

Collaborative OPen Omics (COPO) - Facilitating Metadata Interoperability with Standards, Formats and APIs



Email: ei.copo@earlham.ac.uk
API: www.copo-project.org/api
API docs: www.tinyurl.com/copo-docs-api-usage

 **Aaliyah Providence, Debby Ku, Felix Shaw**

COPO is an online metadata broker for researchers to describe and deposit data (samples, images, sequence reads, assemblies, annotations and other documents) following FAIR (Findability, Accessibility, Interoperability, Reusability) principles. It supports a range of community derived metadata standards, data formats, and API methods, ensuring accurate data description, discoverability, and interoperability across systems. This promotes enhanced collaboration, better data management, and long-term preservation, making scientific outputs more accessible and reusable.

COPO offers diverse metadata standards (Minimum Information about any (x) Sequence (MIxS), Darwin Core (DwC), Tree of Life (ToL), European Nucleotide Archive (ENA)) and supports powerful APIs with outputs in JSON, Research Object Crate (RO-Crate), and CSV formats. By promoting seamless data sharing and interoperability, COPO adheres to FAIR principles, enhancing collaboration, data reuse, and scientific discovery across research domains.

API Formats

```
{
  "@id": "http://identifiers.org/taxonomy:186706",
  "@type": "TAXON",
  "name": "Pseudotaxiphylum elegans",
  "parentTaxon": {
    "@id": "https://copo-project.org/api/sample_field/ORDER_OR_GROUP/"
  },
},
{
  "@id": "#placehzHPvmFTZf",
  "@type": "Place",
  "name": "United Kingdom | Main Argyll | Strath Eachaig | S of Benmore",
  "latitude": "56.01138889",
  "longitude": "-4.981944444"
},
```

RO-Crate format for sample, SAMEA111382928

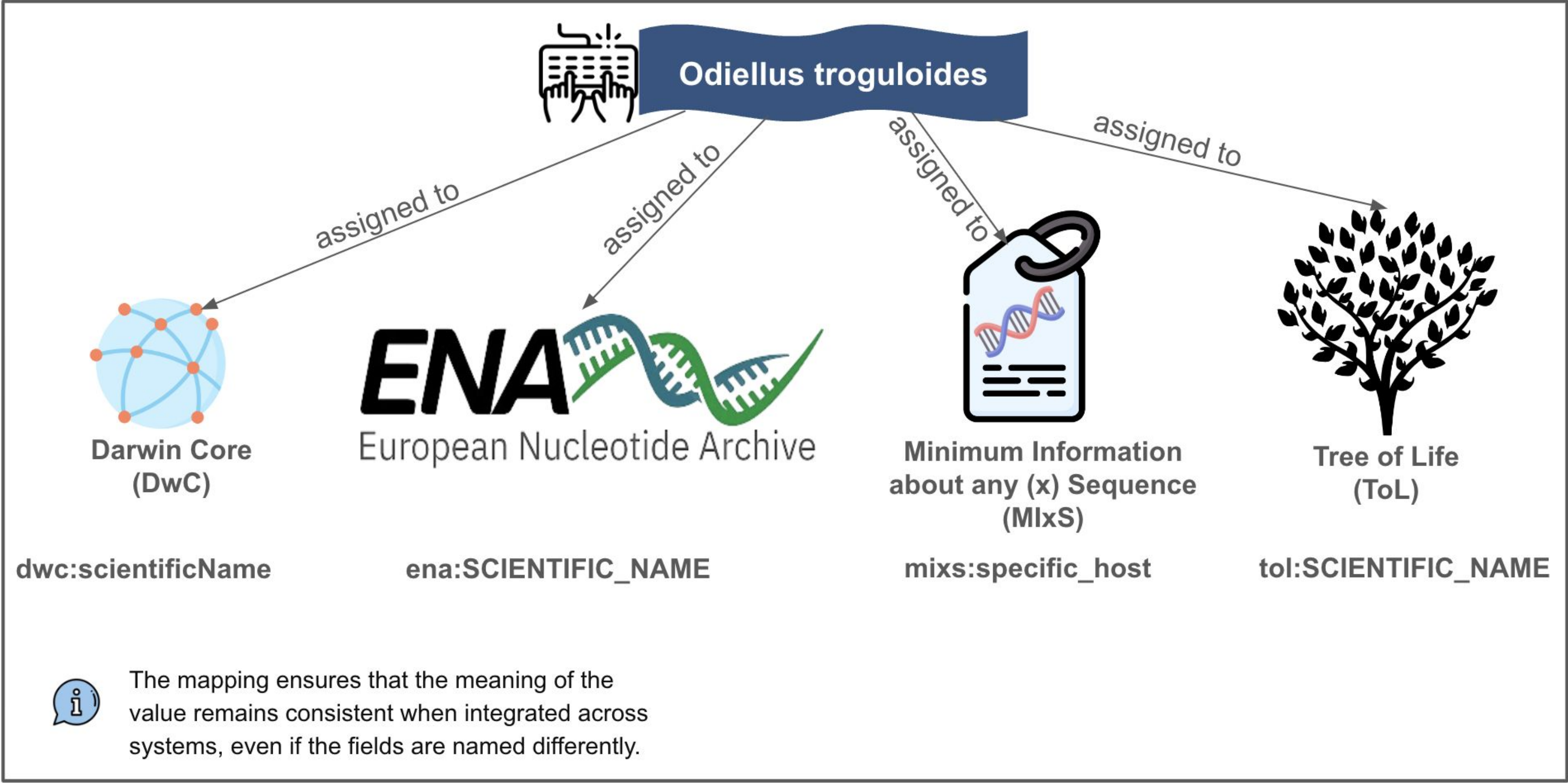
SPECIMEN_ID	ORDER_OR_GROUP	FAMILY	GENUS	TAXON_ID	SCIENTIFIC_NAME
EDTOL01673	Jungermanniales	Gymnomitriaceae	Marsupella	209795	Marsupella aquatica
EDTOL03220	Hypnales	Plagiotheciaceae	Pseudotaxiphylum	186706	Pseudotaxiphylum elegans
EDTOL03258	Anthocerotales	Anthocerotaceae	Anthoceros	3234	Anthoceros punctatus
EDTOL03259	Anthocerotales	Anthocerotaceae	Anthoceros	3234	Anthoceros punctatus
EDTOL03274	Jungermanniales	Saccogynaceae	Saccogyna	362792	Saccogyna viticulosa
EDTOL03275	Jungermanniales	Saccogynaceae	Saccogyna	362792	Saccogyna viticulosa
EDTOL03291	Porellales	Lejeuneaceae	Colura	1187000	Colura calyptrifolia
EDTOL03292	Porellales	Lejeuneaceae	Colura	1187000	Colura calyptrifolia

CSV format for samples

```
"SPECIMEN_ID": "EDTOL03242",
"ORDER_OR_GROUP": "Hookeriales",
"FAMILY": "Daltoniaceae",
"GENUS": "Daltonia",
"TAXON_ID": "759697",
"SCIENTIFIC_NAME": "Daltonia splachnoides",
"TAXON_REMARKS": "",
"INFRA_SPECIFIC_EPITHET": "",
"CULTURE_OR_STRAIN_ID": "",
"COMMON_NAME": "Irish Daltonia",
"COLLECTOR_AFFILIATION": "ROYAL BOTANIC GARDEN EDINBURGH",
"HAZARD_GROUP": "HG1",
"REGULATORY_COMPLIANCE": "Y",
"COLLECTION_LOCATION": "United Kingdom | Main Argyll | Strath Eachaig | S of Benmore"
```

JSON format for sample, SAMEA111382928

Mapping Metadata Standards



Field value mapping for “Odiellus troguloides” across different standards

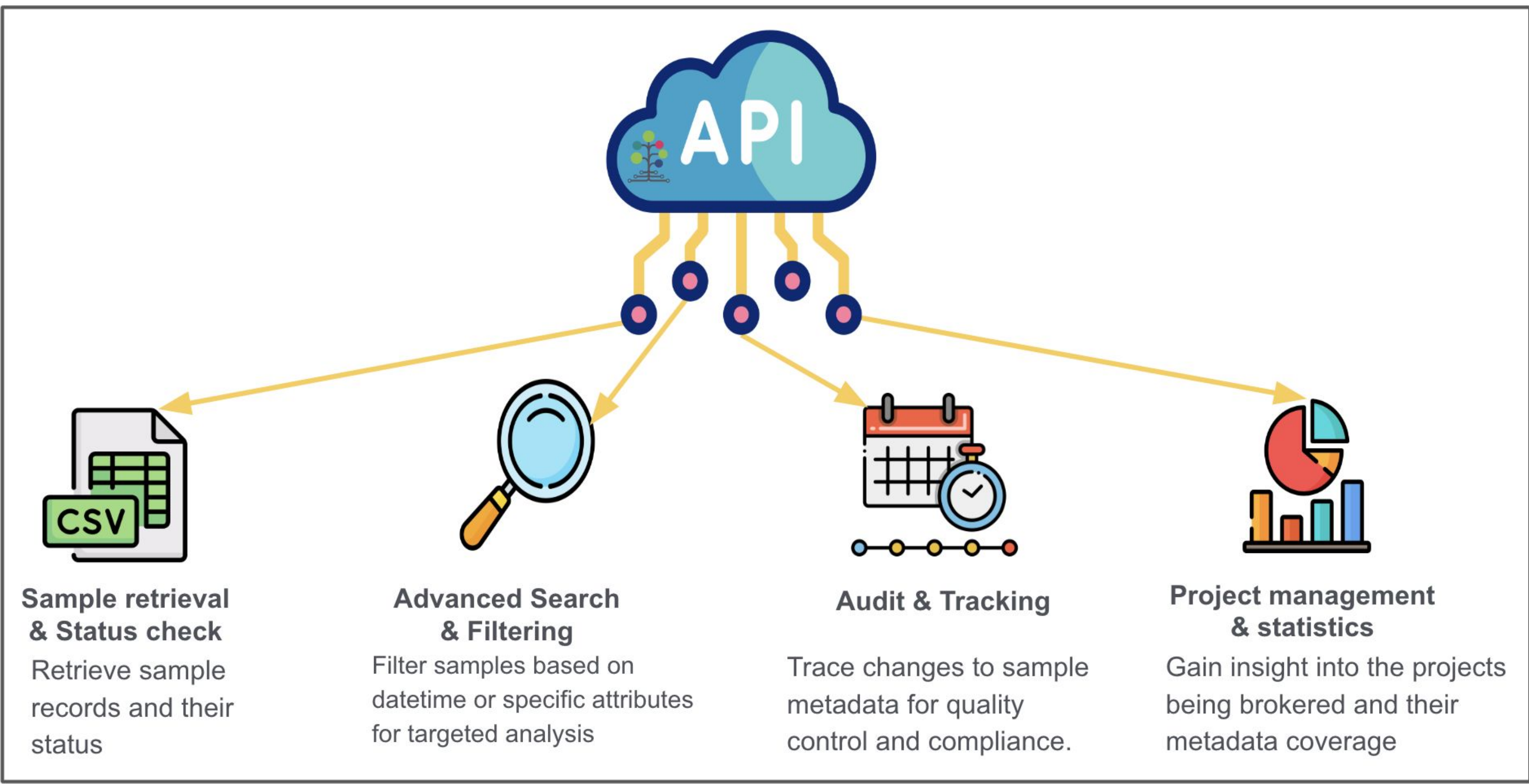
```
measurementID : "PM_A197"
materialSampleID : "PM_A197 | Legs L"
institutionCode : "Departmanet de Genetica, Microbiologia i Immunologia"
order : "Opiliones"
family : "Phalangidae"
genus : "Odiellus"
taxonID : "48695"
scientificName : "Odiellus troguloides"
taxonRemarks : ""
infraspecificEpithet : ""
occurrenceID : ""
vernacularName : "Harvestman|Opili\u00f3n"
lifeStage : "ADULT"
sex : "FEMALE"
organismScope : "LEG"
relationshipRemarks : ""
eventDate : ""
eventTime : "12:00"
higherGeography : "Spain|Catalonia|Lleida"
decimalLatitude : "41.81348"
decimalLongitude : "1.40535"
```

DwC field name map for sample, SAMEA115168223

```
specimen_id : "ERGA_MD_5301_00001"
sample coordinator affiliation : "Departmanet de Genetica, Microbiologia i Immunologia"
TAXON_ID : "48695"
SCIENTIFIC_NAME : "Odiellus troguloides"
culture_or_strain_id : ""
lifestage : "ADULT"
sex : "FEMALE"
organism part : "LEG"
relationship : ""
GAL : "CENTRO NACIONAL DE AN\u00c1LISIS GEN\u00f3MICO"
GAL_sample_id : "PM_A197"
collecting institution : "Departmanet de Genetica, Microbiologia i Immunologia"
collection date : "2023-11-04"
geographic location (latitude) : "41.81348"
geographic location (longitude) : "1.40535"
```

ENA field name map for sample, SAMEA115168223

API Capabilities



An overview of the API functions that COPO offers

Acknowledgements

This work was partnered with the ELIXIR-United Kingdom (ELIXIR-UK) node and supported by Biotechnology and Biological Sciences Research Council (BBSRC), part of UK Research and Innovation, through the Cellular Genomics (CELLGEN) Grant BBS/E/ER/230001A, Data Infrastructure and Algorithms Group Grant EI-G-Data, Decoding Biodiversity (DECOD): Development of Novel Experimental and Bioinformatic Tools for Genomic Diversity and Analysis Grant BBS/E/ER/230002A and EI - Darwin Tree of Life Grant GP182 at the Earlham Institute.

