performs an analysis of stable isotope tracer data to assess metabolic flux profiles in living cells. Isodyn simulates the dynamics of isotopic isomer (isotopomer) distribution in central metabolic pathways, and, by changing its parameters, which reflect the characteristics of corresponding biochemical reactions, fit the simulated dynamics of mass isotopomers to that observed experimentally. The simulated metabolic fluxes that correspond to the best fit are assumed to reproduce the real fluxes in the analyzed biological object and conditions. Isodyn contains tools that check the goodness of fit and perform a statistical analysis of obtained metabolic fluxes.
the source code is available from: <https://github.com/seliv55/isodyn>
- Contributor** [language: en]
Pedro de Atauri
- Contributor** [language: en]
Vitaly Selivanov
- 'definition source'** [language: en]
<http://portal.phenomenal-h2020.eu/app-library/isodyn>
- 'has curation status'**
◆ **'ready for release'**
- 'term editor'** [language: en]
Philippe Rocca-Serra

MSIO support for Fluxomics specific software

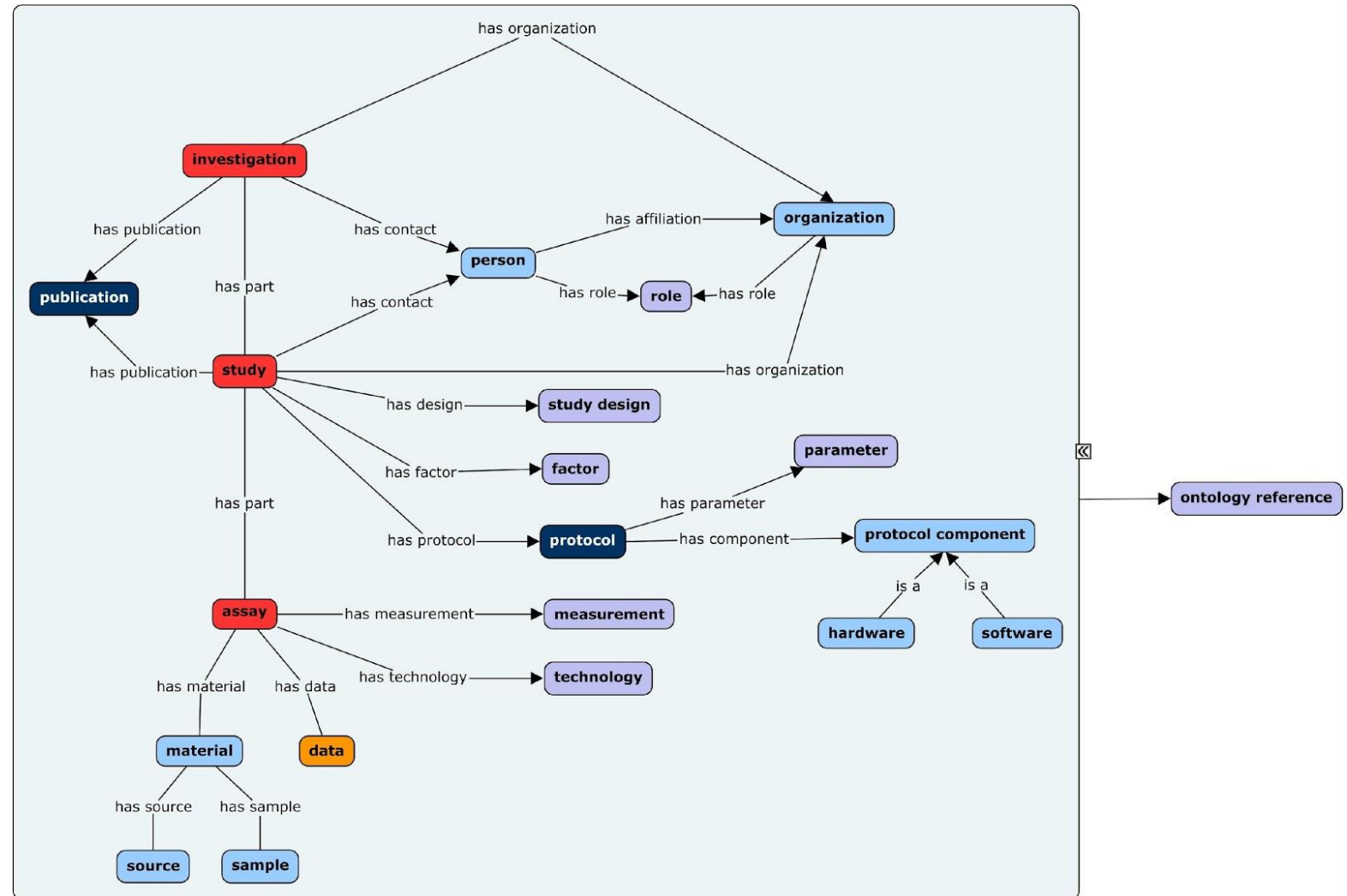


- Need more ?

- **No problem, there a template for that!**
- Take advantage of OBO Foundry Robot tools for submitting, requesting new terms (<http://robot.obolibrary.org/>)
- As simple as filling a spreadsheet
- At minima, provide
 - A name of the software class
 - A definition including a reference to a code repository and/or a DOI / Pubmed ID to a publication (as shown in the example)
- **Log issues / term requests at MSIO github issue tracker**
- <https://github.com/MSI-Metabolomics-Standards-Initiative/MSIO/issues>

ISA Syntax – EMBL-EBI Metabolights Format

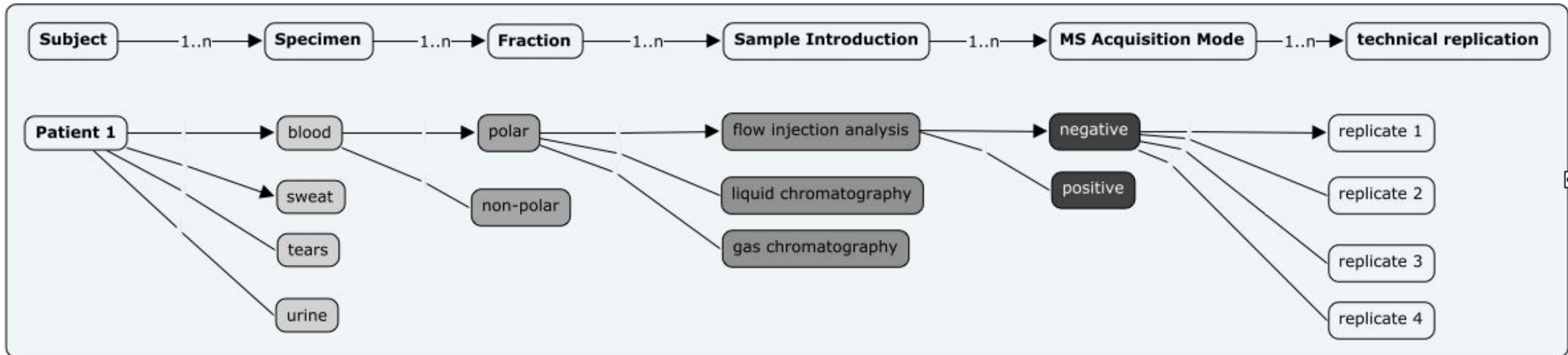
- Information Entity
 - covering actions such as:
 - Measurement types/datum



ISA Syntax – EMBL-EBI Metabolights Format

- Serialization as Tabular text file or JSON document
- JSON-LD context files are available (OBO foundry + Schema.org (not as rich))

Workflows Representations with ISA: Mass Spectrometry Acquisition Case Study



ISA Syntax – EMBL-EBI Metabolights Format

- Fluxomics Specific Annotation Requirements via a set of ISA configurations

ISA-tools / Configuration-Files

Watch 2 Unstar 4 Fork 9

Code Issues 5 Pull requests 0 Projects 0 Wiki Security Insights Settings

Branch: development

Create new file Upload files Find file History

Configuration-Files / isaconfig-SIRM-MTBLs-only-in-vitro_v2015-10-25 /

This branch is 25 commits ahead of master. Pull request Compare

proccaserra fixes to synchronize isaconfig_for_sirm with isa_plugin_config_for_si... Latest commit 53c6a9a on 7 Jun 2018

..

SIRM-isotopologue-distribution_g...	fixes to synchronize isaconfig_for_sirm with isa_plugin_config_for_si...	last year
SIRM-isotopologue-distribution_h...	fixes to synchronize isaconfig_for_sirm with isa_plugin_config_for_si...	last year
SIRM-isotopologue-distribution_lc...	fixes to synchronize isaconfig_for_sirm with isa_plugin_config_for_si...	last year
SIRM-isotopologue-distribution_m...	fixes to synchronize isaconfig_for_sirm with isa_plugin_config_for_si...	last year
SIRM-isotopomer-distribution_nm...	fixes to SIRM configurations, for both the ISAcreator and the plugin	last year
investigation.xml	fixes to SIRM configurations, for both the ISAcreator and the plugin	last year
studySample.xml	fixing to ISA SIRM configurations and ISA data matrix plugin configur...	last year

ISA Syntax – EMBL-EBI Metabolights Format

- Several studies from Marta Cascante's team have been submitted
- <https://www.ebi.ac.uk/metabolights/MTBLS247>
- <https://www.ebi.ac.uk/metabolights/MTBLS412>

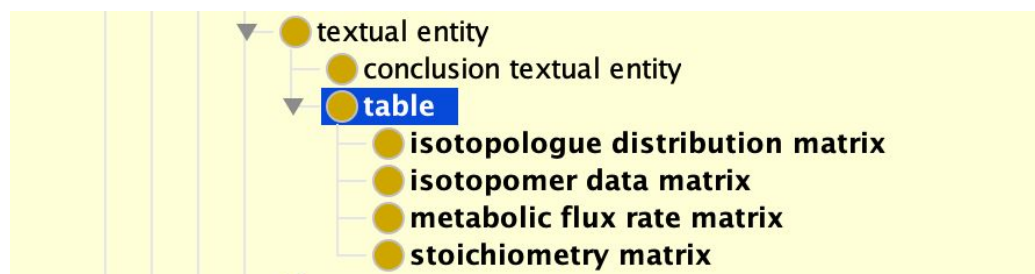
- ISA converter can retrieve Metabolomics Workbench Data and convert to ISA syntax (pb = absence of raw data files from the US repository)
- ISA python code to convert US DCC JSON to ISA format
 - Caveat: free text only, not semantic markup, could be done post-hoc by NER.

Issue 3: Data Matrix Results

- **Again, both syntax and semantics to be agreed upon**

- Dedicated Data Matrix Formats:
 - Option 1: derived from MAF file for compatibility with ISAcreator plugin
 - Validation
 - Layout: wide table + fixed height (i.e. fixed number of molecular dimensions)
 - Option 2: Exploratory JSON Tabular Data packages
 - Layout: long table + fixed width (i.e. fixed number of measurement dimensions)

Modeling Data Matrices



Description: isotopologue distribution matrix

Equivalent To +

SubClass Of +

- 'has part' some ('chemical compound formula' and 'signal intensity' and 'corrected m/z' and 'retention time' and 'chemical entity assignment')

● table

Classes Object properties Data properties Annotation properties Datatypes Individuals

Class hierarchy: measurement datum



- average value
- cartesian spatial coordinate datum
- center value
- consent code
- data about an ontology part
- data set
- data use requirements
- genetic characteristics information
- INCA
- measurement datum
 - categorical measurement datum
 - cell culture doubling time
 - cell seeding density
 - cell volume
 - confluence level
 - dilution factor
 - DNA content
 - dose
 - interacellular metabolite average content
 - m over z
 - corrected m/z
 - reference and adduct corrected m/z
 - reference corrected m/z
 - mass isotopomer fractional abundance
 - metabolite concentration
 - metabolite pool size estimate
 - passage number
 - pH
 - protein concentration
 - protein content
 - retention time

Provisioning for FAIR data matrices – MS

implementation as a ISA Table for ISACreator Plugin:

https://github.com/ISA-tools/Configuration-Files/blob/development/datamatrix-plugin-configuration/configuration_SIRM_ms.xml

```
<!-- OPTIONAL REPORTING OF TRACER MOLECULE INFORMATION -->
<field header="ms_acquisition_ref" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [5 lines]
<field header="isotopic_tracer" data-type="String" is-file-field="false" [4 lines]
<field header="retention_time" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<unit-field data-type="List" is-multiple-value="false" is-required="false" [5 lines]
<field header="cluster_id" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="mass_to_charge" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [5 lines]
<field header="post_collision_mass_to_charge" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="selected_ion_mass_to_charge" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="raw_signal_intensity" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [10 lines]
<unit-field data-type="List" is-multiple-value="false" is-required="false" [5 lines]
<field header="signal_intensity_type" data-type="List" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [6 lines]
<field header="corr_signal_intensity" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [10 lines]
<unit-field data-type="List" is-multiple-value="false" is-required="false" [5 lines]
<field header="signal_intensity_correction_type" data-type="List" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [6 lines]
<field header="norm_signal_intensity" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [10 lines]
<unit-field data-type="List" is-multiple-value="false" is-required="false" [5 lines]
<field header="normalization_type" data-type="List" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [6 lines]
<field header="chemical_formula" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [6 lines]
<field header="molecular_fragment" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [4 lines]
<field header="isotopologue" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [5 lines]
<field header="isotope" data-type="List" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [6 lines]
<field header="corrected_mass_to_charge" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="m_to_z_correction_type" data-type="List" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [6 lines]
<field header="molecular_fragment" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [5 lines]
<field header="chemical_formula" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [5 lines]
<field header="adduct_formula" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="smiles" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="inchi" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="assignment_type" data-type="List" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [6 lines]
<field header="assignment_confidence" data-type="List" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [6 lines]
<field header="database" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="database_identifier" data-type="String" is-file-field="false" is-multiple-value="true" is-required="false" is-hidden="false"> [5 lines]
<field header="isotologue_abundance_relative_concentration" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [6 lines]
<unit-field data-type="List" is-multiple-value="false" is-required="false" [4 lines]
```

Provisioning for FAIR data matrices - NMR

implementation as a ISA Table for ISACreator Plugin:

https://github.com/ISA-tools/Configuration-Files/blob/development/datamatrix-plugin-configuration/configuration_SIRM_nmr.xml

```
<field header="nmr_acquisition_ref" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [5 lines]
<field header="chemical_shift_ppm" data-type="Float" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [4 lines]
<field header="unit" data-type="Float" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [4 lines]
<!-- [10 lines]
<field header="multiplicity" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false"> [3 lines]
<field header="connectivity" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false"> [3 lines]
<field header="satellite_resonances" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false"> [3 lines]
<field header="saturation_factor" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="nb_proton_correction_peak_area" data-type="Integer" is-file-field="false" is-multiple-value="false" is-required="false"> [4 lines]
<field header="deconvoluted_peak_area" data-type="Float" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="peak_deconvolution_method" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="mean_post_rltv_conc_estimate" data-type="Float" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="95pc_upr_bound" data-type="Float" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="95pc_lwr_bound" data-type="Float" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="chemical_formula" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true"> [3 lines]
<field header="smiles" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="inchi" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="database" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [5 lines]
<field header="database_identifier" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [5 lines]
<field header="isotopomer" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [5 lines]
<field header="labeled_position" data-type="String" is-file-field="false" is-multiple-value="false" is-required="true" is-hidden="false"> [5 lines]
<field header="assignment_type" data-type="String" is-file-field="false" is-multiple-value="false" is-required="false"> [4 lines]
<field header="assignment_confidence" data-type="List" is-file-field="false" is-multiple-value="false" is-required="false" is-hidden="false"> [6 lines]
```

Provisioning for FAIR data matrices – Analysis Tools

- Output of Ramid, Midcor and iso2flux implement the current recommendations
 - (credits to Vitaly Selivanov, Pedro de Atauri, Marta Cascante, Silvia Marin, University of Barcelona)

FLUXOMICS TOOLS

Fluxomics

ramid Evaluates the peaks of mass isotopomer distribution (MID), making them ready for correction for natural isotope occurrence. [UBarcelona](#)

midcor corrects ^{13}C mass isotopomers spectra of metabolites for natural occurring isotopes and peaks overlapping. [UBarcelona](#)

iso2flux ^{13}C Metabolic Flux Analysis on a sub-network of a large scale model. [UBarcelona](#)

isodyn C++ program simulating the dynamics of metabolites and their isotopic isomers in central metabolic network using kinetic model [UBarcelona](#)


- Option 2: Exploratory JSON Tabular Data package

ISA-tools / isa-matrix-datapackages



Unwatch 5 Star 0 Fork 1

Code Issues 0 Pull requests 0 Projects 0 Wiki Security Insights Settings

Branch: master isa-matrix-datapackages / src / sirm_datapkg / Create new file Upload files Find file History

 proccaserra alignment work for label, definition and requirement status, now ok b... Latest commit 63deedc on 8 May 2018

..

 ms_maf_datapkg	alignment work for label, definition and requirement status, now ok b...	last year
 nmr_maf_datapkg	alignment work for label, definition and requirement status, now ok b...	last year



Option 2: Exploratory JSON Tabular Data package

ISA-tools / isa-matrix-datapackages Unwatch 5 Star 0 Fork 1

Code Issues 0 Pull requests 0 Projects 0 Wiki Security Insights Settings

Branch: master [isa-matrix-datapackages / src / sirm_datapkg / ms_maf_datapkg / MS_SIRM.csv](#) Find file Copy path

proccaserra minor edits to ms component of sirm jdp + revision of nmr component a... 48e5813 on 11 Oct 2017

[1 contributor](#)

5 lines (5 sloc) | 1.56 KB Raw Blame History

Search this file...

	ms_acquisition_ref	isotopic_tracer	tracer_abundance_pc	retention_time	retention_time_unit	mass_to_charge	cluster_id	raw_signal_intensity	raw_s
1									
2	data_acquisition_1	D-(13C6)Glucose	50	2.57	min	0.002669267		284	arbitr
3	data_acquisition_1	D-(13C6)Glucose	50	2.57	min	0.002669267		285	arbitr
4	data_acquisition_1	D-(13C6)Glucose	50	20.35	min	0.002669267		294	arbitr
5	data_acquisition_1	D-(13C6)Glucose	50	36.19	min	0.002669267		300	arbitr

- Option 2: Exploratory JSON Tabular Data package

```
"publishers": [{
  "name": "Philippe Rocca-Serra",
  "email": "proccaserra@gmail.com",
  "web": "http://www.stato-ontology.org"
}],
"resources": [{
  "name": "MS_SIRM",
  "title": "MS_SIRM",
  "path": "MS_SIRM.csv",
  "schema": {"fields": [
    {
      "name": "ms_acquisition_ref",
      "title": "MS Assay Name",
      "description": "a reference to an ISA mass spectrometry acquisition event",
      "format": " default",
      "type": "string",
      "rdfType": "http://purl.obolibrary.org/obo/OBI_0600013",
      "constraints": {"required": "True"}
    },
    {
      "name": "isotopic_tracer",
      "title": "isotopic tracer",
      "description": "the name of the isotopic tracer molecule used in the labeling experiment. The chemical name of",
      "format": " default",
      "type": "string",
      "rdfType": "http://purl.obolibrary.org/obo/MSIO_0000031",
      "constraints": {"required": "False"}
    }
  ]
}
```

Future Work

- More testing
 - BYOD and creation ISA formatted documents for deposition to Metolights
 - Hackathon with tool developers to implement regularized outputs and self standing data matrices
 - Vocabulary development
 - submit terms where a need arises
 - make some resources used by MSIO more robust with resolvable uri (e.g. NMR cv)

Acknowledgements

- FP7 COSMOS: grant agreement number: EC312941
- H2020 PhenoMeNal: grant agreement number: 654241

All thanks go to:

Marta Cascante, Vitaly Selivanov, Pedro de Atauri, Kenneth Haug, Reza Salek, Claire o'Donovan, Chris Steinbeck, Hunter Moseley, Teresa Fan, Andrew Lane

Questions