



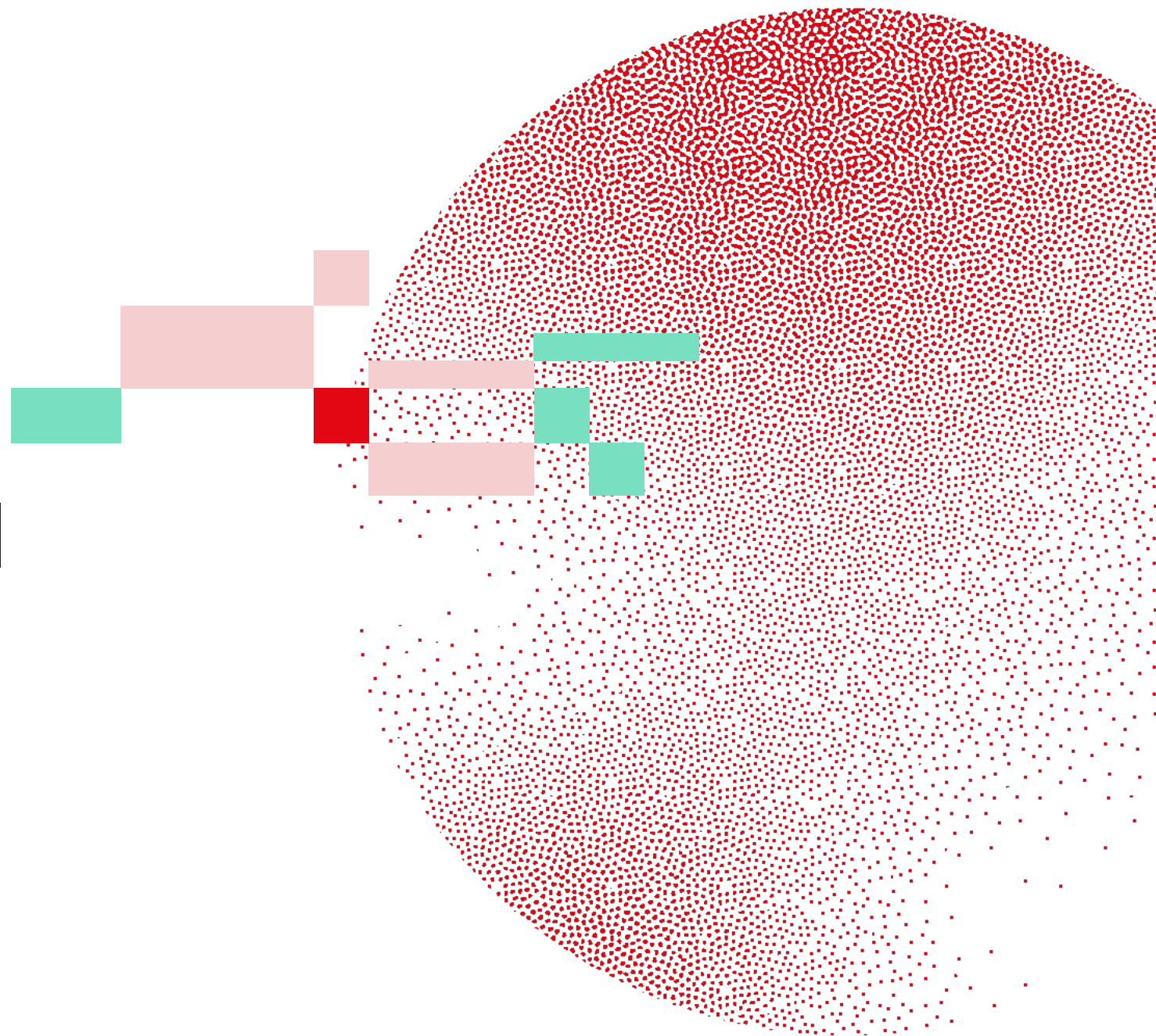
Swiss Institute of
Bioinformatics

EMI semantic model

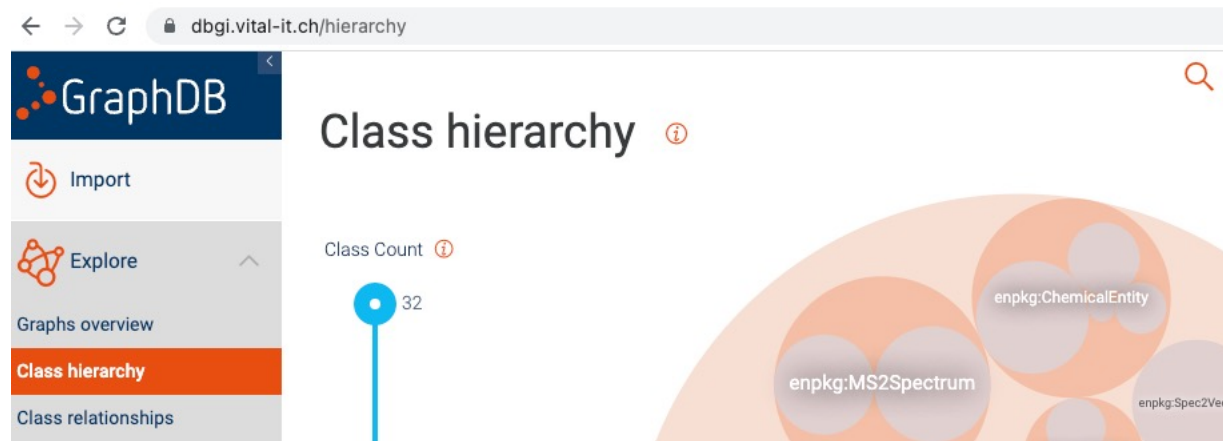
Tarcisio Mendes

08/11/2023

kru@sib.swiss



Testbed and BioQuery



❖ <https://dbgi.vital-it.ch/sparql>

Bio-Query^B: Federated template search over biological databases (DBGI Edition)

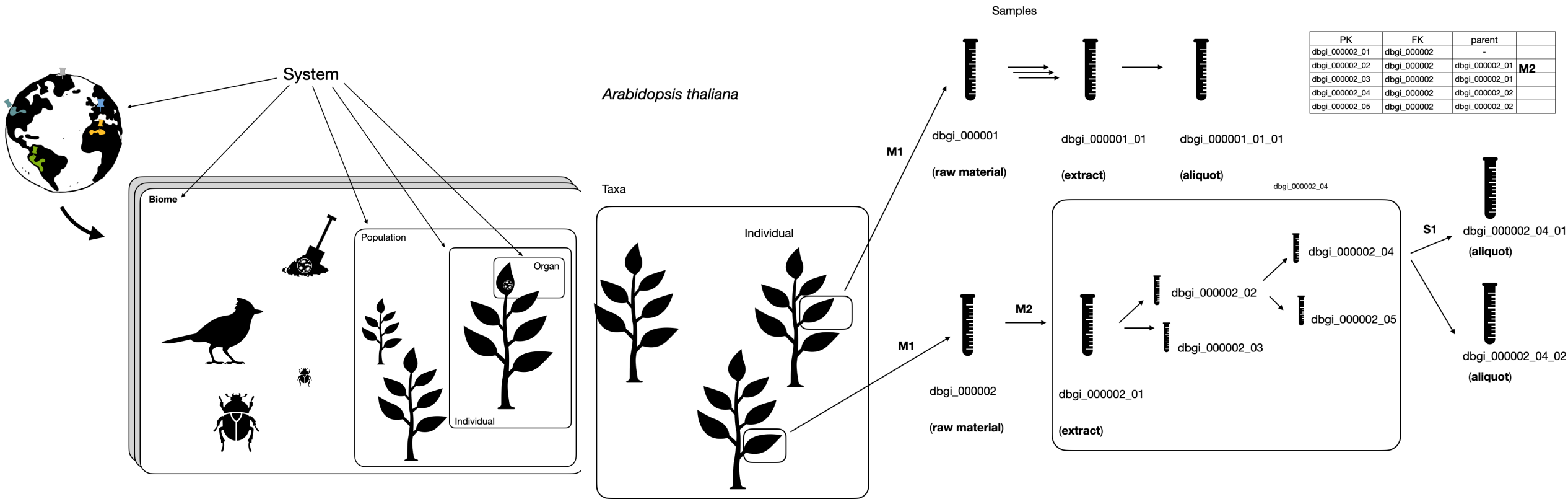
Status of the queried service poi
DBGI ENPKG federation server UniProt Wikidata
Search our queries... Expand All Show SPARQL Query Editor Limited results are on
Reset / Reload About
> Contact form
▼ Sample DBGI Queries
Count all triples in the current DBGI graph
How many features (PI and NI modes) have the same SIRIUS/CSI:FingerID and ISDB annotation?
Which samples have features (PI mode) annotated as Aspidosperma_type alkaloids by CANOPUS with a probability score above 0.5, ordered by the decreasing count of features?

❖ **BioQuery** – Template queries over the DBGI graph

❖ <https://dbgi.vital-it.ch/bioquery>

❖ To provide more real questions and templates

System of systems – from ecosystems to a single individual portion sample



System of systems – from ecosystems to a single individual portion/sample

- ❖ To describe system of systems we rely on **SOSA (Sensor, Observation, Sample, and Actuator)** ontology
 - ❖ Supports a wide range of applications and use cases including scientific observations
 - ❖ Although this ontology was not designed for our problem, it fits well our semantic modelling needs
 - ❖ It is generic enough to accommodate other use cases and extensions
 - ❖ A lightweight version of SSN (Semantic Sensor Network) ontology



System of systems – from ecosystems to a single individual portion sample

- ❖ Modelling examples
- ❖ https://github.com/digital-botanical-gardens-initiative/earth_metabolome_ontology

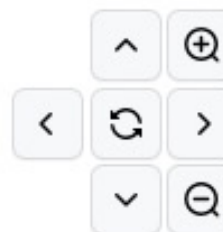
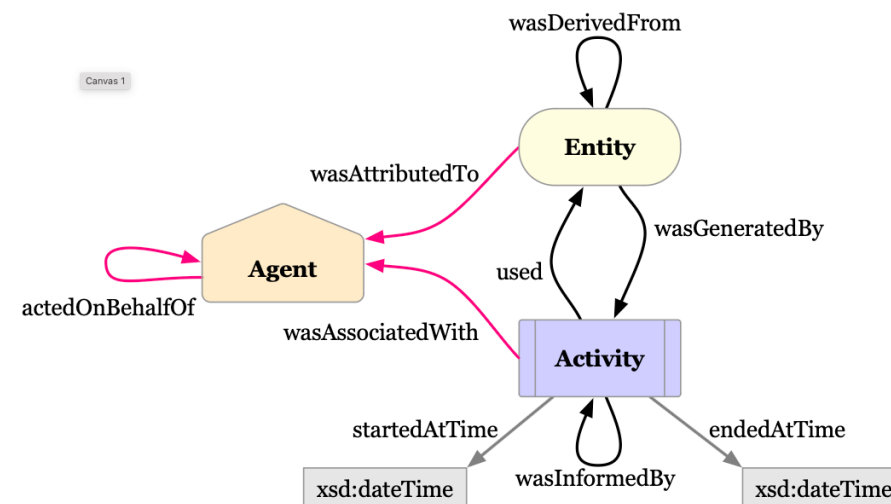
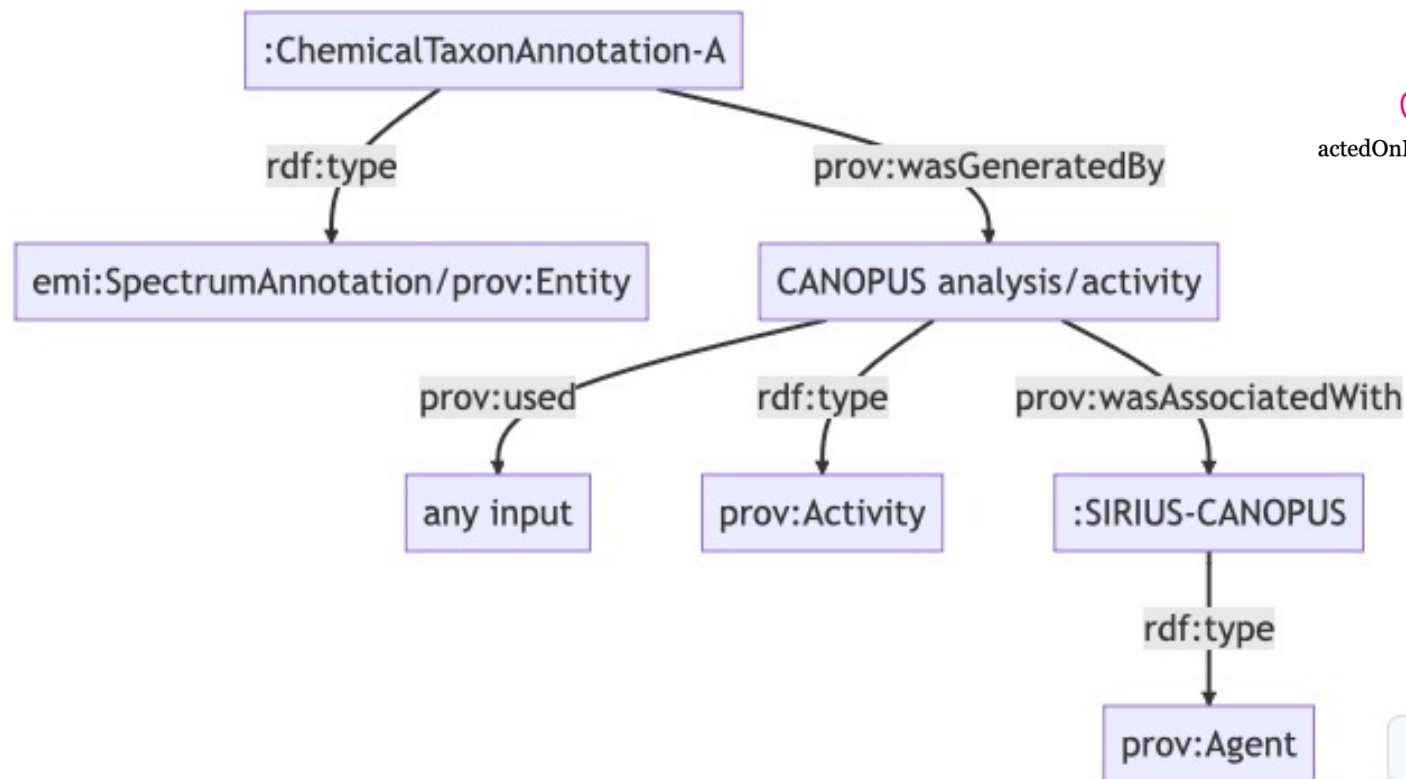
W3C Recommendation **Semantic Sensor Network Ontology** **OGC**[®] **W3C**
Making location count.

W3C Recommendation 19 October 2017 (Link errors corrected 08 December 2017)

This version:
<https://www.w3.org/TR/2017/REC-vocab-ssn-20171019/>

Spectrum annotation provenance [↗](#)

To describe annotation provenance (i.e., information source), we are importing and applying the [PROV-Ontology](#) (a W3C recommendation). Below we show an application example to our knowledge domain:



Curating EMI terms collaboratively with WebProtege:

The screenshot displays the WebProtege interface for the 'emi-semantic-model' project. The top navigation bar includes 'Home' and 'Display' options. Below this, there are tabs for 'Classes', 'Properties', 'Individuals', 'Comments', 'Changes by Entity', and 'History'. The 'Classes' tab is active, showing a 'Class Hierarchy' on the left and a 'Class: GNPS consensus MS2 spectrum' details panel on the right.

Class Hierarchy:

- owl:Thing
 - Agent
 - Any cross-reference
 - Bioassay result
 - Chemical entity
 - Concept
 - Entity
 - FBMN component
 - Information source
 - LCMS feature set
 - Laboratory object
 - MS2 spectrum
 - GNPS consensus MS2 spectrum**
 - LCMS individual MS2 spectrum
 - MS2Spectra pair
 - Observation
 - Organism taxon
 - Sample
 - Spec2vec document
 - Spectrum annotation
 - System

Class: GNPS consensus MS2 spectrum

IRI
<https://purl.org/emi#GNPSConsensusSpectrum>

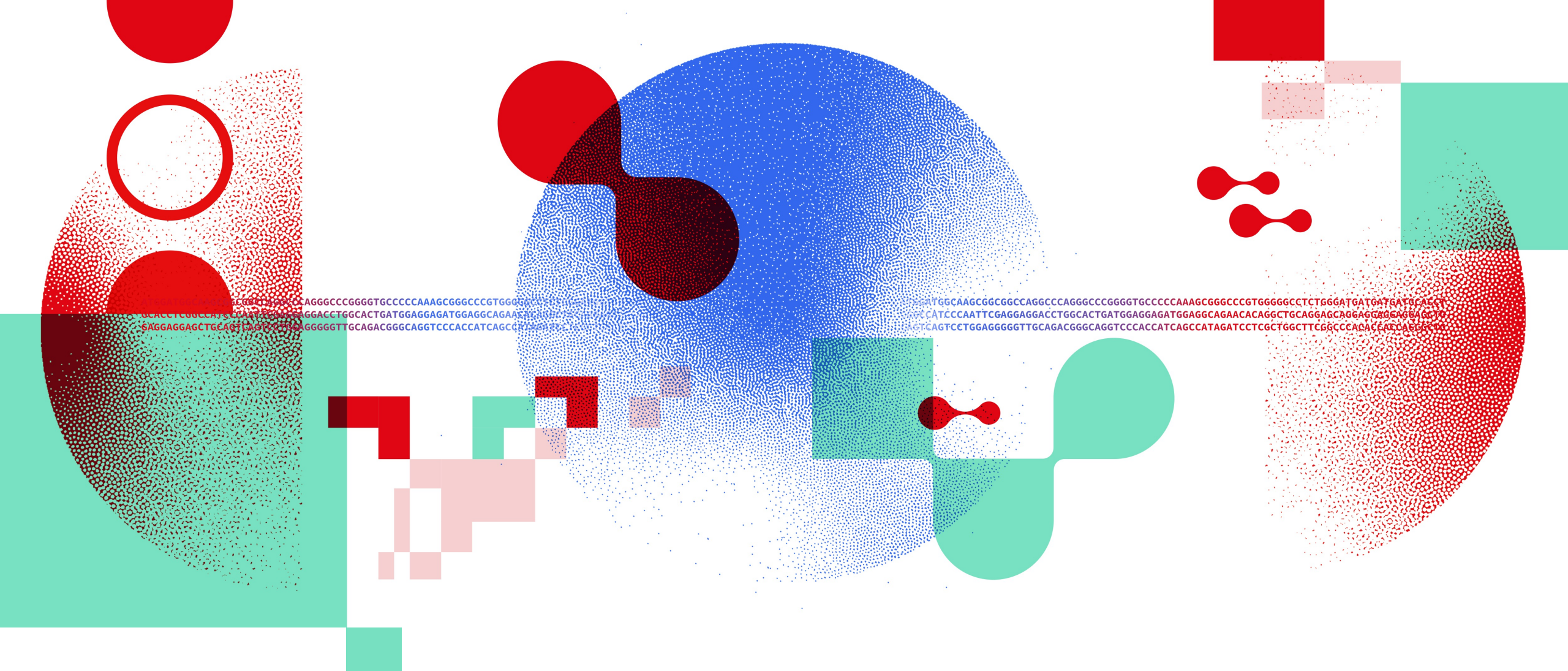
Annotations

- rdfs:label: GNPS consensus MS2 spectrum lang
- rdfs:comment: An MS2 spectrum corresponding to the GNPS consensus spectrum of 1 or more LCMSFeature. Recommendation: instances of this class are ideally built as follows: "https://metabolomics-usi.ucsd.edu/dashinterface/?usi1={id}", e.g.: "https://metabolomics-usi.ucsd.edu/dashinterface/?usi1=mzspec:MassIVE:TASK-773effaad17c4124b9ea9004eaa3f526-spectra/specs_ms.mgf:scan:1028" lang

Parents

- MS2 spectrum

Relationships



Thank you!

DATA SCIENTISTS FOR LIFE

kru@sib.swiss

