Philippe Rocca-Serra<sup>1</sup> (\*), David Johnson(\*)<sup>1</sup>, Ralf J. M. Weber<sup>2</sup>, Luca Pireddu<sup>3</sup>, Pierrick Roger<sup>4</sup>, Alejandra Gonzalez-Beltran<sup>1</sup>, Massimiliano Izzo<sup>1</sup>

Enrico Piras<sup>3</sup>, Pablo Moreno<sup>5</sup> and Susanna-Assunta Sansone<sup>1</sup>.

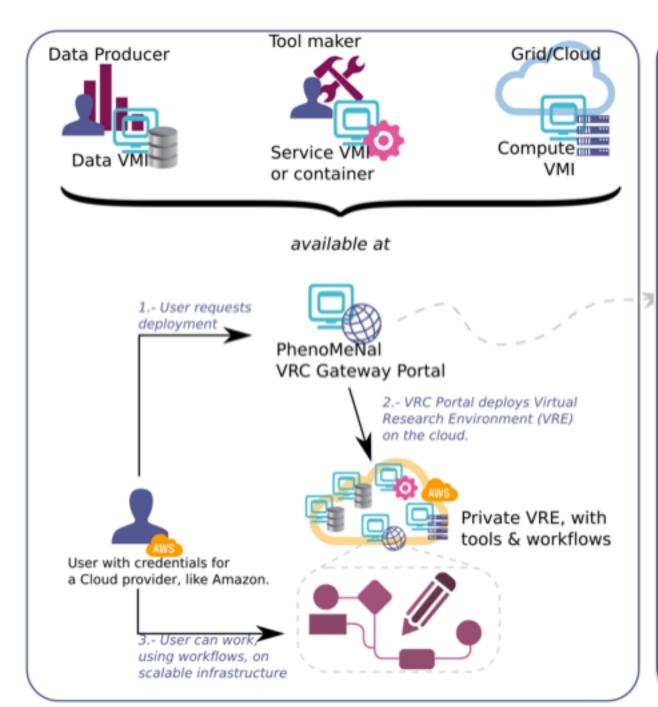
- 1. University of Oxford e-Research Centre, Department of Engineering Science, Oxford, OX1 3QG, UK.
- 2. Phenome Centre Birmingham, University of Birmingham, Birmingham, B15 2TT, UK.
- 3. CRS4, Science and Technology Park of Sardinia, 09010 Pula CA, Italy
- 4. Freelance software engineer
- 5. EMBL-EBI, Wellcome Genome Campus, Hinxton Cambridge, CB10 1SD, UK

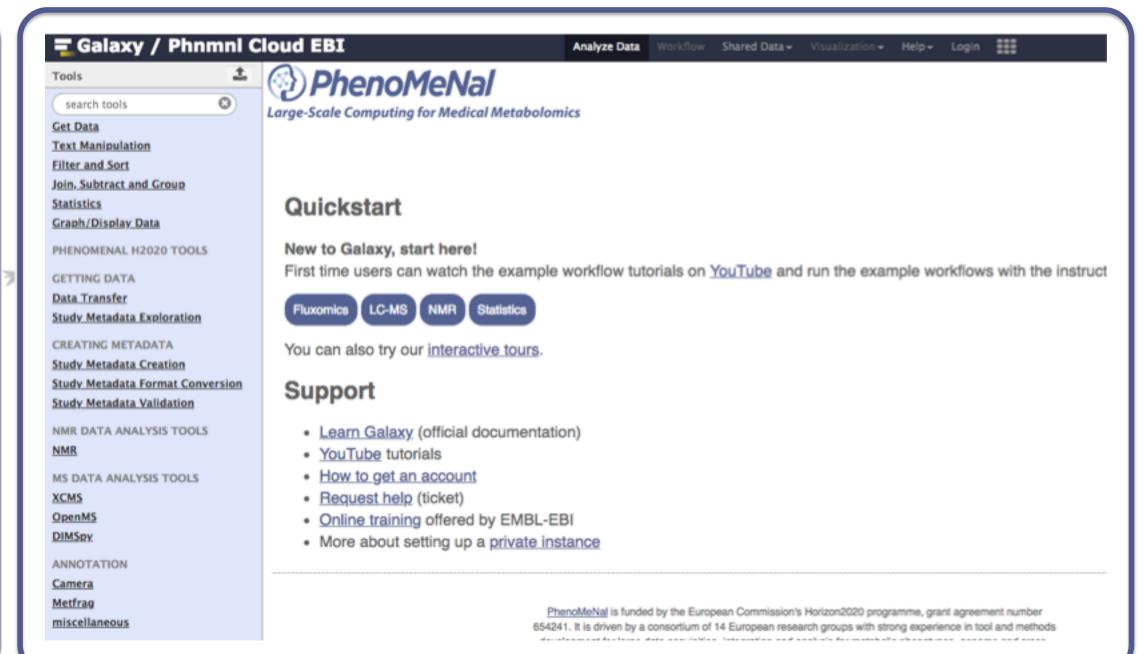


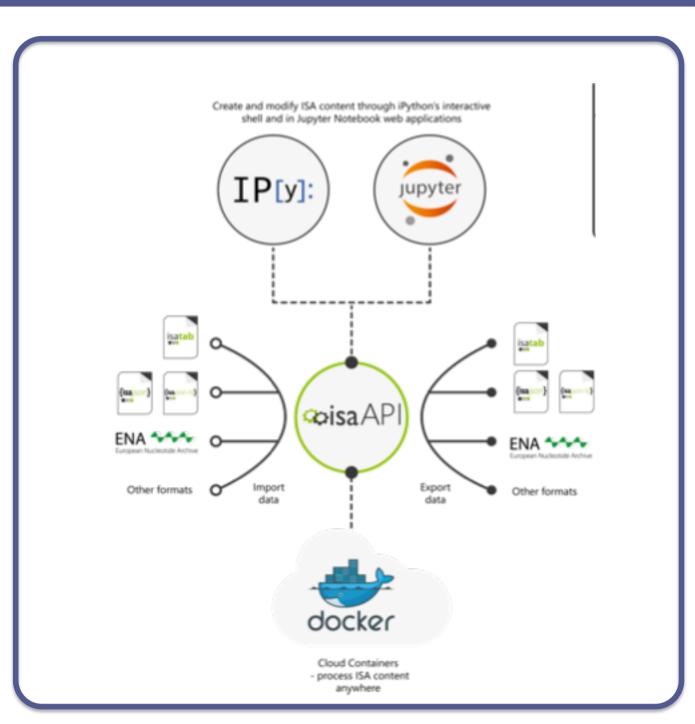


# ISAcreate Galaxy Tool for Prospective Data Management with ISA format support - Application to Metabolomics Datasets

The ISAcreate Galaxy tool is a new component developed as part of the European Union (https://phenomenal-h2020.eu) to help the reporting of experimental data. Built around the ISA syntax and the soon to be available as a native Galaxy datatype, ISAcreate enables the creation of ISA-compliant archives, ready for deposition to prime repositories such as EMBL-EBI MetaboLights.







Using 497.6 K

< Back to Unnamed history

a list of datasets

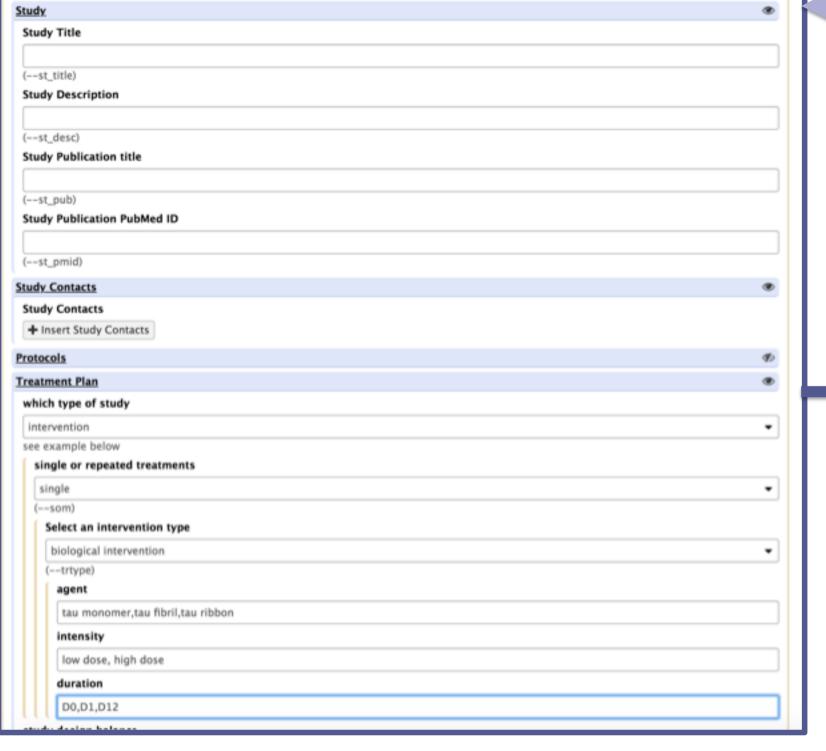
a\_plasma\_non-p

gative\_assay

#### A Galaxy tool definition powered by the ISA-API

reate ISA in Galaxy interactive tool to create ISA archives (Galaxy Version 1.0.0)

https://github.com/ISA-tools/isatools-galaxy



GITHUB.COM/ISA-TOOLS

The Galaxy tool is built around a capability recently added to the ISA API. The ISA Create Mode, by embedding key knowledge about design of experiments and placing study design descriptors at the top of the reporting requirements, allows the rapid creation of regularized data management templates.

### ISA-Tab format as Galaxy data type

mzml2isa Parser to get meta ISA information from mzML files and create an ISA-Tab structure information from nmrML files and udy Metadata Format Conversion Study Metadata Validation NMR DATA ANALYSIS TOOLS implement

annotation requirements

search tools

**Text Manipulation** 

Graph/Display Data

GETTING DATA

Data Transfer

Join, Subtract and Group

PHENOMENAL H2020 TOOLS

Study Metadata Exploration

itudy Metadata Creation

Create ISA in Galaxy

Interactive tool to create ISA

archives based on study design

CREATING METADATA

Filter and Sort

Get Data

Statistics

conventions, provided a set of core rules underwriting the templating engine are followed. We demonstrate the advantages of the approach by developing a component aimed at reporting experiments using metabolomics techniques, and that supports both NMR and Mass spectrometry approaches, and their various data acquisition modalities. It is being tested at some of the UK National Phenome Centres.

Parameter Value[collection event rank] Sample Name

1 Group-001.Subject-002.plasma.1 plasma

1 Group-001.Subject-003.plasma.1 plasma

1 Group-002.Subject-001.plasma.1 plasma

1 Group-002.Subject-002.plasma.1 plasma

1 Group-002.Subject-003.plasma.1 plasma

L Group-003.Subject-002.plasma.1 plasma

I Group-004.Subject-001.plasma.1 plasma

1 Group-009.Subject-001.plasma.1 plasma 1 Group-009.Subject-002.plasma.1 plasma

1 Group-009.Subject-003.plasma.1 plasma

1 Group-010.Subject-001.plasma.1 plasma

1 Group-010.Subject-002.plasma.1 plasma

naming

Group-003.Subject-003.plasma.1

81 lines, 1 comments I Group-004.Subject-002.plasma.1 plasma a-syn monomer format: tsv, database: ? 1 Group-004.Subject-003.plasma.1 plasma a-syn monomer B 0 C a-syn ribbon a-syn ribbon 1 Group-005.Subject-002.plasma.1 plasma 1 Group-005.Subject-003.plasma.1 plasma a-syn ribbon 1 Group-006.Subject-001.plasma.1 plasma Group-001.Subject-001 specimen a-syn ribbon iroup-001.Subject-002 specimen 1 Group-006.Subject-003.plasma.1 a-syn ribbon roup-001.Subject-003 specimen 1 Group-007.Subject-001.plasma.1 plasma a-syn monomer Group-002.Subject-001 specime 1 Group-007.Subject-002.plasma.1 plasma 1 Group-008.Subject-001.plasma.1 plasma a-syn monome a-syn monome

0.01 nM

Semantic markup of ISA templates is achieved by populating many of the select options from an external file that stores all controlled vocabulary and terminology values and feeding those values to the templating engine. Study design information is semantically marked up with ontology terms coming from STATO (https://stato-ontology.org), while other terms are drawn from the Metabolomics Standards Initiative Ontology (https://github.com/ISA-tools/MSIO).

- DUO for Consent and Data reuse
- UBERON for anatomical part

A rich Python API now supports the ISA metadata formats, allowing processing of ISA

documents, either as ISA-Tab or ISA-JSON, such as syntactic + content/light semantic validations

a-syn monome

a-syn monome

a-syn ribbon

a-syn ribbon

a-syn ribbor

a-syn fibril

a-syn fibril

a-syn fibril

a-syn ribbon

a-syn ribbon

a-syn ribbon

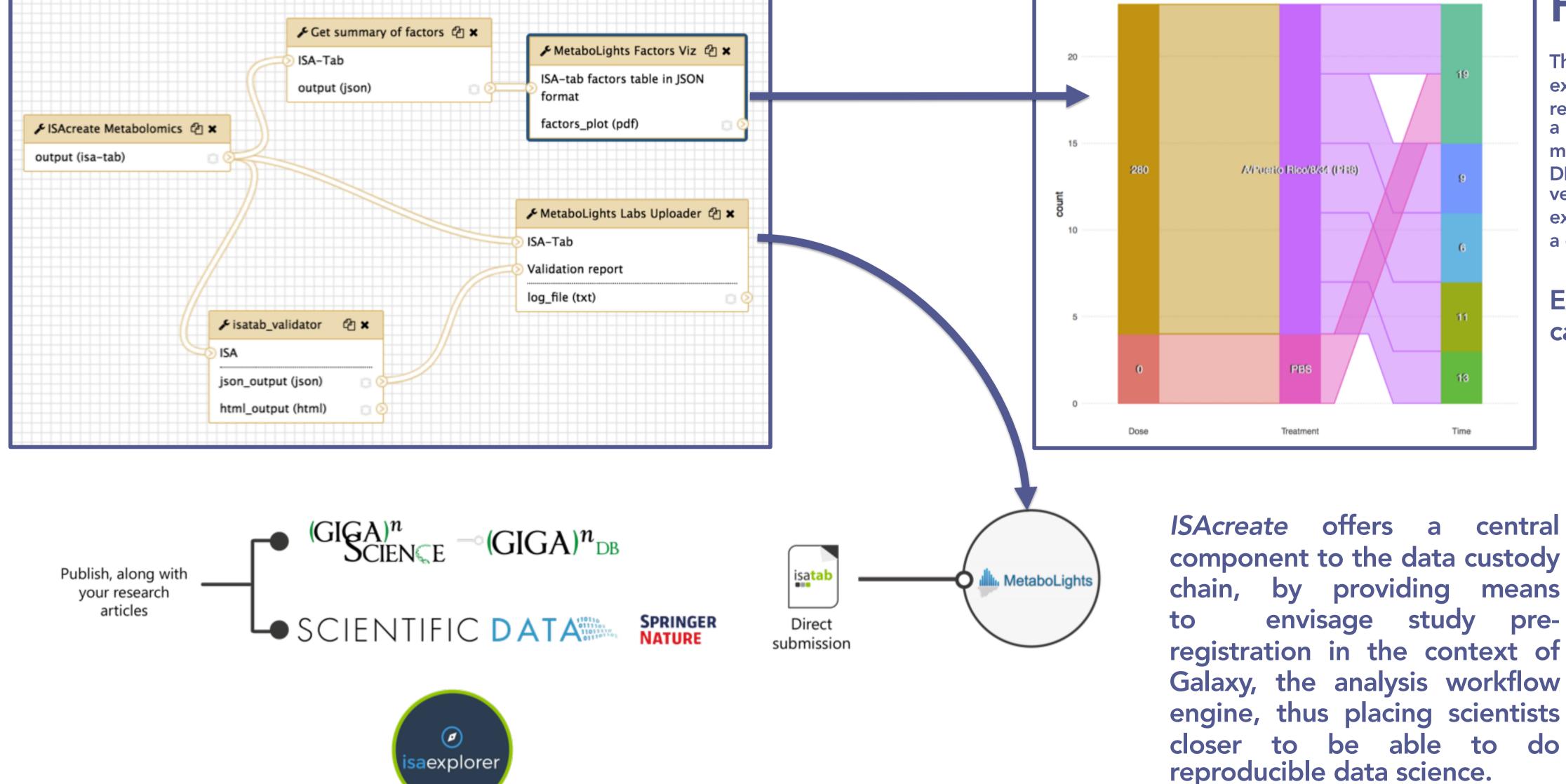
a-syn fibril

a-syn fibri

and conversions (e.g. to well-established formats such SRA or MAGE-TAB when suitable).

- OBI and MSIO for wet/dry lab procedure
- STATO for statistical methods
- Integration of EMBL-EBI MetaboLights curated lists of instrument and parameters

## A Galaxy workflow for study deposition/pre-registration to MetaboLights



#### Future developments:

The flexibility and genericity of the approach is highlighted by extending the initial metabolomics use case to support the reporting of genomics data according to similar principles. Hence, a complementary tool offers similar capabilities for prospective metadata templating for functional genomics applications using DNA microarray or NGS techniques. This further validates the versatility of the ISA specifications, supporting multi-omics experiments, which allow molecular phenotyping to be tracked in a consistent fashion

Expand the tools and workflows ISA aware and capable of producing FAIR datasets:







PhenoMeNal







**London** 







UNIVERSITY<sup>OF</sup> BIRMINGHAM







Funded by European

Union's Horizon2020