



biodiversa+
European Biodiversity Partnership

Making Research Data FAIR: An overview of (Meta)data Harmonisation by

Senem Önen Tarantini &
Martina Pulieri

Capacity Building Workshop on Darwin Core
07-06-2024



FAIRness in biodiversity concept

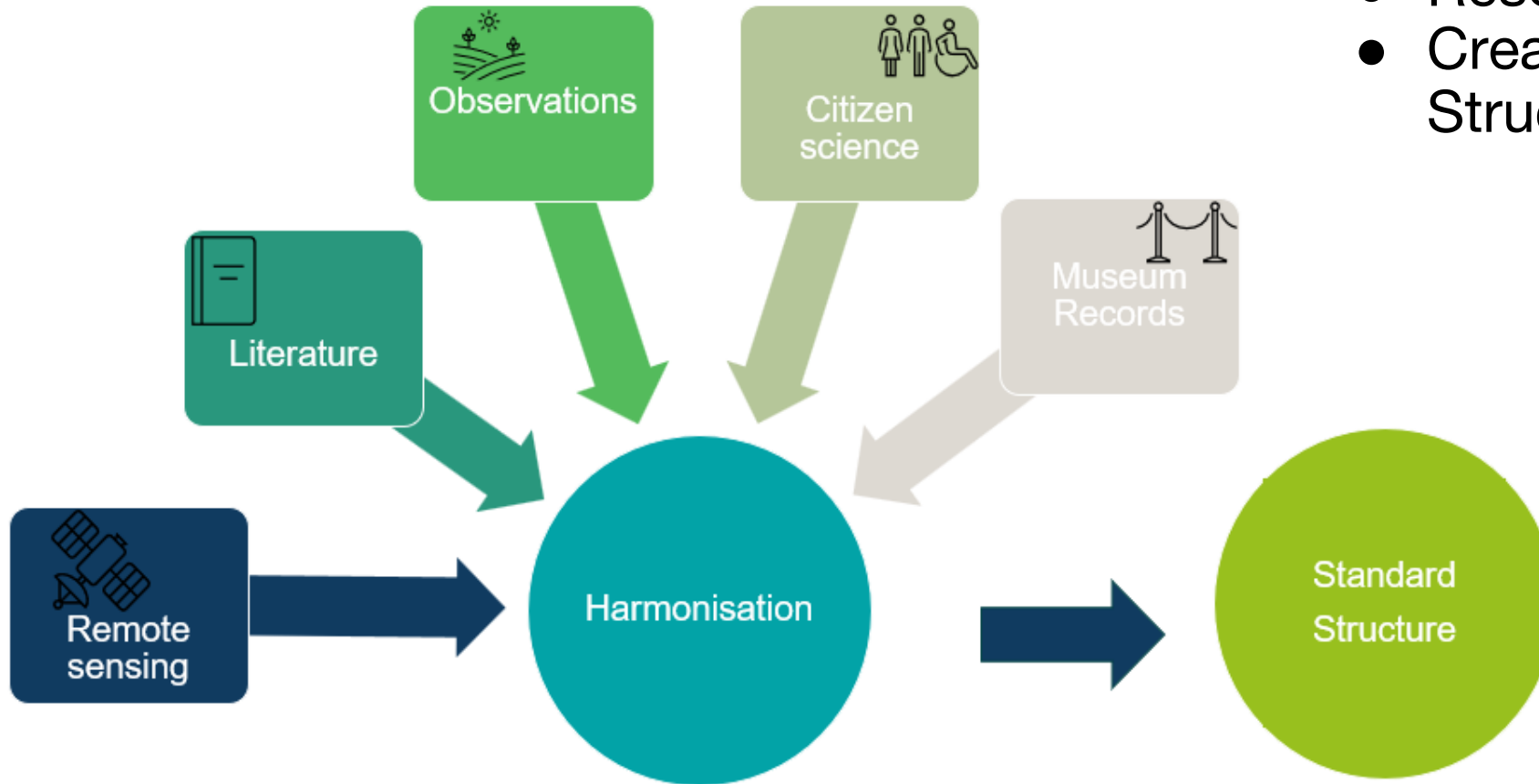
- **Findable:** Detailed metadata; descriptive information about the data and registering datasets in public repositories with search functionalities. Such as species base, area base, habitat base etc.
That facilitates findability of the data both human and computers.
- **Accessible:** Using open data formats whenever possible, and providing clear instructions on how to access and download the data . Moreover data should be stored -long term- so it could be easily accessed
- **Interoperable:** Biodiversity data comes from diverse sources, using different methods and formats. Interoperability ensures that data from various studies can be integrated and compared. Using standardized vocabularies, units, and data structures guaranty interoperability.
- **Reusable:** Data should be usable in new research. This means providing comprehensive documentation about the data collection process, using standard methods, units and also having appropriate licence .

(Meta)data harmonisation

- **Data harmonisation** is the process of bringing together data from various sources and transforming them into a consistent format.
- **Metadata harmonisation** ensures the descriptive information (metadata) about the data is consistent across different sources. The aim of metadata harmonisation is creating a universal instruction manual for biodiversity data.

Data harmonisation

- Unifying formats
- Resolving incoherency
- Creating a Standardize Structure



Data harmonisation: Creating a standard structure

Relazione finale Chiroterri FSP light .XLSX

File Modifica Visualizza Inserisci Formato Dati Strumenti Guida

Menu 100% Arial 12 B I A

1:F1 POR PUGLIA 2000-2006 - MISURA 1.6

	A	B	C	D	E	F
26	L'elenco delle cavità prospectate e della chiroterrofauna rinvenuta è in Appendice.					
27	1. Evidenze pregresse					
28	Sono state raccolte 241 segnalazioni, di cui 187 inequivocabili. Sulla base di queste in Puglia sono segnalate 18 specie (Tab. 1).					
29	Specie (nome comune, nome scientifico)	Berna	Bonn	Habitat	IUCN	
30	Ferro di cavallo maggiore, <i>Rhinolophus ferrumequinum</i>	2	2	2,4	LR:nt	
31	Ferro di cavallo minore, <i>Rhinolophus hipposideros</i>	2	2	2,4	VU:A2c	
32	Ferro di cavallo Euriale, <i>Rhinolophus euryale</i>	2	2	2,4	VU:A2c	
33	Ferro di cavallo di Mehely, <i>Rhinolophus mehely</i>	2	2	2,4	VU:A2c	
34	Serotino comune, <i>Epseticus serotinus</i>	2	2	4	LR:lc	
35	Pipistrello di Savi, <i>Hypsugo savii</i>	2	2	4	LR:lc	
36	Miniottero, <i>Miniopterus schreibersi</i>	2	2	2,4	LR:nt	
37	Vespertilio di Capaccini, <i>Myotis capaccini</i>	2	2	2,4	VU:A2c	
38	Vespertilio di Daubenton, <i>Myotis daubentoni</i>	2	2	4	LR:lc	
39	Vespertilio smarginato, <i>Myotis emarginatus</i>	2	2	2,4	VU:A2c	
40	Vespertilio maggiore/minore, <i>Myotis myotis/blythii</i>	2	2	2,4	LR:nt/lc	
41	Nottola di Leisler, <i>Nyctalus leisleri</i>	2	2	4	LR:nt	
42	Nottola comune, <i>Nyctalus noctula</i>	2	2	4	VU:A2c	
43	Pipistrello albolimbato, <i>Pipistrellus kuhlii</i>	2	2	4	LR:lc	
44	Pipistrello nano/pigmeo, <i>Pipistrellus pipistrellus/pygmaeus</i>	2	2	4	LR:lc/DD	
45	Orecchione bruno/grigio, <i>Plecotus auritus/austriacus</i>	2	2	4	LR:lc	
46	Molosso di Cestoni, <i>Tadarida teniotis</i>	2	2	4	LR:lc	

Tabella 1. Check-list, status legale (Convenzione di Berna, Convenzione di Bonn, Direttiva Habitat) e minaccia (IUCN) delle singole specie di Chiroterri. Legenda: Berna 2: Allegato 2 della Convenzione di Berna; Bonn 2: Allegato 2 della Convenzione di Bonn; Habitat 2: Allegato 2 della Direttiva Habitat; Habitat 4: Allegato 4 della Direttiva Habitat; IUCN: LR= Low Risk (Basso Rischio); VU= Vulnerable (Vulnerabile); DD=

- In Italian
- Limited geographical information
- Sampling protocol within pdf text
- Limited taxonomic information

Data harmonisation: Creating a standard structure

Relazione finale Chiroterri final .XLSX

File Modifica Visualizza Inserisci Formato Dati Strumenti Guida

Menu 100% 123 Calibri 10 B I A

P10

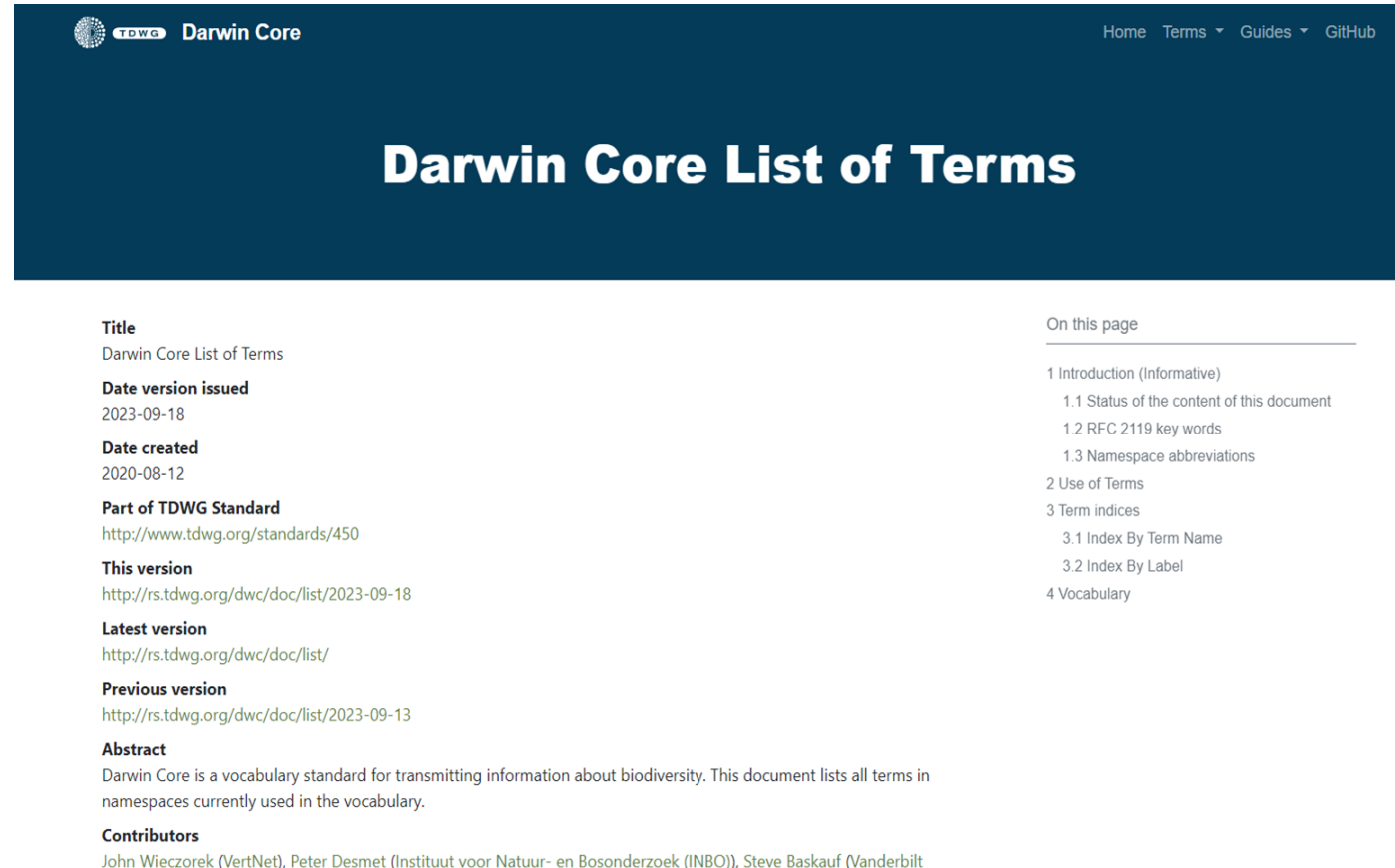
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	countryCode	stateProvince	location	municipality	verbatimlocality	locality	verbatimidentification	phylum	class	order	family	genericName	air pollution	feces	organismQuantity
2	darwin core	darwin core	darwin core	darwin core	darwin core	darwin core	darwin core	darwin core	darwin core	darwin core	darwin core	darwin core	NALT Core	NALT Core	darwin core
3	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/	http://rs.tdwg.org/
4	IT	Puglia	5	BA	Grotta della	Putignano	Chiroptera sp	Chordata	Mammalia	Chiroptera			2	N	A
5	IT	Puglia	8	BA	Grotte di Castellana	Castellana Grotte	Myotis blythii	Chordata	Mammalia	Chiroptera	Vespertilionidae	Myotis	4	N	B
6	IT	Puglia	8	BA	Grotte di Castellana	Castellana Grotte	Myotis capaccinii	Chordata	Mammalia	Chiroptera	Vespertilionidae	Myotis	4	N	B
7	IT	Puglia	8	BA	Grotte di Castellana	Castellana Grotte	Myotis blythii	Chordata	Mammalia	Chiroptera	Vespertilionidae	Myotis	4	N	B
8	IT	Puglia	8	BA	Grotte di Castellana	Castellana Grotte	Rhinolophus euryale	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	4	N	B
9	IT	Puglia	16	BA	Grave Santa Lucia	Monopoli	Rhinolophus	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	4	T	A
10	IT	Puglia	18	BA	Grotta di Cristo	Cassano delle Murge	Myotis myotis/blythii	Chordata	Mammalia	Chiroptera	Vespertilionidae	Myotis	1	T	C
11	IT	Puglia	18	BA	Grotta di Cristo	Cassano delle Murge	Rhinolophus	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	1	T	C
12	IT	Puglia	18	BA	Grotta di Cristo	Cassano delle Murge	Rhinolophus hipposideros	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	1	T	C
13	IT	Puglia	23	BAT	Pozzo naturale presso	Spinazzola	n.d.	Chordata	Mammalia	Chiroptera			4	N	A
14	IT	Puglia	25	BA	Pulo di Altamura	Altamura	Rhinolophus	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	2	N	B
15	IT	Puglia	27	BA	Pulo di Molfetta	Molfetta	n.d.	Chordata	Mammalia	Chiroptera	n.d.	n.d.	4	N	A
16	IT	Puglia	30	BAT	Grotta di S. Michel	Minervino Murge	Rhinolophus euryale	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	n.a	N	A
17	IT	Puglia	31	BA	Grave di Faraualla	Gravina in Puglia	Myotis blythii	Chordata	Mammalia	Chiroptera	Vespertilionidae	Myotis	2	N	B
18	IT	Puglia	31	BA	Grave di Faraualla	Gravina in Puglia	Myotis myotis	Chordata	Mammalia	Chiroptera	Vespertilionidae	Myotis	2	N	B
19	IT	Puglia	38	BA	Grotta di	Acquaviva delle Fonti	n.d.	Chordata	Mammalia	Chiroptera	n.d.	n.d.	1	N	A
20	IT	Puglia	77	BA	Grotta della Masse	Conversano	Miniopterus schreibersii	Chordata	Mammalia	Chiroptera	Miniopteridae	Miniopterus	1	N	B
21	IT	Puglia	77	BA	Grotta della Masse	Conversano	Rhinolophus euryale	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	1	N	B
22	IT	Puglia	77	BA	Grotta della Masse	Conversano	Rhinolophus	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	1	N	B
23	IT	Puglia	113	LE	Grotta Grande di	Gagliano del	Rhinolophus	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	1	P	C
24	IT	Puglia	113	LE	Grotta Grande di	Gagliano del	Miniopterus schreibersii	Chordata	Mammalia	Chiroptera	Miniopteridae	Miniopterus	1	P	C
25	IT	Puglia	152	LE	Grotta della	Otranto	Miniopterus schreibersii	Chordata	Mammalia	Chiroptera	Miniopteridae	Miniopterus	1	N	E
26	IT	Puglia	152	LE	Grotta della	Otranto	Rhinolophus	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	1	N	E
27	IT	Puglia	201	FG	Grotta di	San Marco	Rhinolophus euryale	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	4	N	C
28	IT	Puglia	201	FG	Grotta di	San Marco	Rhinolophus	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	4	N	C
29	IT	Puglia	209	FG	Grotta Campana	Mattinata	n.d.	Chordata	Mammalia	Chiroptera	n.d.	n.d.	4	N	A
30	IT	Puglia	265	FG	Grotta Occhiopinto	Manfredonia	Rhinolophus	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	2	N	E
31	IT	Puglia	265	FG	Grotta Occhiopinto	Manfredonia	Rhinolophus hipposideros	Chordata	Mammalia	Chiroptera	Rhinolophidae	Rhinolophus	2	N	E
32	IT	Puglia	265	FG	Grotta Occhiopinto	Manfredonia	Myotis myotis	Chordata	Mammalia	Chiroptera	Vespertilionidae	Myotis	2	N	E
33	IT	Puglia	265	FG	Grotta Occhiopinto	Manfredonia	Miniopterus schreibersii	Chordata	Mammalia	Chiroptera	Miniopteridae	Miniopterus	2	N	E
34	IT	Puglia	266	FG	Complesso Scaloria (26	Manfredonia	n.d.	Chordata	Mammalia	Chiroptera	n.d.	n.d.	2	M	A
35	IT	Puglia	276	FG	Grava di	San Giovanni	Myotis blythii, Myotis	Chordata	Mammalia	Chiroptera	Vespertilionidae	Myotis	2	P	D

- In English
- Detailed geographical information
- Detailed taxonomic information
- Detailed sampling findings

How (meta)data harmonisation works in biodiversity research

Data harmonisation focuses on bringing the data itself into a consistent format, while metadata harmonisation specifically tackles ensuring the descriptive information (metadata) about the data is consistent across different sources .

- Using standardised vocabularies: Darwin Core (DwC) is a widely used vocabulary for sharing information on biodiversity.
- Standard meta(data) structure
- Shared definition for describing the (meta)data



The screenshot shows the 'Darwin Core List of Terms' page. The header is dark blue with the TDWG logo and 'Darwin Core' text on the left, and navigation links 'Home', 'Terms', 'Guides', and 'GitHub' on the right. The main title 'Darwin Core List of Terms' is centered in large white font. Below the title, the page is divided into two columns. The left column contains metadata: Title (Darwin Core List of Terms), Date version issued (2023-09-18), Date created (2020-08-12), Part of TDWG Standard (with a URL), This version (with a URL), Latest version (with a URL), Previous version (with a URL), Abstract (a paragraph about the standard), and Contributors (a list of names). The right column, titled 'On this page', contains a table of contents with links to sections: 1 Introduction (Informative), 1.1 Status of the content of this document, 1.2 RFC 2119 key words, 1.3 Namespace abbreviations, 2 Use of Terms, 3 Term indices, 3.1 Index By Term Name, 3.2 Index By Label, and 4 Vocabulary.

Title
Darwin Core List of Terms

Date version issued
2023-09-18

Date created
2020-08-12

Part of TDWG Standard
<http://www.tdwg.org/standards/450>

This version
<http://rs.tdwg.org/dwc/doc/list/2023-09-18>

Latest version
<http://rs.tdwg.org/dwc/doc/list/>

Previous version
<http://rs.tdwg.org/dwc/doc/list/2023-09-13>

Abstract
Darwin Core is a vocabulary standard for transmitting information about biodiversity. This document lists all terms in namespaces currently used in the vocabulary.

Contributors
John Wieczorek (VertNet), Peter Desmet (Instituut voor Natuur- en Bosonderzoek (INBO)), Steve Baskauf (Vanderbilt)

On this page

- 1 Introduction (Informative)
 - 1.1 Status of the content of this document
 - 1.2 RFC 2119 key words
 - 1.3 Namespace abbreviations
- 2 Use of Terms
- 3 Term indices
 - 3.1 Index By Term Name
 - 3.2 Index By Label
- 4 Vocabulary

Metadata harmonisation: Creating a standard structure

End Date	A single time stamp indicating the end of a time period.	1			
Taxonomic Coverage		0..∞		Yes	
Taxon Rank Name	The name of the taxonomic rank for which the Taxon rank value is provided (e.g., kingdom, class, order, family, genus, etc.).	Phylum, Class, Order, Family, Genus, Species			
Taxon Rank Value	The taxonomic rank name being described (e.g., Acer).	1			
Common Name	Specification of applicable common names (e.g., insects, vertebrate, grasses, etc.).	Reptilia	Amphibia		
Taxon ID	The identifier for this taxon from an authority, such as ITIS (https://www.itis.gov) or USDA Plant Database (https://plants.usda.gov).	https://www.itis.gov/servlet/SingleRpt/SingleRpt?search_topic=TSN&search_value=173747#null	https://www.itis.gov/servlet/SingleRpt/SingleRpt#null		
License Information					
Intellectual Rights	<u>Intellectual property rights regarding usage and licensing of the resource. E.g., This data package is released to the “public domain” under Creative Commons CC0 1.0 “No Rights Reserved” (see: https://creativecommons.org/publicdomain/zero/1.0/).</u>	LifeWatch Metadata			
License Name	The official name of the license applied to the data and metadata described in this metadata record. The name should match the name of a well-known license from the SPDX license vocabulary or a similar persistent vocabulary.	1	Yes		

(Meta)data harmonisation and FAIR (interoperability)

Interoperable data means it can be integrated with different datasets, application and workflows.

It has different layers as

- **Organisational:** Governance, policies
- **Legal:** Licences, copyright, data protection
- **Technical:** Infrastructure, services, technologies
- **Semantic:** Community standards, ontologies & metadata

Towards interoperability

Data interoperability



Implies **semantic interoperability**, which ensures that the *precise format and meaning of exchanged data and information* is preserved and understood throughout exchanges between parties.

Semantic interoperability: problems

- Lack of common explicit definitions

Among different vocabularies there are different definitions about the same concepts.

- Lack of common semantic artefacts across communities

More than 540 semantic artefacts were found with numerous overlapping in topics. A great number were created *ad hoc* for projects and never “connected” to the existing semantic artefacts, impeding interoperability.

- Lack of expertise and skills related to semantics

Negatively influences the use of common definitions.

Data and semantic interoperability: recommendations

- Clear and precise definitions for the concepts

Use of commonly agreed vocabularies to allow an easier data exchange and integration.

- Thoroughly document your data

What, Who, Where, When, How to not lose information and allow the reuse of data.



biodiversa+
European Biodiversity Partnership

EUROPEAN PARTNERSHIP



Co-funded by
the European Union

Thank you!



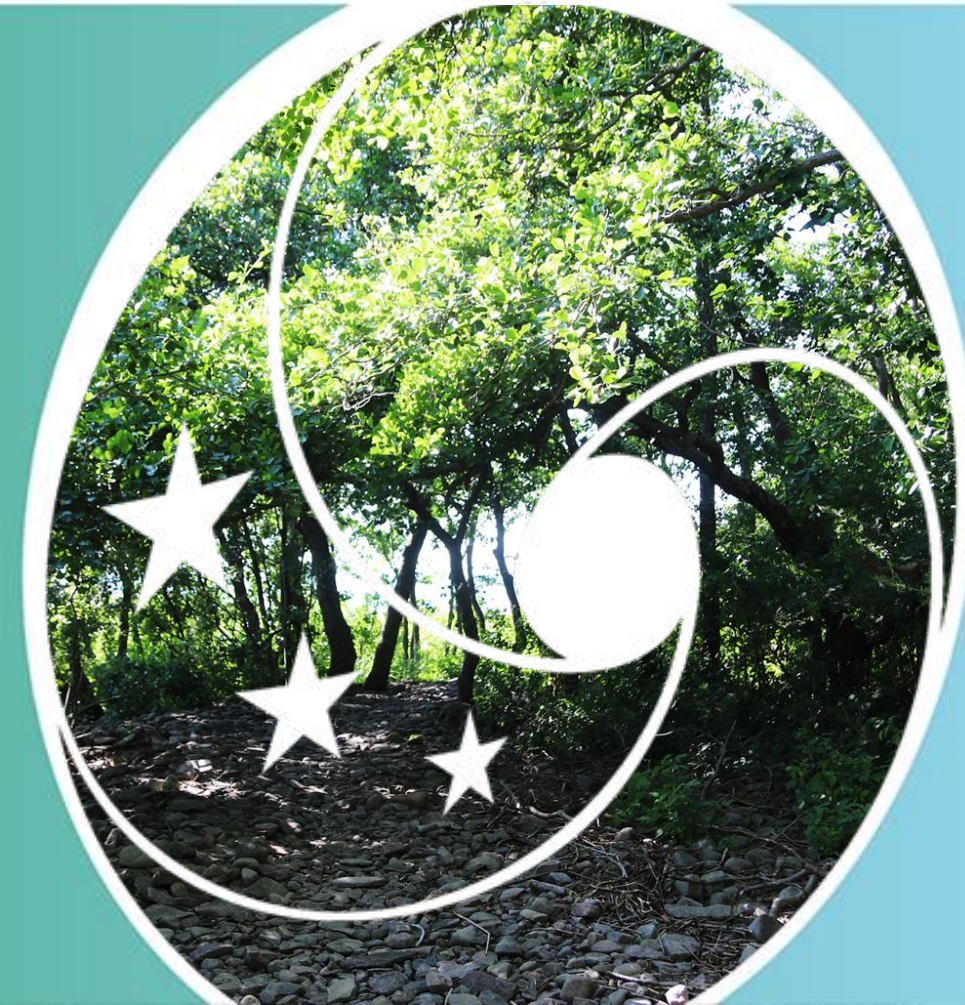
www.biodiversa.eu



contact@biodiversa.eu



BiodiversaPlus



Semantic technologies and their usage for biodiversity data: opportunities and challenges

Naouel Karam

Institute for Applied Informatics (InfAI), Leipzig University

Biodiversa+ Capacity building workshop on Darwin Core standard

7th of June 2024

Agenda

Why

... do we need Semantic technologies?

Ambiguity of Natural Language
Keywords vs Concepts

What

... are terminologies? ... terminologies do we need?

Terminology?

Terminologies for Semantic Search

Terminologies for Metadata Harmonization

Terminologies for Data Interoperability

How

... do we use terminologies?

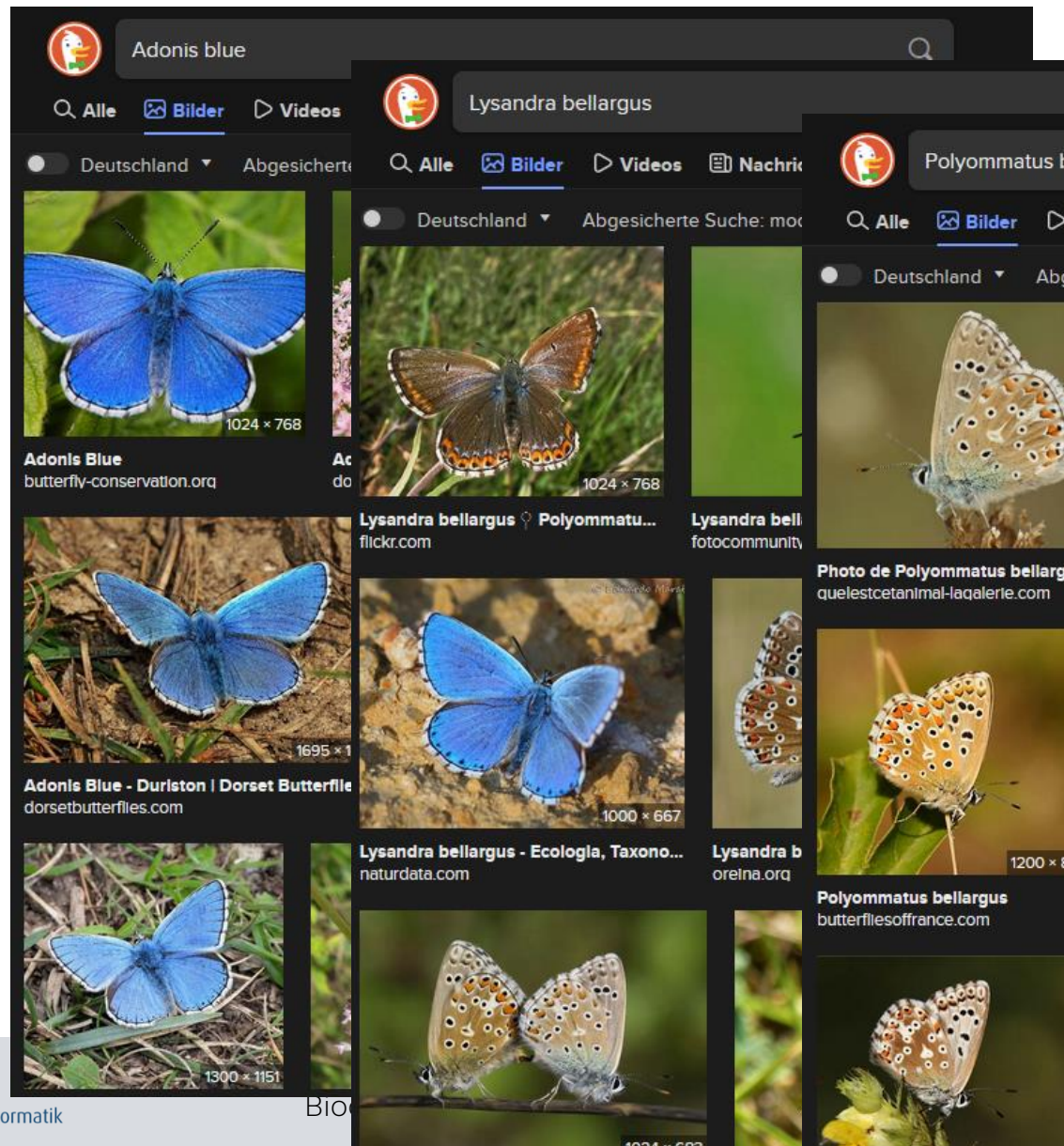
Semantic Services for Biodiversity

Requirements on a Terminology Service

Ongoing Work

Why Ambiguity of Natural Language

> Different words for the same concept (Synonymy)



NCBI Entrez PubMed Nucleotide Protein Genome Stru

Search for as complete name

Display 3 levels using filter: none

Lysandra bellargus

Taxonomy ID: 138070 (for references in articles please use NCBI:txid138070)

current name

Lysandra bellargus (Rottemburg, 1775)

basionym: *Papilio bellargus* Rottemburg, 1775

homotypic synonym: *Polyommatus bellargus*

Genbank common name: **Adonis blue**

NCBI BLAST name: **butterflies**

Rank: **species**

Genetic code: [Translation table 1 \(Standard\)](#)

Mitochondrial genetic code: [Translation table 5 \(Invertebrate Mitochondrial\)](#)

Other names:

heterotypic synonym

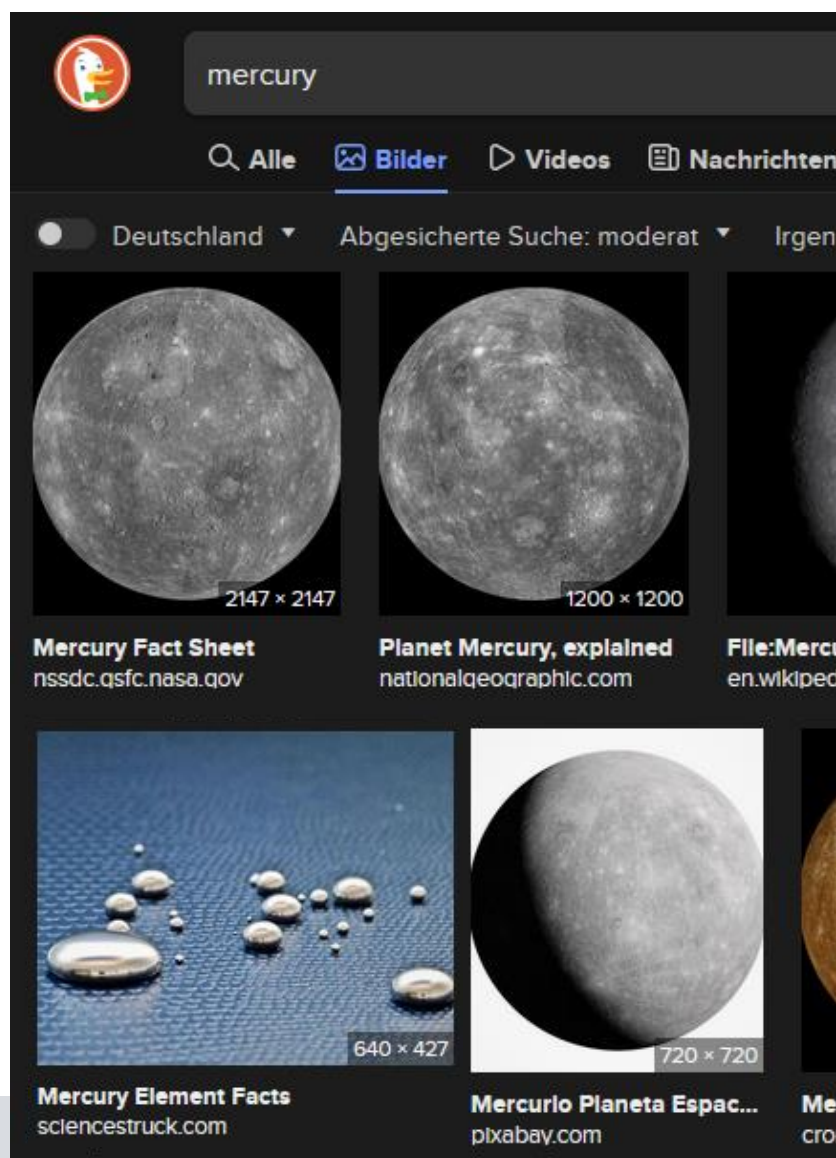
Lysandra belargus

common name(s)

Adonis blue butterfly

Why Ambiguity of Natural Language

> Same word
for different
concepts
(Polysemy)



The ChEBI entry for mercury(0) (CHEBI:16170) is displayed. The header shows the ChEBI logo and the name 'mercury(0)'. Below the header, there are tabs for 'Main', 'ChEBI Ontology', 'Automatic Xrefs', 'Reactions', and 'Pathways'. The 'Main' tab is selected, showing the chemical symbol 'Hg' and the ChEBI Name 'mercury(0)'. The ChEBI ID is 'CHEBI:16170'. The Definition is 'Elemental mercury of oxidation state zero.' The Roles Classification section shows 'Biological Role(s): neurotoxin' and a description: 'A poison that interferes with the functions of the nervous system.'

Why Keywords vs Concepts

Search results for **apis** | **apis mellifica** Western honey bee — Insect

Western honey bee
Insect

Overview Lower classifications Diet Characteristics

Lower classifications

East African lowland hone...	Italian bee	Carniolan honey bee	European dark bee	Buckfast bee	Caucasian honey bee
Apis mellifera adansonii	Cape honey bee	Apis mellifera iberiensis	Russian honey bee	Apis mellifera cecropia	Apis mellifera macedonica

About

The western honey bee or European honey bee is the most common of the 7–12 species of honey bees worldwide. The genus name *Apis* is Latin for "bee", and *mellifera* is the Latin for "honey-bearing" or "honey carrying", referring to the species' production of honey. [Wikipedia](#)

Scientific name: *Apis mellifera*

Lifespan: 30 – 60 days (Female, Worker, In Spring, In Fall), 21 – 32 days (Male, In Spring, In Summer, Drones)

Class: Insecta

Family: Apidae

Order: [Hymenoptera](#)

Domain: Eukaryota

Meaning
(**Semantics**) is
expressed with the
help of knowledge
representation
(**Terminologies**)

What is a Terminology?

informal

formal

Controlled Vocabulary

List of terms

Glossary

List of terms +
informal definitions of
their meaning



Absence Observation
DrawingOrPhotograph
Earth Science Specimen
Fossil
Human Observation
Literature
Living Specimen
Machine Observation
MaterialSample
Mineral Specimen
MultimediaObject
OtherSpecimen
Preserved Specimen
Unknown

Taxonomy

List of terms
organised in a
hierarchical
structure



Aves
ACCIPTRIFORMES
OTIDIFORMES
PASSERIFORMES
Acanthisittidae
Acanthizidae
Thraupidae
Acanthidops
Anisognathus
Bangsia
Nesospiza
Nesospiza acunhae
Nesospiza questi
Nesospiza wilkinsi

Thesaurus

Controlled vocabulary
connected via **relations**
between terms
(e.g. "narrower/broader",
"related term", "synonym")



physical environment
aquatic environment
atmosphere
biosphere

biodiversity

broader biosphere
related animal resources
plant resources
protected species

Ontology

Formal representation of a
set of concepts and the
relationships between
them using **logical axioms**

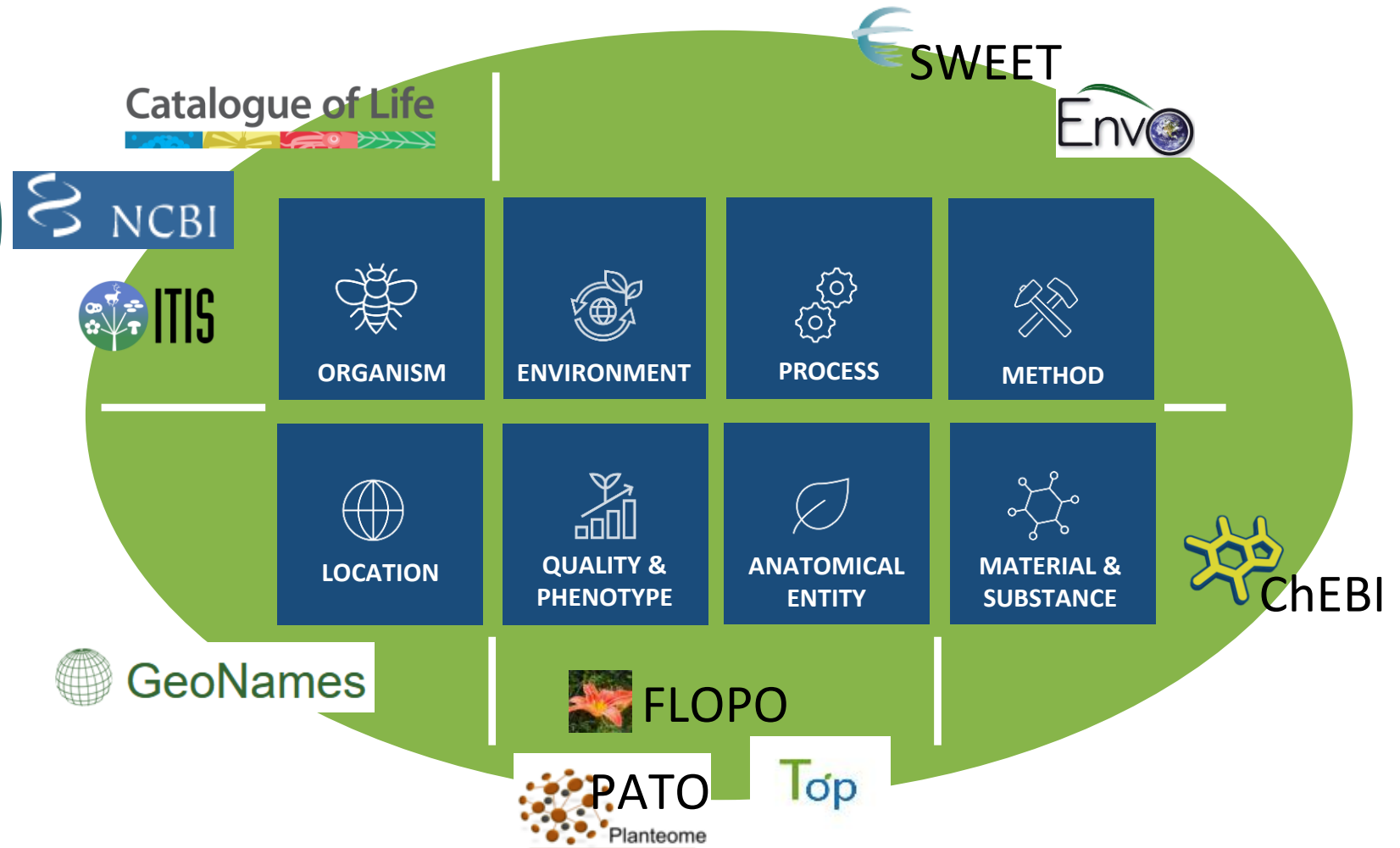


mercury molecular entity
elemental mercury
mercury cation
mercury(0)

Preferred Name	mercury(0)
Synonyms	mercury Hg Hgn mercury(0) Elemental mercury metallic mercury Mercury Hg(0)
monoisotopicmass	201.97064
subClassOf	elemental mercury
has role	some neurotoxin

What Terminologies for Semantic Search

Do butterflies occur in calcareous grassland?
How does agriculture affect the ground water composition?
Is there data on the influence of geographic elevation on the growth rate of Zea mays?



F. Löffler, C. Pfaff, N. Karam, D. Fichtmüller, F. Klan: What do Biodiversity Scholars Search for? Identifying High-Level Entities for Biological Metadata. S4BioDiv@ISWC 2017

What Terminologies for Metadata Harmonization



NFDI 4
BIODIVERSITY

ABCD Access to Biological
Collection Data



```
17 ],
18 "http://rs.tdwg.org/abcd/terms/name": [
19 {
20   "@value": "Botanic Garden and Botanical Museum Berlin-Dahlem"
21 }
22 ],
23 "http://rs.tdwg.org/abcd/terms/abbreviation": [
24 {
25   "@value": "BGBM"
26 }
27 ],
28 "@type": [
29   "http://rs.tdwg.org/abcd/terms/Organization"
30 ]
31 },
32 {
33   "@id": "http://www.bgbm.org/Herbarium_Berolinense",
34   "http://rs.tdwg.org/abcd/terms/hasContact": [
35 {
36   "@id": "http://www.bgbm.org"
37 }
38 ],
39   "http://rs.tdwg.org/abcd/terms/details": [
40 {
41   "@value": "The herbarium of the Botanic Garden and Botanical Museum Berlin-Dahlem (
collection of more than 3.5 million preserved specimens. All plant groups - flowering plant
and lichens - are represented in the collections which are worldwide in scope. Associated w
fruits and seeds, wood samples, and specimens preserved in alcohol. The collections of the
conducted by staff, and through gifts, acquisitions, and exchanges of specimens from other
42   }
43 ],
44 "@type": [
45   "http://rs.tdwg.org/abcd/terms/DataSet"
46 ],
47 "http://rs.tdwg.org/abcd/terms/title": [
48 {
49   "@value": "Herbarium Berolinense"
50 }
```

Schema.org

3 {

4 {

5 "@context": "http://schema.org/",

6 "@id": "https://doi.org/10.1594/PANGAEA.718130",

7 "@type": "Dataset",

8 "identifier": "https://doi.org/10.1594/PANGAEA.718130",

9 "url": "https://doi.pangaea.de/10.1594/PANGAEA.718130",

10 "creator": [

11 {

12 "@id": "https://orcid.org/0000-0001-8807-4597",

13 "@type": "Person",

14 "name": "Frédéric Gazeau",

15 "familyName": "Gazeau",

16 "givenName": "Frédéric",

17 "identifier": "https://orcid.org/0000-0001-8807-4597",

18 "email": "gazeau@obs-vlfr.fr"

19 },

20 {

21 "@type": "Person",

22 "name": "Christophe Quiblier",

23 "familyName": "Quiblier",

24 "givenName": "Christophe"

25 },

26 {

27 "@type": "Person",

28 "name": "Jeroen M Jansen",

29 "familyName": "Jansen",

30 "givenName": "Jeroen M"

31 },

32]

Property	Type	Description
abstract	Text	An abstract is a short description that summarizes a CreativeWork.
accessMode	Text	The human sensory perceptual system or cognitive faculty through which a person may process or perceive information. Values should be drawn from the approved vocabulary.

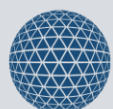
NFDI 4 BIODIVERSITY

QUDT

[Tópico](#)



Plant Trait Database



InfAI[®]
Institut für Angewandte Informatik

Biodive

How Semantic Services for Biodiversity

Goal: Development of **semantic standards**, **services** and **tools** for the discovery, integration, transformation and harmonization of heterogeneous data

2014 2015



gfbio Terminology Service
Up and running
<https://terminologies.gfbio.org>

2020



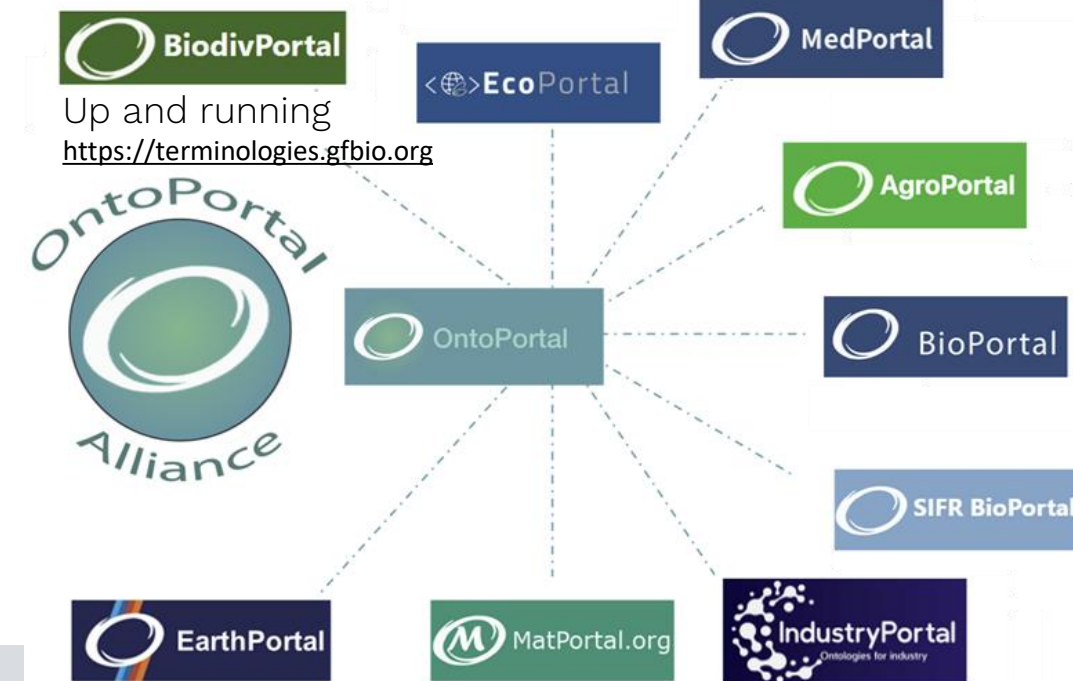
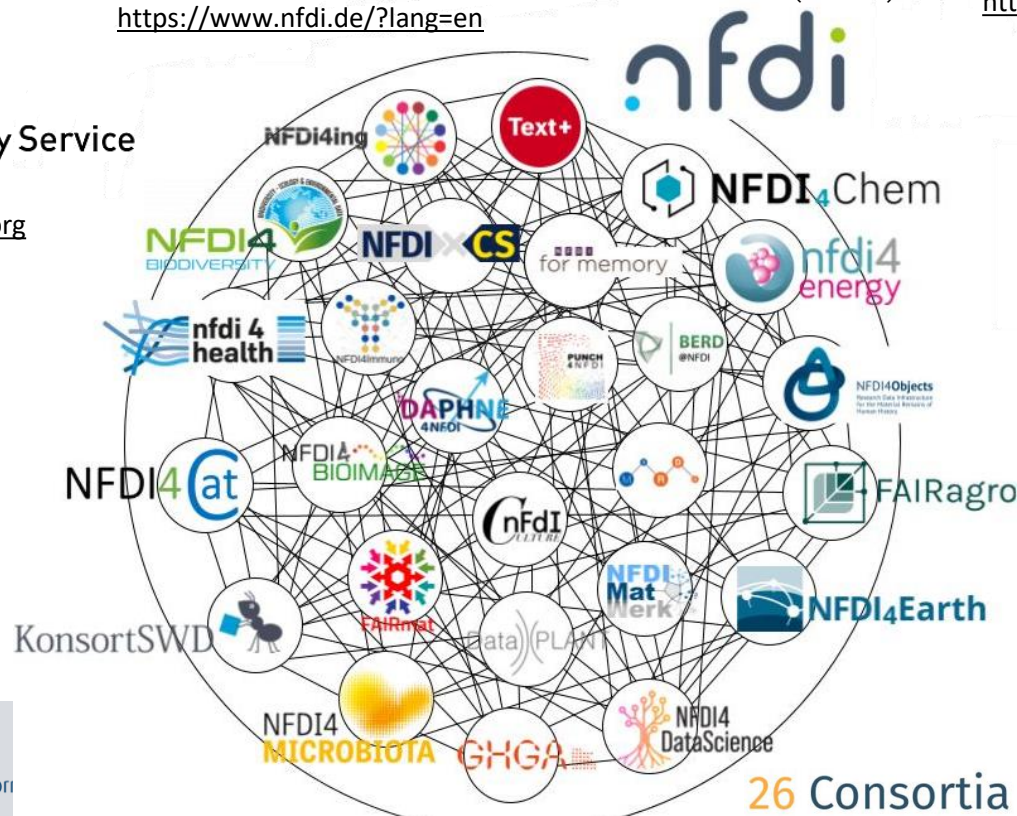
NFDI 4
BIODIVERSITY

National Research Data Infrastructure (NFDI)
<https://www.nfdi.de/?lang=en>

2022

Joined the
OntoPortal Alliance
<https://ontoportal.org/about/>

2023



Up and running
<https://terminologies.gfbio.org>



26 Consortia

win Core standard, 7th of June 2024

Slide 10



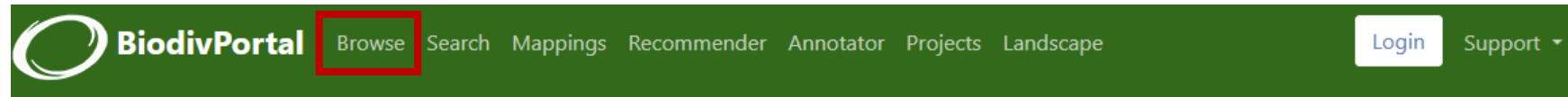
- ▶ **Repository** to store, search, browse and visualize terminologies
- ▶ Offer efficient **semantic annotation** for text and tabular data
- ▶ Automatically generate and store **mappings** between terminologies
- ▶ Provide a single access point (**API**) to heterogenous terminological resources
- ▶ Offer an environment for the **development**, curation, and publication of project terminologies
- ▶ Provide efficient terminology **versioning** and evolution mechanisms
- ▶ Offer **Community feedback** mechanisms

How Requirements on a Terminology Service



NFDI 4
BIODIVERSITY

Repository to store, search, browse and visualize terminologies



Browse

Browse the library of ontologies ?

Search... Showing 29 of 29 Sort: Popular

Submit New Ontology

Entry Type

☒ Ontology (29)

☐ Ontology View (0)

Uploaded in the Last

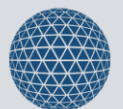
Category

☐ anatomical (4)

Thysanoptera Ontology (THYSANOPTERA) Working list: taxon names of the order Thysanoptera (Insecta). Uploaded: 7/31/23	classes 8,367	FAIR score 244
Bohlmann Ontology (BOHLMANN) An ontology listing the natural substances that occur in the plant family of Compositae (Asteraceae). Uploaded: 11/6/23	classes 53,761	FAIR score 244
The lithologs rock names ontology for igneous rocks (LIT_I) The lithologs ontology for igneous rocks provides SKOSified information on the classification of igneous rocks extracted from the IUGS recommendations published by Le Maitre (2002): https://www.iugc.org/terminology/ Uploaded: 7/31/23	concepts 1,871	FAIR score 225

<https://biodivportal.globe.org/ontologies>

Biodiversity Ontology building workshop on Darwin Core standard, 7th of June 2024



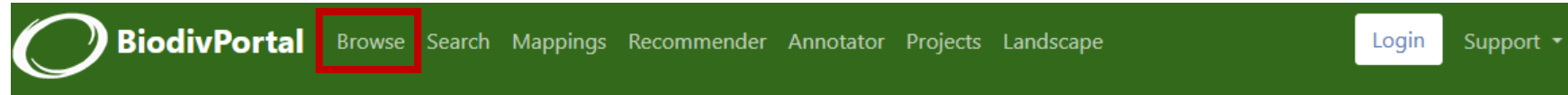
InfAI®
Institut für Angewandte Informatik

How Requirements on a Terminology Service



NFDI 4
BIODIVERSITY

Repository to store, search, browse and visualize terminologies



Browse

Browse the library of ontologies ?

contributing
terminologies

Submit New
Ontology

Entry Type

☒ Ontology (29)

NFDI4Biodiversity
Community

Submit New Ontology

Name *

Acronym *

Viewing Restriction ?

* fields are required

Public

Administrat

naouel x

Categories

Select category (domain)

Groups

Select group

This ontology is a view of: ☐

Select an ontology to create a view on

Subscribe to email notifications for new notes ☐

<https://biodivportal.globio.org/ontologies>

How Requirements on a Terminology Service



NFDI 4
BIODIVERSITY

Offer efficient **semantic annotation** for text and tabular data



BiodivPortal

[Browse](#)

[Search](#)

[Mappings](#)

[Recommender](#)

[Annotator](#)

[Projects](#)

[Landscape](#)

[Login](#)

[Support](#)

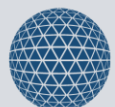
Annotator

Get annotations for biomedical text with classes from the ontologies

This dataset provides palaeoecological data for the glacial and Holocene sediment sequence retrieved from Lake Uddelermeer (The Netherlands) in 2012. Counts of pollen and chironomids are presented against both depth (cm) and age (cal yr. BP), and a loss-on ignition record is presented against depth. A total of 110 samples were analysed for their pollen content; a subset of 32 samples was analysed for sterol and stanol compounds. A total of 1412 samples were analysed for their organic content, approximated through loss-on-ignition. The data provides information on changes in the local and regional vegetation, changes in the within-lake ecosystem and changes in minerogenic input into the lake.

Get annotations

CLASS <small>filter</small>	ONTOLOGY <small>filter</small>	TYPE <small>filter</small>	UMLS SEM TYPE	CONTEXT	MATCHED CLASS <small>filter</small>	MATCHED ONTOLOGY <small>filter</small>
Dataset	Schema.org Vocabulary	direct		This dataset provides palaeoecological data ...	Dataset	Schema.org Vocabulary
glaciation	The Environment Ontology	direct		... for the glacial and Holocene sediment ...	glaciation	The Environment Ontology
sediment	The Environment Ontology	direct		... and Holocene sediment sequence retrieved from ...	sediment	The Environment Ontology
Sequence	ABCD Base Ontology	direct		... Holocene sediment sequence retrieved from Lake ...	Sequence	ABCD Base Ontology
saline evaporation pond	The Environment Ontology	direct		... retrieved from Lake Uddelermeer (The Netherlands) ...	saline evaporation pond	The Environment Ontology



InfAI®
Institut für Angew

How Requirements on a Terminology Service



NFDI 4
BIODIVERSITY

Automatically generate and store **mappings** between terminologies

BiodivPortal Browse Search **Mappings** Recommender Annotator Projects Landscape Login

Catalogue of Life Checklist
Last uploaded: April 23, 2024

Summary Classes Properties

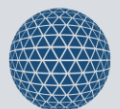
ONTOLOGY

- Integrated Taxonomic Information System
- The Environment Ontology

BiodivPortal is currently

← Previous 123456789...129130Next →

CATALOGUE OF LIFE CHECKLIST	INTEGRATED TAXONOMIC INFORMATION SYSTEM	SOURCE
Festuca rubra	Festuca rubra	LOOM
Tetragastris balsamifera	Tetragastris balsamifera	LOOM
Viola blanda	Viola blanda	LOOM
Malvaviscus arboreus	Malvaviscus arboreus	LOOM
Epilobium hirsutum	Epilobium hirsutum	LOOM
Alchornea latifolia	Alchornea latifolia	LOOM
Tanacetum bipinnatum	Tanacetum bipinnatum	LOOM
Araucaria araucana	Araucaria araucana	LOOM
Heteranthera dubia	Heteranthera dubia	LOOM
Erigeron bonariensis	Erigeron bonariensis	LOOM



How Requirements on a Terminology Service

Provide a single access point (API) to heterogenous terminological resources



```
{
  - links: {
    annotator: https://data.biodivportal.gfbio.org/annotator,
    categories: https://data.biodivportal.gfbio.org/categories,
    groups: https://data.biodivportal.gfbio.org/groups,
    documentation: https://data.biodivportal.gfbio.org/documentation,
    mappings: https://data.biodivportal.gfbio.org/mappings,
    metrics: https://data.biodivportal.gfbio.org/metrics,
    notes: https://data.biodivportal.gfbio.org/notes,
    ontologies: https://data.biodivportal.gfbio.org/ontologies,
    ontologies_full: https://data.biodivportal.gfbio.org/ontologies\_full,
    analytics: https://data.biodivportal.gfbio.org/analytics,
    submissions: https://data.biodivportal.gfbio.org/submissions,
    projects: https://data.biodivportal.gfbio.org/projects,
    property_search: https://data.biodivportal.gfbio.org/property\_search,
    provisional_classes: https://data.biodivportal.gfbio.org/provisional\_classes,
    provisional_relations: https://data.biodivportal.gfbio.org/provisional\_relations,
    recommender: https://data.biodivportal.gfbio.org/recommender,
    recommender_v1: https://data.biodivportal.gfbio.org/recommender\_v1,
    replies: https://data.biodivportal.gfbio.org/replies,
    reviews: https://data.biodivportal.gfbio.org/reviews,
    search: https://data.biodivportal.gfbio.org/search,
    slices: https://data.biodivportal.gfbio.org/slices,
    submission_metadata: https://data.biodivportal.gfbio.org/submission\_metadata,
    ontology_metadata: https://data.biodivportal.gfbio.org/ontology\_metadata,
    users: https://data.biodivportal.gfbio.org/users,
  }
}
```

<https://data.biodivportal.gfbio.org>

API integration in



Auswahl aus dem GFBio Terminology Service ITIS

Maximale Treffer: 10 Treffer

Suchbegriff eingeben: leucoma

Leucoma

Leucoma salicis

http://terminologies.gfbio.org/ITIS/Taxa_709248

Leucoma salicis

- Animalia
- Bilateria
- Protostomia
- Ecdysozoa
- Arthropoda
- Hexapoda
- Insecta
- Pterygota
- Neoptera

terminologi

tree

Aloe wollastor

terminologi

Leucomalt

Leucomalt

Megachile l

Corinnaetu

Philodromu

Rhytidophy


Spiroplasm

Pleurotome

How Ongoing Work

- ▶ Offer an environment for the **development**, curation, and publication of project terminologies
- ▶ Provide efficient terminology **versioning** and evolution mechanisms
- ▶ Offer **Community feedback** mechanisms

Contact:
karam@infai.org

 naouelkaram



www.nfdi4biodiversity.org

Approaches for an Open and FAIR Research Lifecycle

Ilaria Rosati, Mariantonietta La Marra, Alexandra Muresan, Martina Pulieri, Andrea Tarallo



UNIVERSITÀ
DEL SALENTO



Data management and capacity building workshop on Darwin Core standard – 7 June 2024

Open Science & FAIR Principles: what?

Data and other research outputs are available in the public domain or under copyright and licensed under an open licence that allows access, re-use, repurpose, adaptation and distribution under specific conditions.

UNESCO Recommendation on Open Science 2022 version 1.
<https://doi.org/10.54677/UTCD9302>

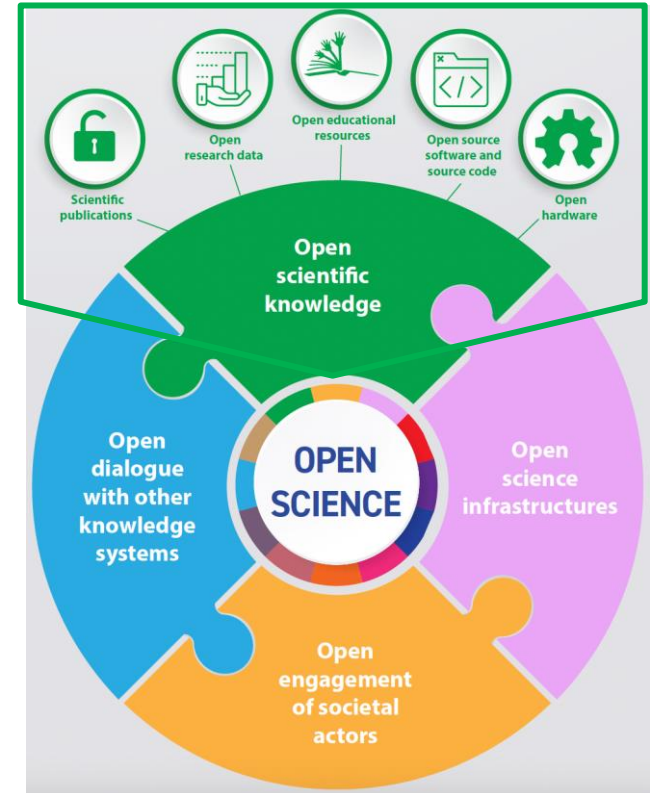


Image: UNESCO Open Science brochure

Open Science & FAIR Principles: what?

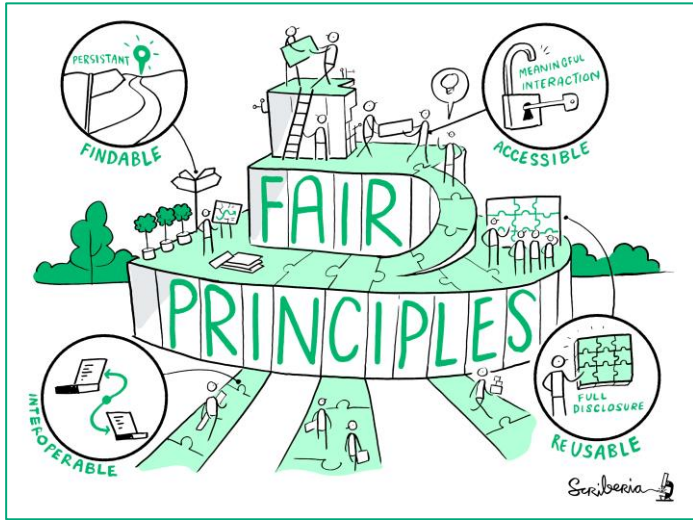


Image: Illustrations from the Turing Way book dashes. Zenodo.
<http://doi.org/10.5281/zenodo.3695300>

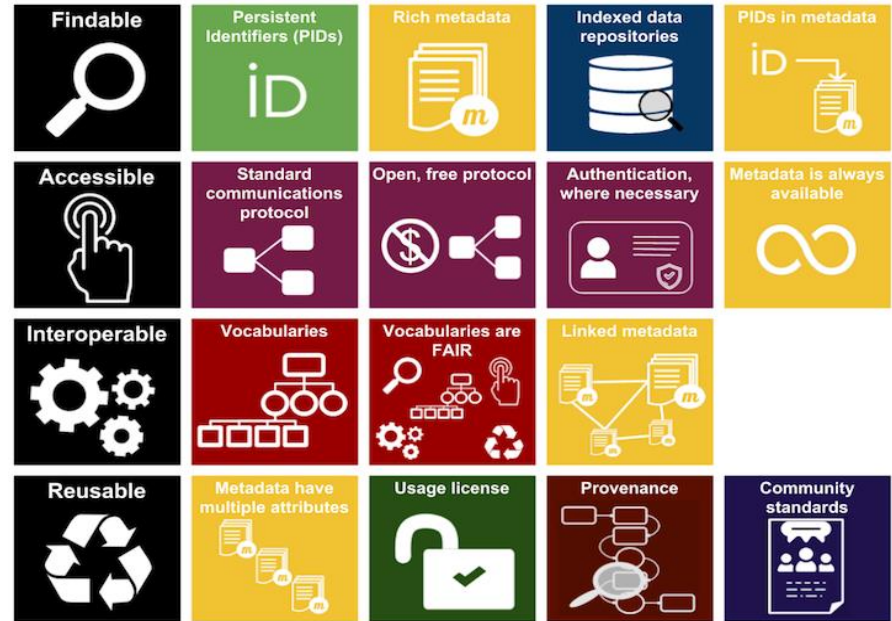


Image: Icons by [Freepik](https://www.flaticon.com) from www.flaticon.com and ARDC
<https://conference.eresearch.edu.au/fair-go-new-resources-to-support-fair-data/>

Open Science & FAIR Principles: what?

FAIR  **Open**

"FAIR is not the equivalent of open, but open needs to be FAIR to be useful"

Making your data/digital resources openly available does not translate to it being reusable!

To do so, we need clear and detailed information and data description.

Data can be FAIR but not Open! FAIR data maxim is: "as open as possible, as closed as necessary"

Ideally we want FAIR data/digital resources openly!

Open Science & FAIR Principles: why?



Bildquelle: University of Cape Town, Research Support Hub [CC-BY, Creative Commons Attribution 4.0 International License](#)

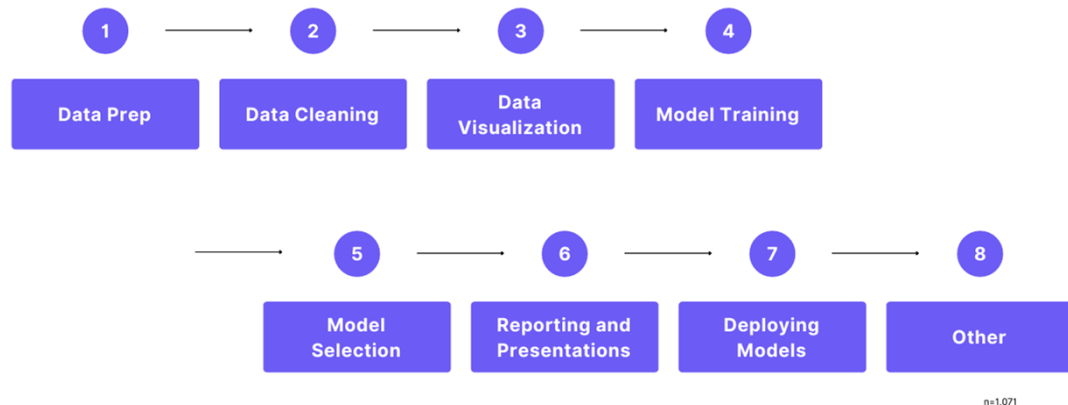
Open Science & FAIR Principles: who?



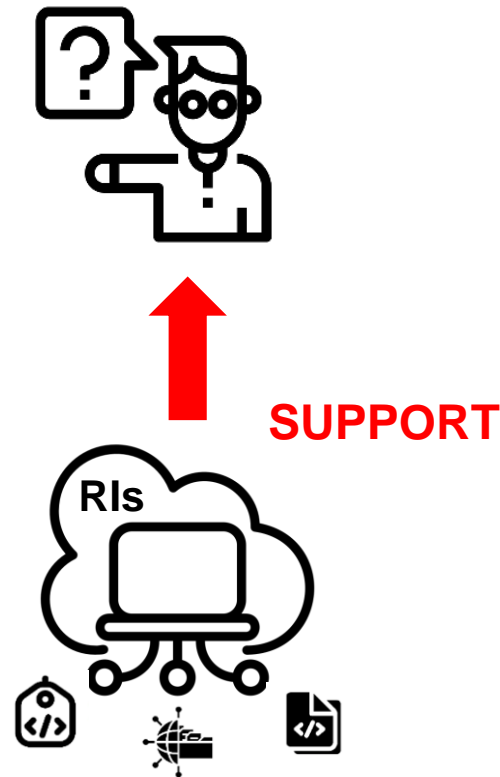
Gomes DGE et al. 2022 *Why don't we share data and code? Perceived barriers and benefits to public archiving practices*. Proc. R. Soc. B 289: 20221113. <https://doi.org/10.1098/rspb.2022.1113>

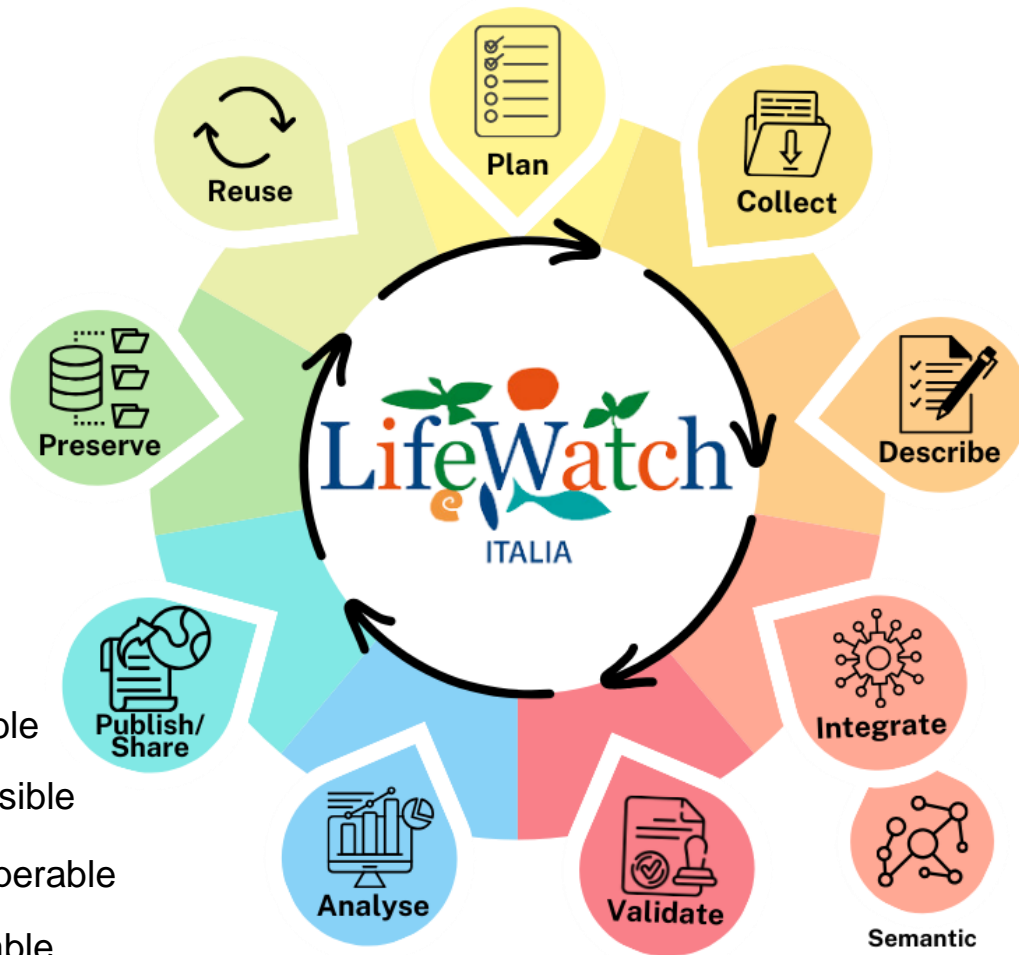
Open Science & FAIR Principles: who?

Thinking about your current role, what tasks are most time consuming?








Anaconda's "The State of Data Science 2023" report

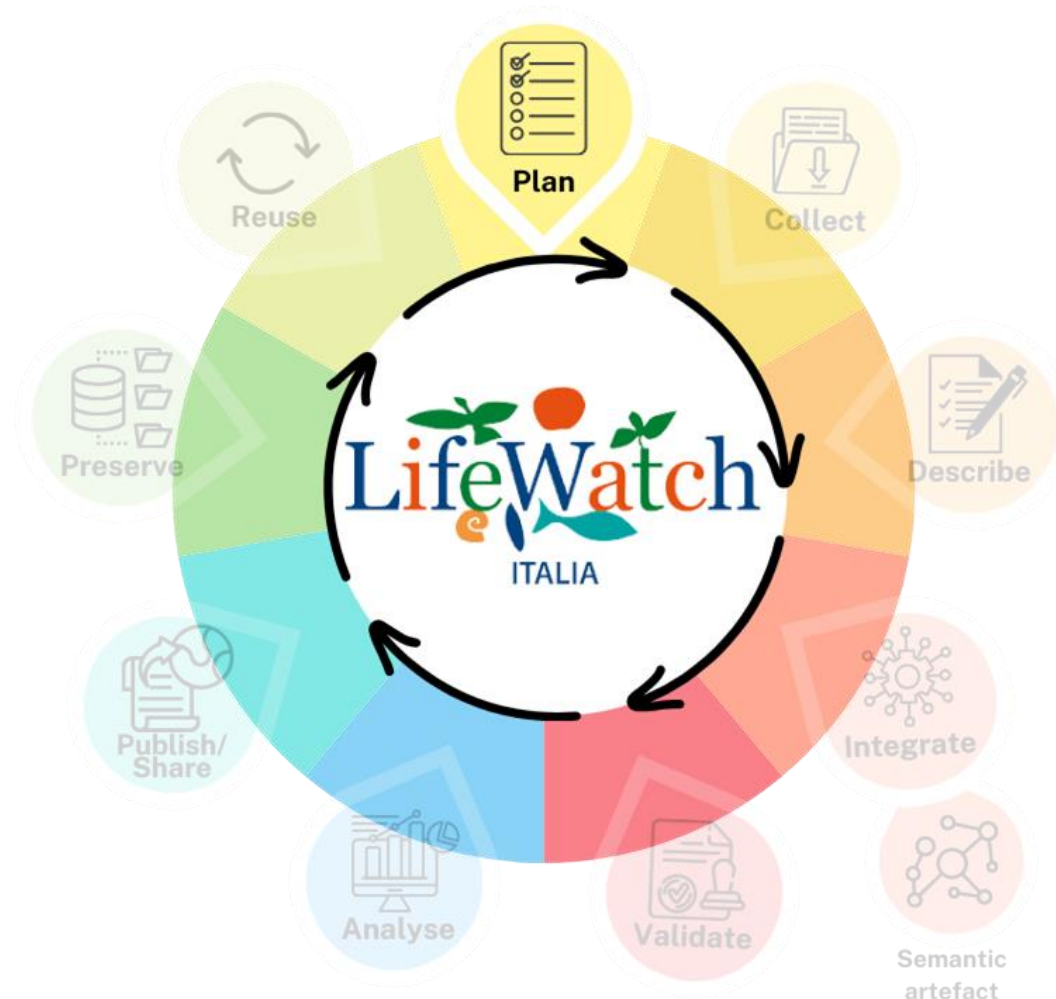
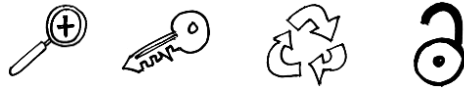




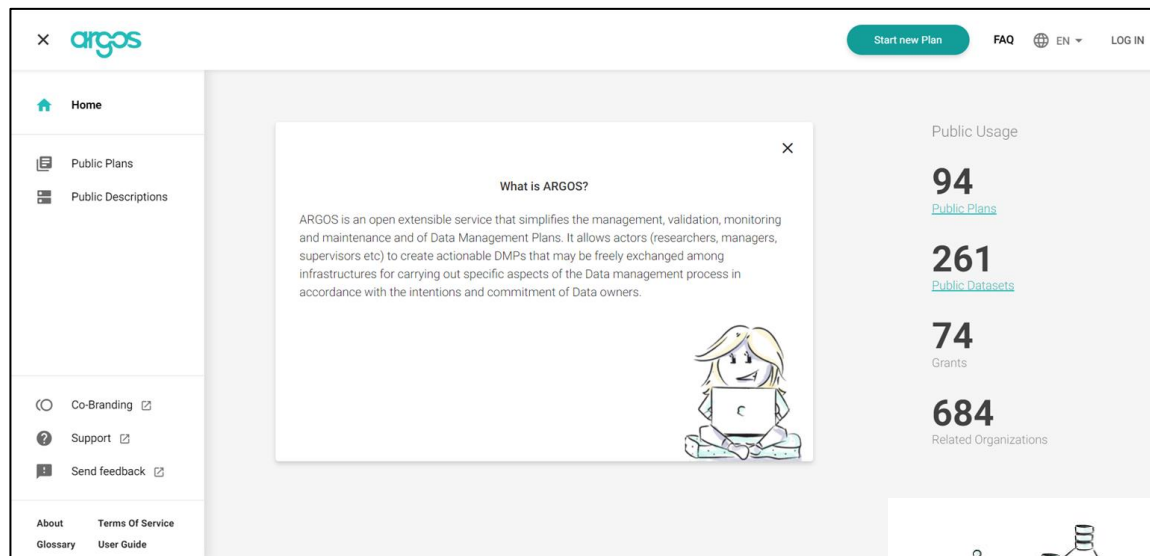
Open & FAIR Research Lifecycle

-  Findable
-  Accessible
-  Interoperable
-  Reusable
-  Open

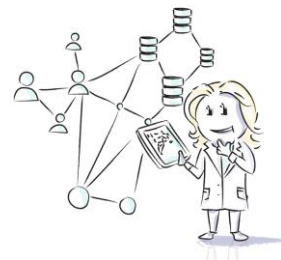
Data Management Plan



Plan the management of data and other research products



<https://argos.openaire.eu/home>



- Create machine actionable DMPs
- Configure to best fit your discipline
- Link to EOSC components out of the box
- Share easily in your repository

Data Portal



Datasets



Collect data



Biological invasions today represent one of the greatest threats to global biodiversity, with negative impacts on human health and the 'economy of our country. The USEit project addresses this issue by initiating a shared process among several marine and terrestrial institutes of the National Research Council (CNR) that aims to improve Italian research on invasive species and transform this environmental challenge into opportunities and new growth strategies.

Datasets


Now showing 1 - 1 of 1

LifeWatch Italy; 2022. [Individual and population-scale carbon and nitrogen isotopic signatures of *Procambarus clarkii* in invaded freshwater ecosystems](https://handle.stage.datacite.org/10.80186/dataportal/3)
<https://handle.stage.datacite.org/10.80186/dataportal/3>



Creation and customisation of sub-portals for gathering datasets produced of the same project or initiative.

Describe your data

 Drop files to attach them to the item, or [browse](#)

General information

Person organisation

Coverage

License

Project Detail

Methods

Ecological Metadata Language 2.2.0 as metadata schema for describing ecological data.

Every section contains a set of attributes, e.g. Title, Publication date, Temporal coverage, Taxonomic coverage, ect, to thoroughly describe your data.

Integration with semantic artefacts



EcoPortal
integration to
annotate
(meta)data and
to semantically
enrich them.

Hierarchical tree view

Data table

Attribute Name		Attribute Label	
<input type="text" value="shape"/>		<input type="text" value="Shape"/>	
The name of the attribute.		A label for displaying an attribute name.	
Attribute Definition		Storage Type	
<input type="text" value="The approximate 3 dimensional shape of a cell, described as one or more gec"/>		<input type="text"/>	
Precise definition of the attribute.		Storage type for data in a RDBMS or other data management system.	
Property Label	Property URI	Value Label	Value URI
<input type="text" value="contains measurements of type"/>	<input type="text" value="http://ecoinformatics.org/oboe/obo"/>	<input type="text" value="Shape"/>	<input type="text" value="https://kos.lifewatch.eu/thesauri/pl"/>
The persistent URI used to identify a property from a vocabulary.		The persistent URI used to identify a value from a vocabulary.	

Environmental Thesaurus

- variable
- biodiversity loss

Metadata validation

EML Schema

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <eml:eml xmlns:eml="https://eml.ecoinformatics.org/eml-2.2.0" xmlns:stmm1="http://www.xml-cml.org/schema/stmm1"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" packageId="8e6eb663-3d62-4453-aea5-35afd799bcad"
  system="https://data.lifewatchdev.eu" xsi:schemaLocation="https://eml.ecoinformatics.org/eml-2.2.0
  https://eml.ecoinformatics.org/eml-2.2.0/eml.xsd" scope="system">
3   <dataset>
4     <title>Biodiversity data of Italian marine ecosystems gathered for the alien species study case of LifeWat
5     <creator>
6       <individualName>
7         <givenName>Ilaria</givenName>
8         <surName>Rosati</surName>
9       </individualName>
```

Eml validation

Validation Successful

The xml file is checked against the EML schema validator (.xsd file)

Data validation: Taxonomic check

↑ Drop files to attach them to the item, or [browse](#)



**LW-ITA Data Schema based on
DarwinCore and controlled vocabularies**

scientificName	samplingProtocol	bodyLenght
<i>Idotea balthica</i>	hand operated ekman grab	4.5



scientificName
<i>Idotea balthica</i>



Taxonomic check



Validated data



**Italian Taxonomic
Backbone**



Data validation: Taxonomic check

Target Data

Select Domain

Select Environment

Select Taxonomy

Fauna

Brackish

Italian

Global

Back

Matching summary result

File name: e4cae767-c998-494c-a371-732bcd2f0bb2 - Dataset: Macrozoobenthos data collected in the Acquatina lagoon, Apulia, Italy

1

Select Actions

2

Confirm Actions

This matching resulted in **6%** compliance in Italian Backbone.
The remaining **94% (no match)** have this matching: 42% in CoL, 1% in WFO, 99% in WoRMS.

No Match

Match with warning

N. records: 13892/14735

N. records: 0/14735

0 records to delete

0 requests to send

Scientific name	CoL	WFO	WoRMS	Actions
Hydrobia - 778 records			✓	Keep the rows
Idotea balthica - 25 records			✓	Keep the rows
Lekanesphaera hookeri - 10 records			✓	Keep the rows
Lekanesphaera monodi - 142 records			✓	Keep the rows
Leptochelia savignyi - 108 records			✓	Keep the rows

Discover data

The screenshot shows a data discovery interface. On the left, a 'Filter by' panel lists various filters: Keyword, Creator, Organization, Year, Location, Taxon, Identifier, and Portal. The 'Portal' filter is currently selected. To the right of the filters, a search bar is present with the text 'Search phrase ...'. Below the search bar, a list of results is displayed, showing details for three different datasets. A red arrow points from the top left towards the search bar. Another red arrow points from the 'Portal' filter towards the map. A third red arrow points from the 'Keyword' filter towards the map. A fourth red arrow points from the 'Portal' filter towards the map. A fifth red arrow points from the 'Identifier' filter towards the map. A sixth red arrow points from the 'Portal' filter towards the map. A seventh red arrow points from the 'Identifier' filter towards the map. A eighth red arrow points from the 'Portal' filter towards the map. A ninth red arrow points from the 'Identifier' filter towards the map. A tenth red arrow points from the 'Portal' filter towards the map. A eleventh red arrow points from the 'Identifier' filter towards the map. A twelfth red arrow points from the 'Portal' filter towards the map. A thirteenth red arrow points from the 'Identifier' filter towards the map. A fourteenth red arrow points from the 'Portal' filter towards the map. A fifteenth red arrow points from the 'Identifier' filter towards the map. A sixteenth red arrow points from the 'Portal' filter towards the map. A seventeenth red arrow points from the 'Identifier' filter towards the map. An eighteenth red arrow points from the 'Portal' filter towards the map. A nineteenth red arrow points from the 'Identifier' filter towards the map. A twentieth red arrow points from the 'Portal' filter towards the map. A twenty-first red arrow points from the 'Identifier' filter towards the map. A twenty-second red arrow points from the 'Portal' filter towards the map. A twenty-third red arrow points from the 'Identifier' filter towards the map. A twenty-fourth red arrow points from the 'Portal' filter towards the map. A twenty-fifth red arrow points from the 'Identifier' filter towards the map. A twenty-sixth red arrow points from the 'Portal' filter towards the map. A twenty-seventh red arrow points from the 'Identifier' filter towards the map. A twenty-eighth red arrow points from the 'Portal' filter towards the map. A twenty-ninth red arrow points from the 'Identifier' filter towards the map. A thirtieth red arrow points from the 'Portal' filter towards the map. A thirty-first red arrow points from the 'Identifier' filter towards the map. A thirty-second red arrow points from the 'Portal' filter towards the map. A thirty-third red arrow points from the 'Identifier' filter towards the map. A thirty-fourth red arrow points from the 'Portal' filter towards the map. A thirty-fifth red arrow points from the 'Identifier' filter towards the map. A thirty-sixth red arrow points from the 'Portal' filter towards the map. A thirty-seventh red arrow points from the 'Identifier' filter towards the map. A thirty-eighth red arrow points from the 'Portal' filter towards the map. A thirty-ninth red arrow points from the 'Identifier' filter towards the map. A fortieth red arrow points from the 'Portal' filter towards the map. A forty-first red arrow points from the 'Identifier' filter towards the map. A forty-second red arrow points from the 'Portal' filter towards the map. A forty-third red arrow points from the 'Identifier' filter towards the map. A forty-fourth red arrow points from the 'Portal' filter towards the map. A forty-fifth red arrow points from the 'Identifier' filter towards the map. A forty-sixth red arrow points from the 'Portal' filter towards the map. A forty-seventh red arrow points from the 'Identifier' filter towards the map. A forty-eighth red arrow points from the 'Portal' filter towards the map. A forty-ninth red arrow points from the 'Identifier' filter towards the map. A fiftieth red arrow points from the 'Portal' filter towards the map. A fifty-first red arrow points from the 'Identifier' filter towards the map. A fifty-second red arrow points from the 'Portal' filter towards the map. A fifty-third red arrow points from the 'Identifier' filter towards the map. A fifty-fourth red arrow points from the 'Portal' filter towards the map. A fifty-fifth red arrow points from the 'Identifier' filter towards the map. A fifty-sixth red arrow points from the 'Portal' filter towards the map. A fifty-seventh red arrow points from the 'Identifier' filter towards the map. A fifty-eighth red arrow points from the 'Portal' filter towards the map. A fifty-ninth red arrow points from the 'Identifier' filter towards the map. A sixtieth red arrow points from the 'Portal' filter towards the map. A sixty-first red arrow points from the 'Identifier' filter towards the map. A sixty-second red arrow points from the 'Portal' filter towards the map. A sixty-third red arrow points from the 'Identifier' filter towards the map. A sixty-fourth red arrow points from the 'Portal' filter towards the map. A sixty-fifth red arrow points from the 'Identifier' filter towards the map. A sixty-sixth red arrow points from the 'Portal' filter towards the map. A sixty-seventh red arrow points from the 'Identifier' filter towards the map. A sixty-eighth red arrow points from the 'Portal' filter towards the map. A sixty-ninth red arrow points from the 'Identifier' filter towards the map. A seventieth red arrow points from the 'Portal' filter towards the map. A seventy-first red arrow points from the 'Identifier' filter towards the map. A seventy-second red arrow points from the 'Portal' filter towards the map. A seventy-third red arrow points from the 'Identifier' filter towards the map. A seventy-fourth red arrow points from the 'Portal' filter towards the map. A seventy-fifth red arrow points from the 'Identifier' filter towards the map. A seventy-sixth red arrow points from the 'Portal' filter towards the map. A seventy-seventh red arrow points from the 'Identifier' filter towards the map. A seventy-eighth red arrow points from the 'Portal' filter towards the map. A seventy-ninth red arrow points from the 'Identifier' filter towards the map. An eightieth red arrow points from the 'Portal' filter towards the map. An eighty-first red arrow points from the 'Identifier' filter towards the map. An eighty-second red arrow points from the 'Portal' filter towards the map. An eighty-third red arrow points from the 'Identifier' filter towards the map. An eighty-fourth red arrow points from the 'Portal' filter towards the map. An eighty-fifth red arrow points from the 'Identifier' filter towards the map. An eighty-sixth red arrow points from the 'Portal' filter towards the map. An eighty-seventh red arrow points from the 'Identifier' filter towards the map. An eighty-eighth red arrow points from the 'Portal' filter towards the map. An eighty-ninth red arrow points from the 'Identifier' filter towards the map. A ninetieth red arrow points from the 'Portal' filter towards the map. A ninety-first red arrow points from the 'Identifier' filter towards the map. A ninety-second red arrow points from the 'Portal' filter towards the map. A ninety-third red arrow points from the 'Identifier' filter towards the map. A ninety-fourth red arrow points from the 'Portal' filter towards the map. A ninety-fifth red arrow points from the 'Identifier' filter towards the map. A ninety-sixth red arrow points from the 'Portal' filter towards the map. A ninety-seventh red arrow points from the 'Identifier' filter towards the map. A ninety-eighth red arrow points from the 'Portal' filter towards the map. A ninety-ninth red arrow points from the 'Identifier' filter towards the map. A hundredth red arrow points from the 'Portal' filter towards the map.

Filter by

- Keyword +
- Creator +
- Organization +
- Year +
- Location +
- Taxon +
- Identifier +
- Portal +

Reset filters

Keyword

- ☐ Biodiversity 2
- ☐ Aquatic Habitats 1
- ☐ Biosphere 1
- ☐ Body size 1
- ☐ Coastal Habitat 1
- ☐ Coastal lagoons 1
- ☐ Intertidal 1
- Processes
- ☐ Intertidal Zone 1
- ☐ Invertebrates 1
- ☐ Lagoon 1

Search keyword

Name

Portal

- ☐ California MPA Monitoring Data Portal 4
- ☐ MCR LTER: Coral Reef: Rates of benthic coral reef community metabolism from 2007 ongoing 2
- ☐ Biological invasions USEit 1

Search portal

Name

Now showing 1 - 10 of 11

2023. Frugivoria: A global database for birds and mammals in contiguous Neotropical moist forests
<https://data.lifewatchitaly.eu/handle/1234567890>

Michela, Caggiano, LifeWatch Italy; 2023. Gas and meteorological measurements at the CTC site and Birch Hill in Fairbanks, Alaska, during the ALPACA-2022 field study
doi:10.18739/A27D2Q87W

LifeWatch Italy; 2022. Individual and population-scale carbon and nitrogen isotopic signatures of *Procambarus clarkii* in invaded freshwater ecosystems
<https://handle.stage.datacite.org/10.80186/dataportal/3>

Ecology Unit; 2020. Macrozoobenthos Cadres Project 2004-2005 LeafPack
<https://data.lifewatchitaly.eu/handle/123456789/1349>

Show all

Reuse data



DATA
Advanced Search
Semantic Search
Statistics
Portals
SUBMIT DATA
SUPPORT
ABOUT

Login

Phytoplankton communities data collected in Maldivian lagoon ecosystems in 2012

Home / Phytoplankton communities data collected in Maldivian lagoon ecosystems in 2012



Copy Citation Download Eml



Download

Name	File Type	Size	Download all	Data Table
Phytoplankton_Progetto_Strategico_2009_2012_Maldives.csv	CSV	8.06 MB		More info
Phytoplankton_communities_data_collected_in_Maldivian_lagoon_ecosystems_in_2012	JSON / JSON-linked-data	55 KB		More info



DataLabs



Scripts



Services



Datasets



Describe script



[HOMEPAGE](#)

[PROJECTS](#)

[SERVICES](#)

[HELP DESK](#)

[TRAINING](#)

[FAQS](#)

CREATE PROJECT



Title* ?

Phyto_TraitsComputation

Version* ?

1.0

Programming language* ?

r-4.3.0-cran

Programming language

python-3.10.0

r-4.3.0-cran

matlab-r2023a

Creator

Ilaria Rosati

Category* ?

data analysis



data processing

data curation

data validation

data annotation

data integration

data cleaning

data transformation

data visualization

Analyse data



HOMEPAGE
PROJECTS
SERVICES
HELP DESK
TRAINING
FAQS

```
if taxlev != 'community':  
    if len(dataset[taxlev].unique()) == 1:  
        taxlev = 'community'  
  
# In case the mandatory fields are not provided, the script provides an empty output  
if 'density' not in dataset.columns:  
    dataset['density'] = 1  
if 'biolume' not in dataset.columns:
```

Simple

0 0 0

Mode: Command index.ipynb

Input

Name

Size

Actions

Phytoplankton__Progetto_Strategico_2009_2012_United_Kingdom.csv

7.59 MB



Attach new file input

Select your file from your device or Data Portal

Local

Data Portal

Output

Warnings and Alerts

Team

Analyse data



[HOMEPAGE](#)
[PROJECTS](#)
[SERVICES](#)
[HELP DESK](#)
[TRAINING](#)
[FAQS](#)

File Edit View Run Kernel Git Tabs Settings Help CPU: 10% Mem: 155 / 3906 MB

index.ipynb

Python3.10.0

```
plt.plot(xx, den, marker='o')
if param == 'biovolume':
    plt.xlabel('average biovolume (µm³)')
elif param == 'cellcarboncontent':
    plt.xlabel('average cell carbon content (pg C)')
plt.ylabel('density (cell·L⁻¹)')
plt.title('Whole dataset')
plt.savefig(file_graph)

mat.index = ['density', 'average biovolume', 'average cell carbon content']
mat.to_csv('output/sizedensity_DATA.csv', sep=';', index=True, header=False, quoting=0, encoding='lat

plt.show()
```

Last executed at 2023-11-15 11:50:28 in 29ms

```
[ ]: calc_fun(datain, cluster, taxlev, param)
```

Last executed at 2023-11-15 11:50:34 in 1.18s

cluster: country, locality, year, month, day, parenteventid, eventid

cluster*scientificname

ty (cell·L⁻¹)

10²

16.1*M^{-0.17}
R²=0.06

Simple 0 1 Mode: Command index.ipynb

Validate & Publish metadata



LIFEWATCH ITALY
METADATA CATALOGUE



Scripts



Services

Categories ▾

Group ▾

Cancel

Save & close

Save metadata ▾

Request for publishing



Identification info

Title *

PhytoVRE_TraitsComputation

Date *

Publication ▾

2023/10/20



--:--:-- --



Abstract *

Mandatory field

B *I* U %

Programming Language *

Dropdown menu for Programming Language with options: R, Python, MatLab

Resource Constraints



Access constraints *

Overview

Choose or drop an image here

Download

Choose or drop a file here

Associated resources

+ Add ▾

Programming Language

Any other descriptive information about the dataset



Metadata Catalogue



Datasets



Research Site



Audio



Scripts



Services



Workflows



VREs



Describe

The screenshot displays the LifeWatch Italy Metadata Catalogue interface. The top navigation bar includes the LifeWatch logo, the site name, and links for Search, Metrics, FAIRness Assessment, and FAQs. The user profile 'Lucia Vaira ADMINISTRATOR' is in the top right. The 'Admin console' dropdown menu is open, showing options like Editor board, Add new record, Import new records, Manage directory, Batch editing, and Access rights. The 'Harvesting' option is highlighted with a red box and labeled 'API'. The main content area features a search bar, a resource count of 177, and a 'Browse resources' section with icons for Scripts (22), Datasets (70), Audio (9), VREs (24), Workflows (18), Services (24), and Research Sites (10).

LifeWatch Italy Metadata Catalogue

Search Metrics FAIRness Assessment FAQs

Contribute Admin console

Lucia Vaira ADMINISTRATOR

Search ...

177 resources for biodiversity and ecosystem re

Browse resources

Scripts 22

Datasets 70

Audio 9

VREs 24

Workflows 18

Services 24

Research Sites 10

Admin console dropdown menu:

- Editor board
- Summary
- Metadata and templates
- Users and groups
- Harvesting API
- Statistics and status
- Reports
- Classification systems
- Settings
- Tools
- Filters configurator
- Manage resource

Describe



LifeWatch Italy Metadata Catalogue

Search ▾

Metrics

FAIRness Assessment

FAQs

Contribute ▾

Admin console ▾

Add a new record

Create a **Script**



Scripts



Datasets



Audio



VREs

From ...

LIFEWATCH ERIC METADATA PROFILE FOR
SCRIPTS (CUSTOMISED ISO 19139)

In group ...

LifeWatch Italy ▾

+ Create ▾

✕ Cancel

Validate

The screenshot shows a web form with a top navigation bar containing buttons for 'Categories', 'Group', 'Cancel', 'Save & close', 'Save metadata', 'Request for publishing', and an eye icon. The main form area has several sections:

- A text input field with a rich text editor toolbar (B, I, U, link icon).
- A 'Programming Language' dropdown menu with a red asterisk indicating it is mandatory. A dropdown menu is open, showing a list of options: Copyright, Intellectual property rights, License, Other restrictions, Patent, Pending patent, Restricted, and Trademark. The 'Copyright' option is highlighted with a blue bar and a checkmark.
- A 'Mandatory field' label in red text.
- A 'Use limitation' dropdown menu with a red asterisk and a 'Mandatory field' label in red text.
- A text input field with a rich text editor toolbar (B, I, U, link icon).
- An 'Other constraints' section with a text input field and a rich text editor toolbar (B, I, U, link icon).

On the right side, there is a sidebar with three sections:

- 'Overview' with a green button labeled 'Choose or drop an image here'.
- 'Download' with a green button labeled 'Choose or drop a file here'.
- 'Associated resources' with a text input field and a '+ Add' button.
- 'Suggestions' with a gear icon.

At the bottom right, there are two green circular buttons: one with an upward arrow and one with a hamburger menu icon.

Mandatory fields

Recommended values

Controlled vocabularies

Publish

Categories

Group

Cancel

Save & close

Save metadata

Request for publishing

DefaultProjectMethodsData Tables

Identification

Alternate Identifier

cc6e6d1e-03c7-40ee-8b63-bf1845e699d2

Publication Date *

2024-05-29

Title *

Phytoplankton communities data collected in United Kingd

Short Name

Abstract *

B I U ↺

This dataset provide a broad perspective and information on taxonomic and morphological (shape and size) structure of phytoplankton communities in Scottish lagoon ecosystem.

Dataset Language

en

Dataset Creator

Organization Name *

Request for publishing

Please, send the request only if you are sure that your metadata record includes all mandatory and recommended fields.

Remember that you can use the FAIRness report to verify that your metadata record follows the FAIRness criteria. Use the button "Save & close" to see the detailed view of your metadata record and click on the button "FAIRness Assessment".

If your metadata record does not have a DOI, you can ask for it right now by selecting the following checkbox. The DOIs are assigned only upon verification and validation of the metadata record.

☐ I want to get a DOI

Cancel

Send

Discover & Reuse



LifeWatch Italy Metadata Catalogue

Search

Metrics

FAIRness Assessment

FAQs

Contribute

biodiversity



Active filters



any biodiversity



6 Results

Sorted by relevancy



Filter

Expand Collapse

TYPE OF RESOURCES

- ☐ Workflow (3)
- ☐ Service (2)
- ☐ Dataset (1)

CONTACT FOR THE RESOURCE

PUBLICATION YEAR

- ☐ 2021 (3)
- ☐ 2018 (1)
- ☐ 2015 (1)

STATUS

GROUPS

NEW FILTER_WORKFLOW



Biodiversity data of Italian transitional waters gathered for the alien species study case of LifeWatch Italy.

The dataset contains more than 3800 occurrence data divided among 9 Eunis species groups (Algae, Amphibian, Birds, Cyanobacteria, Fishes, Flowering Plants, Invertebrates, Protists, Reptiles). Data refer to alien species and native species distributed within 23 sites and 4 Eunis habitats (level 2) all around Italy.

2 links

Category: Datasets



ICTS Doñana monitoring data

It shows information collected on the state of biodiversity and natural processes in the different areas collected according ICTS Monitoring Program.

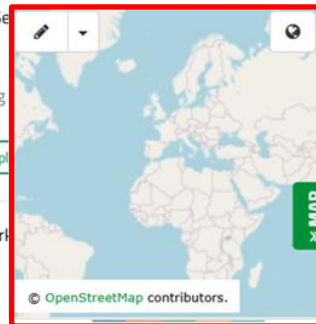
1 link

Category: Services



Metabarcoding Workflow - Detection of Non-native Invasive Species with metabarcoding

Category: Workflows



MAP

OpenStreetMap contributors.

Discover & Reuse



LifeWatch Italy Metadata Catalogue

Search

Metrics

FAIRness Assessment

FAQs

Contribute

Admin console

Lucia Vaira
ADMINISTRATOR

Back to search

Edit

Delete

Manage record

Download

Copy citation

FAIRness Assessment

Display mode

Back



Macrozoobenthos data collected in the Acquatina lagoon, Apulia, Italy.

The dataset contains abundance and body size data of benthic macroinvertebrate community of the Acquatina lagoon (Apulia, Italy). Data refer to 50 taxa and 14735 individuals.

- Fairness assessment resource
- Fairness assessment catalogue



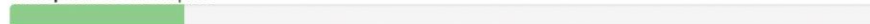
Findable: 69% complete



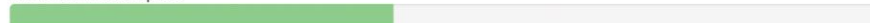
Accessible: 83% complete



Interoperable: 20% complete



Reusable: 44% complete



Passed 31 checks out of 65.

Warning for 30 checks. Please review these warning.

Failed 34 checks. Please correct these issues.

“To make progress in science, we need to be open and share.”

THANK YOU!

Ilaria Rosati

ilaria.rosati@cnr.it



UNIVERSITÀ
DEL SALENTO





French BIF - Introduction to Darwin Core

Guillaume BODY | Sophie PAMERLON



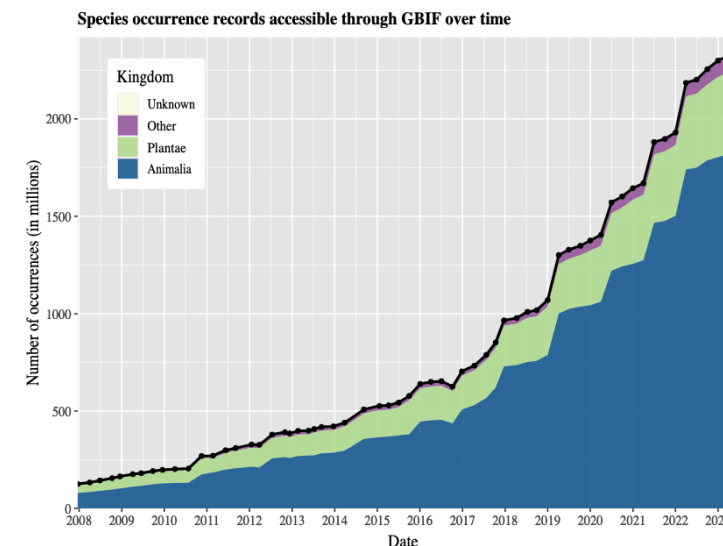
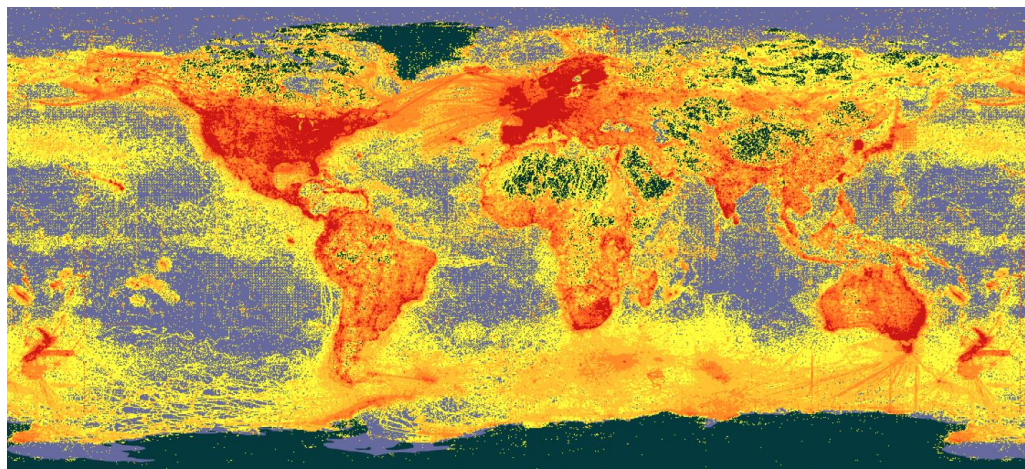
Global Biodiversity Information Facility (GBIF)

- Intergovernmental programme & data infrastructure, created in 2001 by the OECD scientific committee
- Objectives: promote & facilitate the **free and open access to biodiversity data**

More than **2,9 mds data**, cited in **10 229** scientific publications | 173,8 mds data downloaded per month

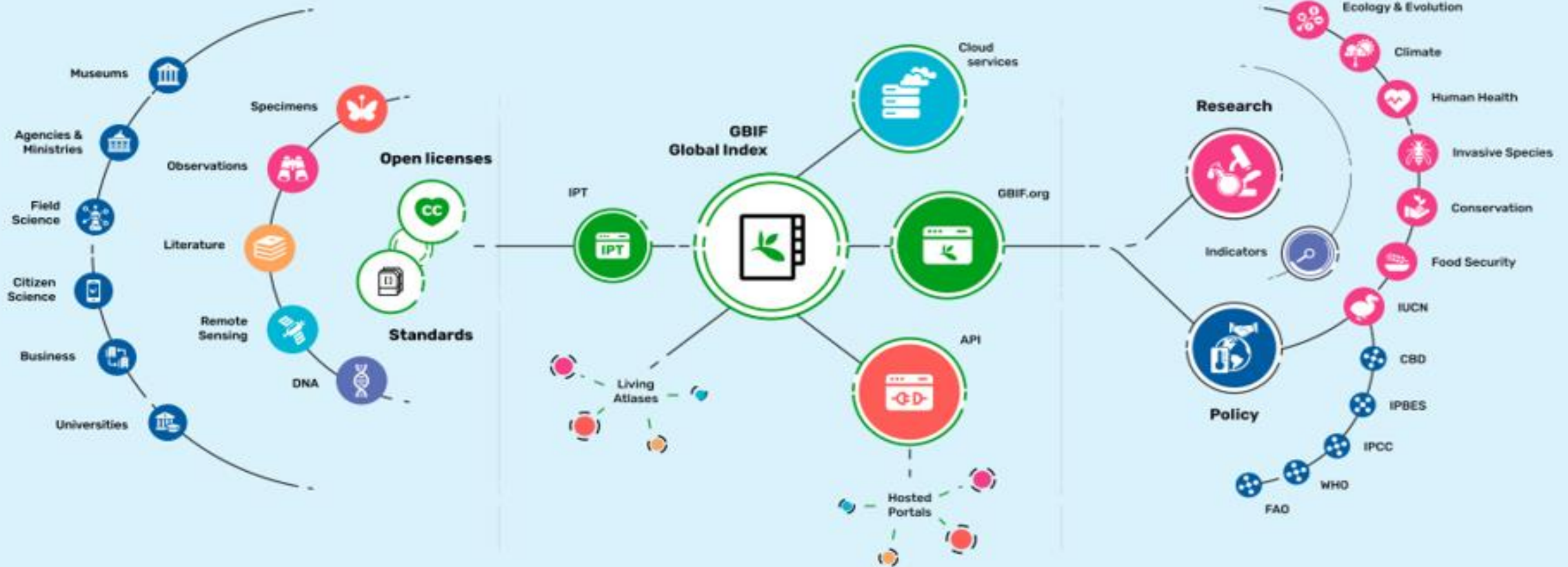
→ A window on where, when and by whom species were recorded

- Collaboration through an agreement protocol
- **105 members** (62 countries, 43 associate organisations)
- **142 contributing countries**
- **2170 data publishers**
- Secretariat based in Copenhagen, Denmark





PROVIDING BIODIVERSITY EVIDENCE FOR RESEARCH AND POLICY





STANDARDS: LET'S AGREE TO AGREE

”

Standardisation does not mean that we all wear the same color and weave of cloth, eat standard sandwiches, or live in standard rooms with standard furnishing. Homes of infinite variety of design are built with a few types of bricks, and with lumber of standard sizes, and with water and heating pipes and fitting of standard dimensions.

W. Edwards Deming



WHAT IS A STANDARD?

An agreed way of doing something

Convention
Restriction
Rule Requirement
Norm
Specification

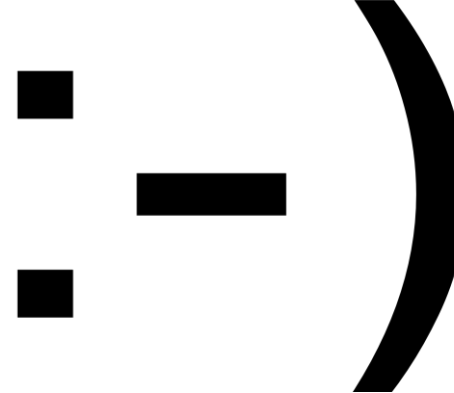
EVERYDAY STANDARDS



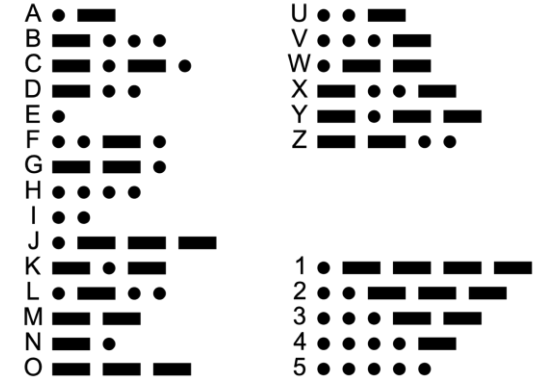
Units of Measurement
(Metric, Imperial)



Alphabets



Emojis



Morse Code



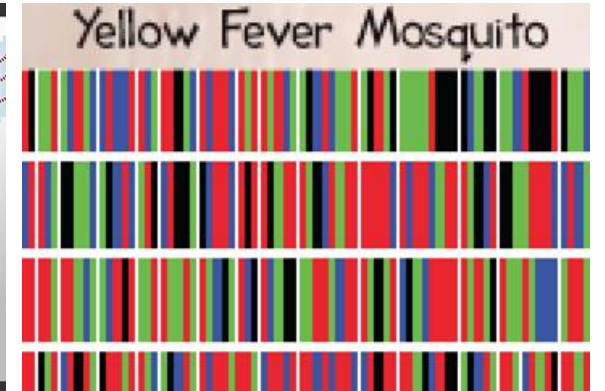
Numeral Systems
(Hindu-Arabic; Roman Numerals)



Languages



Postal addressing



Barcoding

RULES AND RESTRICTIONS

- Type of data
restrictions on the category of the field
- Encoding schema
restrictions on the range of values in the field
- Format
restrictions on the representation of the data
- Character encoding
rules for interpreting bytes

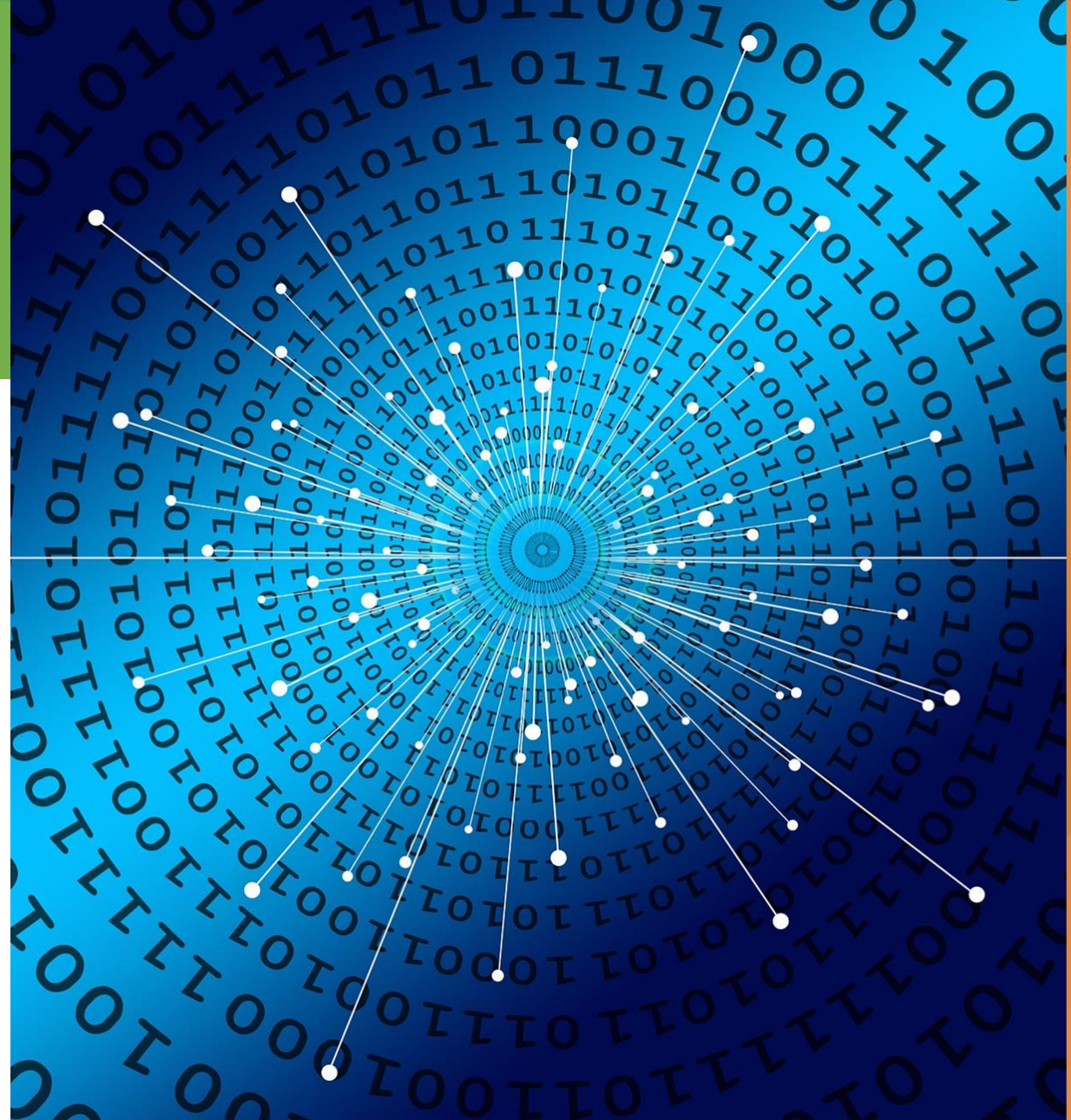
Image by Bird Explorers



STANDARDS FOR DATA TRANSFER

- Application schema
Specific combinations of data standards for a particular purpose
- Format
Restrictions in the dataset structure
- Transfer protocol
Where and how to send content

Image by Gerd Altmann





BIODIVERSITY INFORMATION STANDARDS

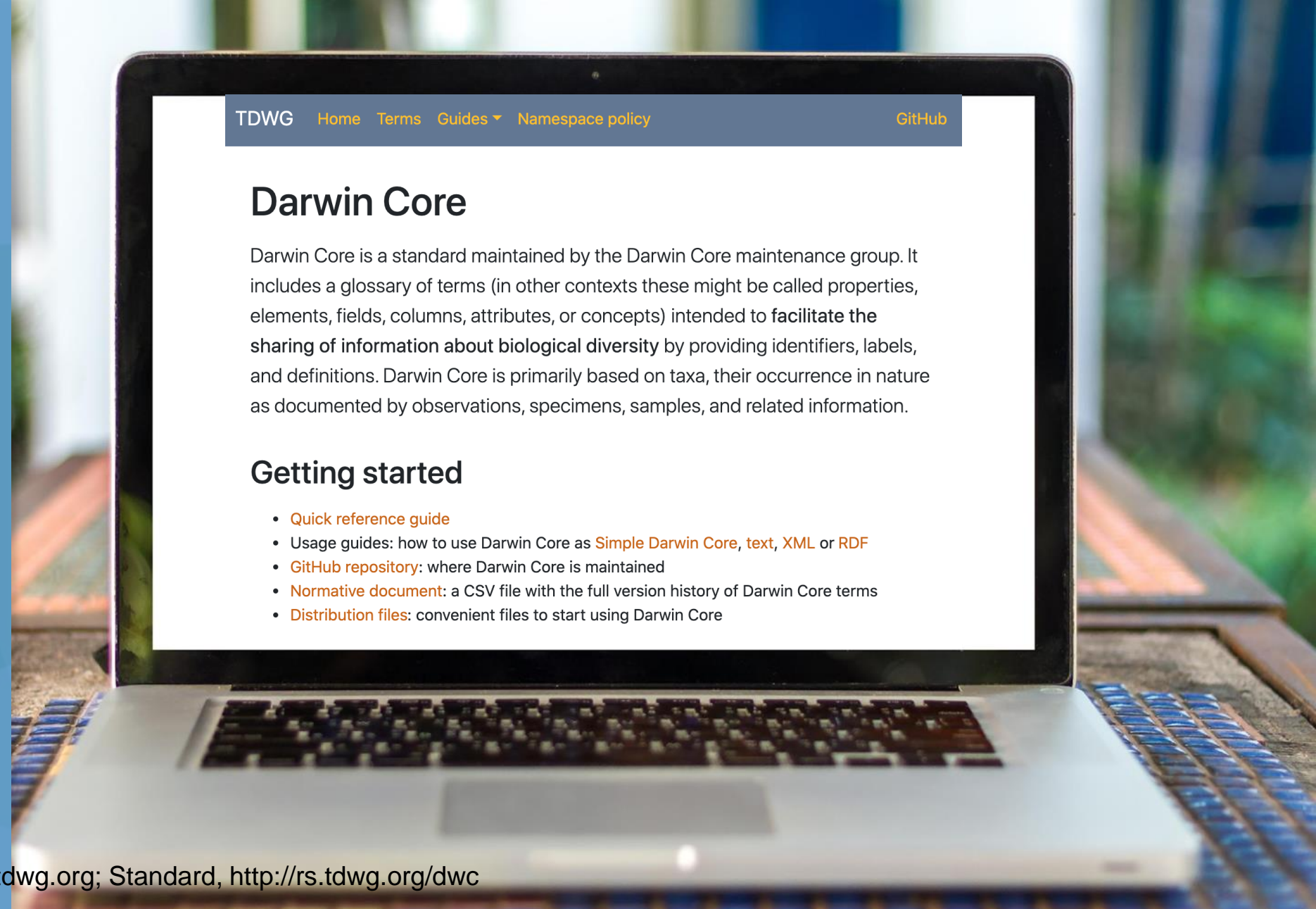
“Data standards are the rules by which data are described and recorded. In order to share, exchange, and understand data, we must standardize the format as well as the meaning.” (USGS)

Ecological Metadata Language Standard (EML)
Audubon Media Description (aka Audubon Core)
Global Genome Biodiversity Network (GGBN)
Ocean Data Standards and Best Practices Project (ODSBP)
Darwin Core

WHAT IS DARWIN CORE?

"List of fields and their definitions,
as they relate to biodiversity
data."

Reference(s): Governance, <http://www.tdwg.org>; Standard, <http://rs.tdwg.org/dwc>



SIMPLE DARWIN CORE

Field classes

- Record & Dataset
- Occurrence
- Organism
- Material Sample
- Event
- Location
- Geological Context
- Identification
- Taxon

Auxiliary classes:

- ResourceRelationship
- MeasurementOrFact

<https://dwc.tdwg.org/simple/>
Image by Vladimir Tkalčić



DWC QUICK REFERENCE GUIDE

<https://dwc.tdwg.org/terms/>

Darwin Core quick reference guide

This page provides a list of all currently recommended terms of the Darwin Core standard. Categories such as **Occurrence** or **Event** correspond to Darwin Core classes which group other terms. Convenient **files of these terms** and **their full history** can be found in the **Darwin Core repository**.

Record-level

type	modified	language	license	rightsHolder	accessRights
bibliographicCitation	references	institutionID	collectionID		
datasetID	institutionCode	collectionCode	datasetName		
ownerInstitutionCode	basisOfRecord	informationWithheld			
dataGeneralizations	dynamicProperties				

type

Property

Record-level

Occurrence

Organism

MaterialSample

Event

Location

GeologicalContext

Identification

Taxon

MeasurementOrFact

ResourceRelationship

UseWithIRI

LivingSpecimen

PreservedSpecimen

FossilSpecimen

HumanObservation

MachineObservation



DWC TERMS: BASISOFRECORD

basisOfRecord		Property
Identifier	http://rs.tdwg.org/dwc/terms/basisOfRecord	
Definition	The specific nature of the data record.	
Comments	Recommended best practice is to use the standard label of one of the Darwin Core classes.	
Examples	PreservedSpecimen, FossilSpecimen, LivingSpecimen, MaterialSample, Event, HumanObservation, MachineObservation, Taxon, Occurrence	

Record-level

Occurrence

Organism

MaterialSample

Event

Location

<https://dwc.tdwg.org/terms/#dwc:basisOfRecord>



DWC TERMS: OCCURRENCEID

occurrenceID		Property
Identifier	http://rs.tdwg.org/dwc/terms/occurrenceID	
Definition	An identifier for the Occurrence (as opposed to a particular digital record of the occurrence). In the absence of a persistent global unique identifier, construct one from a combination of identifiers in the record that will most closely make the occurrenceID globally unique.	
Comments	Recommended best practice is to use a persistent, globally unique identifier.	
Examples	http://arctos.database.museum/guid/MSB:Mamm:233627 , 000866d2-c177-4648-a200-ead4007051b9 , urn:catalog:UWBM:Bird:89776	

Record-level

Occurrence

Organism

MaterialSample

Event

Location

GeologicalContext

Identification

<https://dwc.tdwg.org/terms/#dwc:occurrenceID>



DWC TERMS: COUNTRY AND COUNTRYCODE

country Property	
Identifier	http://rs.tdwg.org/dwc/terms/country
Definition	The name of the country or major administrative unit in which the Location occurs.
Comments	Recommended best practice is to use a controlled vocabulary such as the Getty Thesaurus of Geographic Names.
Examples	Denmark , Colombia , España

countryCode Property	
Identifier	http://rs.tdwg.org/dwc/terms/countryCode
Definition	The standard code for the country in which the Location occurs.
Comments	Recommended best practice is to use an ISO 3166-1-alpha-2 country code.
Examples	AR , SV

MaterialSample

Event

Location

GeologicalContext

Identification

Taxon

MeasurementOrFact

ResourceRelationship

UseWithIRI

LivingSpecimen

PreservedSpecimen

<https://dwc.tdwg.org/terms/#dwc:country>; <https://dwc.tdwg.org/terms/#dwc:countryCode>

DARWIN CORE EXTENSIONS

- Audubon Media Description (aka Audubon Core)
- Measurements or Facts
- Identification History
- And many more!

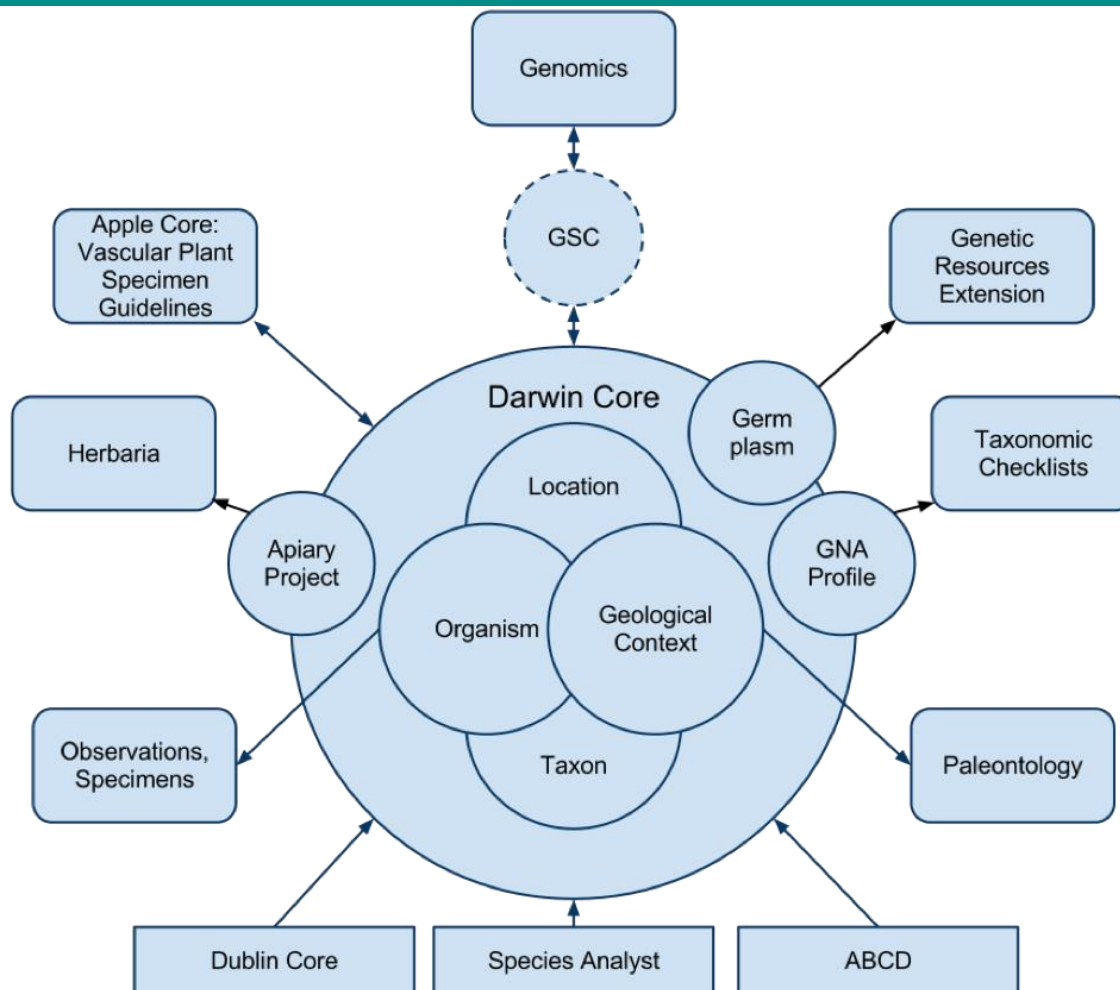
<https://tools.gbif.org/dwca-validator/extensions.do>

Image by Elizabeth Byers

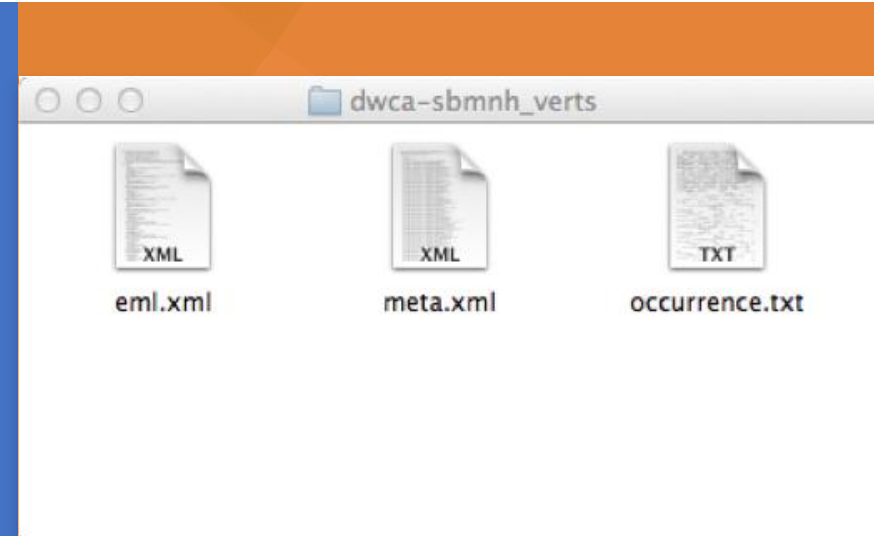
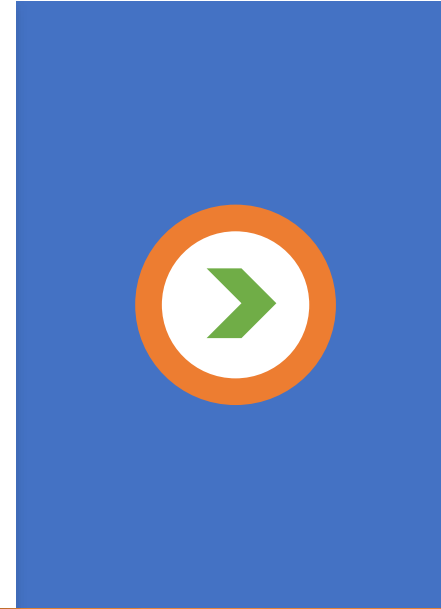
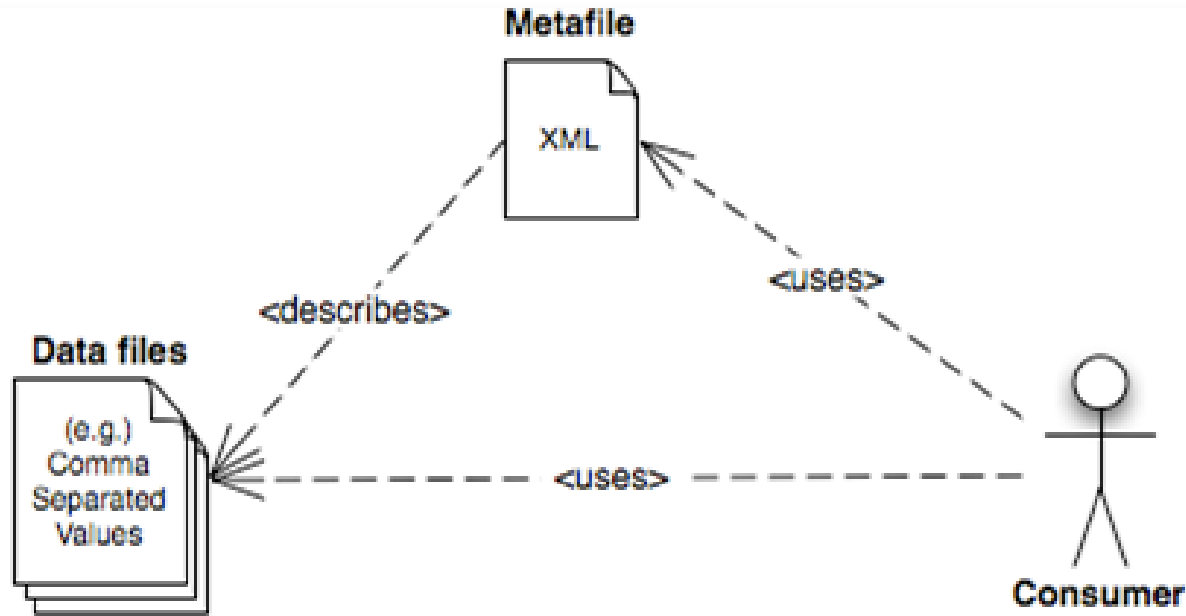




COMMUNITY AND STANDARDS RELATIONSHIPS



DARWIN CORE ARCHIVES



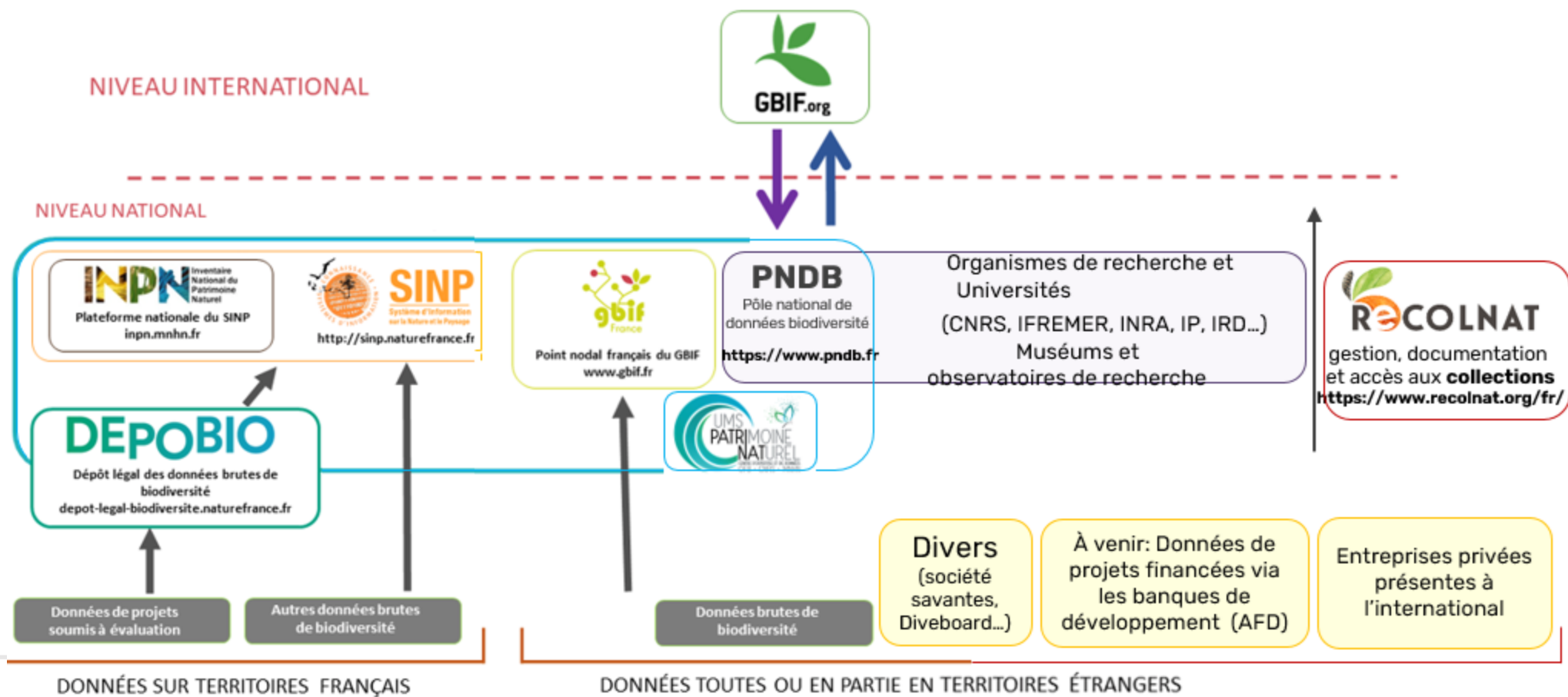
A DwC Archive is an expression of the Darwin Core text guide. It is a compressed file containing a minimum of three files. It is encoded as UTF-8.

<https://dwc.tdwg.org/text/>



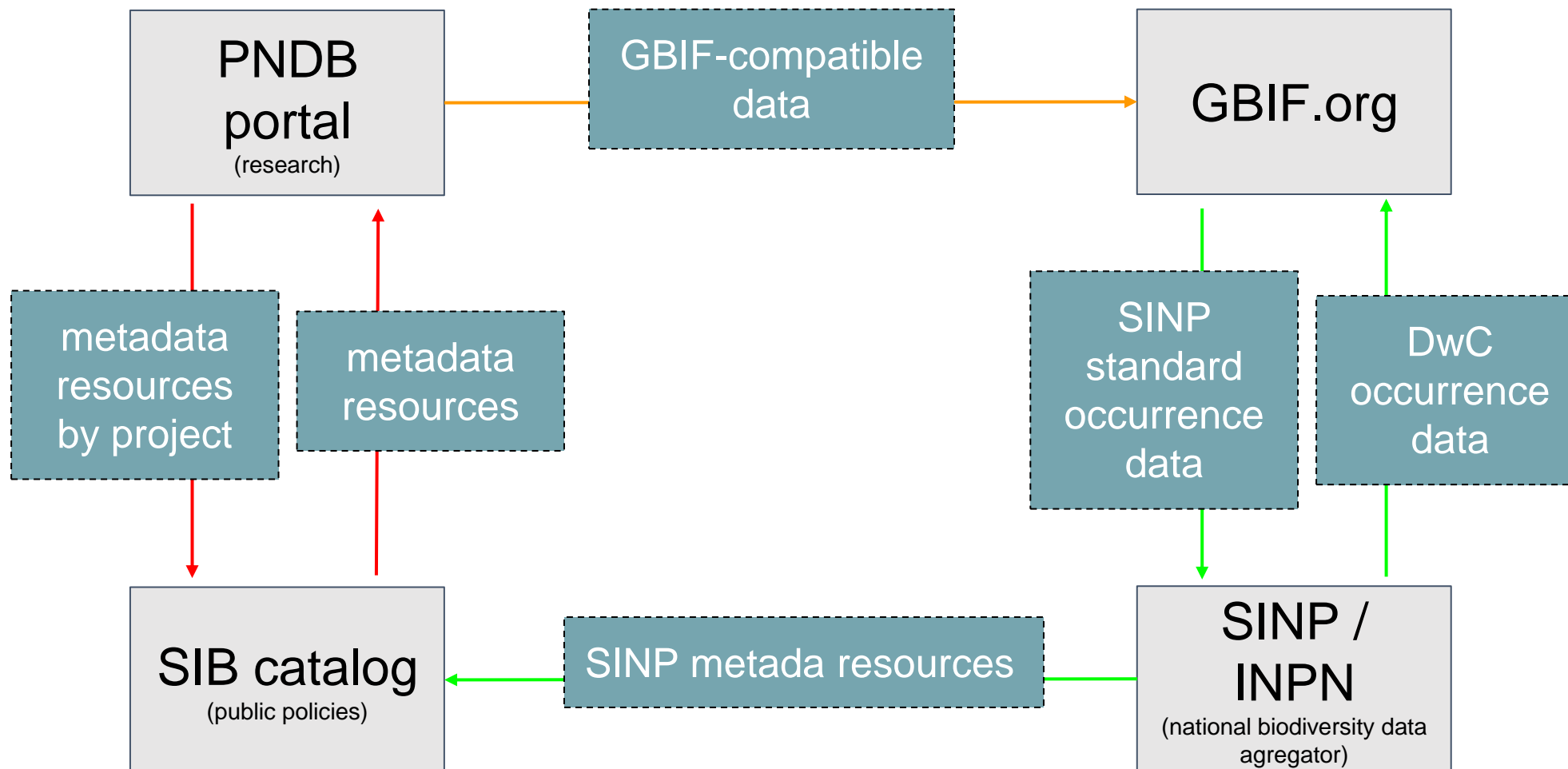
FRENCH NATIONAL DATA WORKLOWS

DIFFUSION DES DONNÉES BRUTES DE BIODIVERSITÉ





FRENCH NATIONAL DATA WORKFLOWS





Darwin core and monitoring data

Electrofishing protocol





Darwin core and monitoring data

Electrofishing monitoring data

Code Sandre de la station	Nom de la station	Code ASPE du point de prélèvement	Code Sandre du point de prélèvement	Numéro d'opération	Date de l'opération	Niveau de validation	Qualification de la donnée	Bassin hydrographique	Cours d'eau				
4613007	R DEME A LOUESTAULT		6	90170	12/10/23	Validé niveau 1	Correcte	LOIRE-BRETAGNE	la Dême				
4613007	R DEME A LOUESTAULT		6	90170	12/10/23	Validé niveau 1	Correcte	LOIRE-BRETAGNE	la Dême				
4613007	R DEME A LOUESTAULT		6	90170	12/10/23	Validé niveau 1	Correcte	LOIRE-BRETAGNE	la Dême				
Code de l'entité hydrographique	Région	Département	Commune	Lieu-dit	Coordonnées X du point de prélèvement	Coordonnées Y du point de prélèvement	Système de projection des coordonnées du point	Objectif de pêche	Espèce ciblée				
M1345800	CENTRE-VAL-DE-LOIRE	Indre-et-Loire	BEAUMONT-LA-RONCE	PONT DE LA D29 - LIEU-DIT MOQUE-SOURIS	524028.00000000	6728114.00000000	RGF93 / Lambert 93	RRP – Réseau de Référence Pérenne					
M1345800	CENTRE-VAL-DE-LOIRE	Indre-et-Loire	BEAUMONT-LA-RONCE	PONT DE LA D29 - LIEU-DIT MOQUE-SOURIS	524028.00000000	6728114.00000000	RGF93 / Lambert 93	RRP – Réseau de Référence Pérenne					
M1345800	CENTRE-VAL-DE-LOIRE	Indre-et-Loire	BEAUMONT-LA-RONCE	PONT DE LA D29 - LIEU-DIT MOQUE-SOURIS	524028.00000000	6728114.00000000	RGF93 / Lambert 93	RRP – Réseau de Référence Pérenne					
Protocole de pêche	Numéro du Passage	Type de point	Nom de l'ambiance	Surface échantillonnée (m²)	Identifiant du lot	Numéro du lot	Type du lot	Effectif de la ligne	Effectif du lot				
Pêche complète à un ou plusieurs passages	1			351.00	5648513	1	N	1	1				
Pêche complète à un ou plusieurs passages	1			351.00	5648575	63	I	1	3				
Pêche complète à un ou plusieurs passages	1			351.00	5648569	57	S/L	1	158				
Poids du lot (g)	Taille maximale des individus du lot (mm)	Taille minimale des individus du lot (mm)	Type de longueur du lot	Identifiant du poisson	Taxon - Code alternatif	Taxon - Code taxref	Taxon - Nom commun	Taille de l'individu (mm)	Mesure réelle de la taille (non dégroupée)	Poids du poisson (g)	Présence de pathologie	Sexe de l'individu	Opération confidentielle
36			Totale	22483116	TRF	67778	Truite de rivière	162	Oui		Non		Non
34			Totale	22483352	LOF	67550	Loche franche	56	Oui		Non		Non
864			Totale	22483172	VAI	67404	Vairon	79	Oui		Non		Non

40 columns

Figure 5: Sample of electrofishing raw data extracted from ASPE database (data has been subdivided for easier reading)



Darwin core and monitoring data

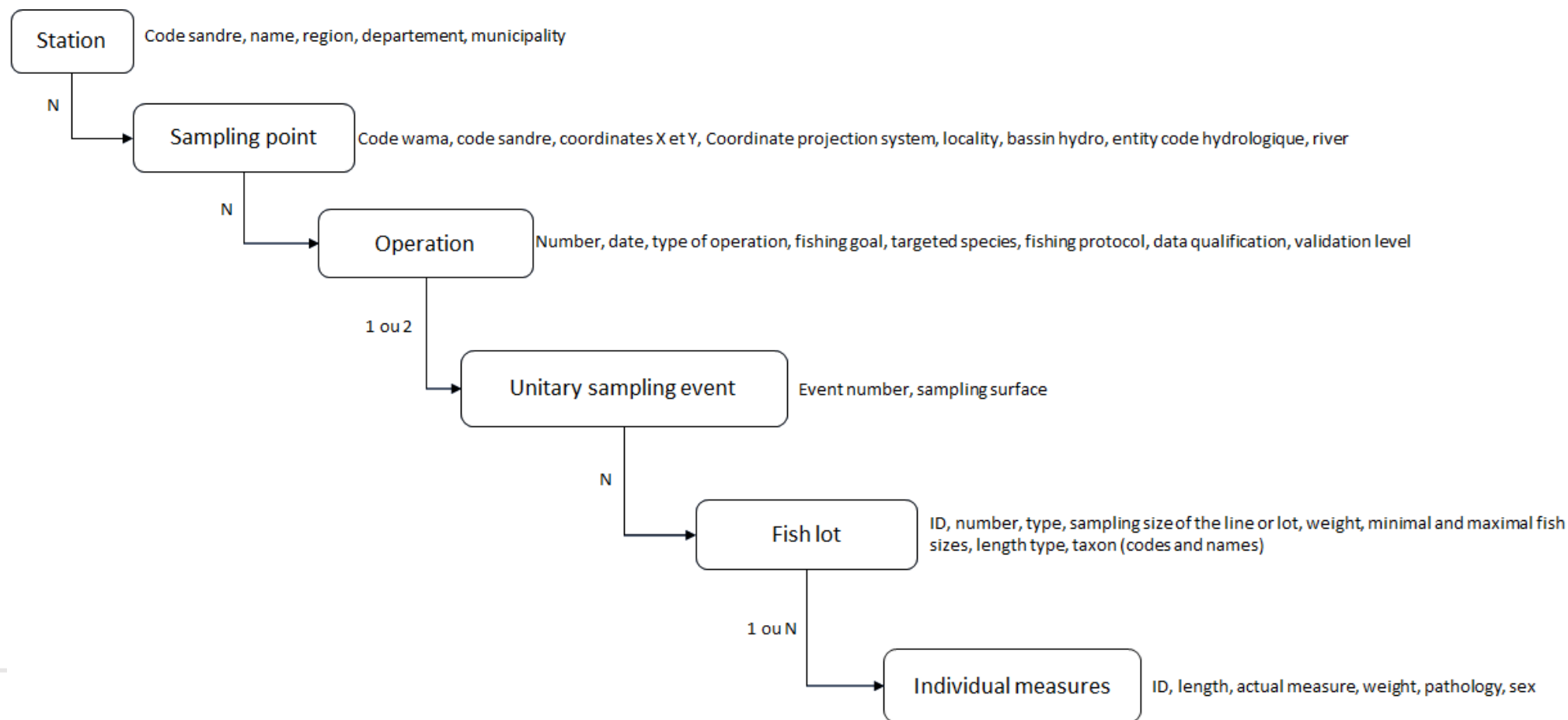
Electrofishing protocol





Darwin core and monitoring data

Electrofishing protocol





Event core:
for sampling event description

eventID	term2	term3
value	value	value

Occurrence extension:
for observations

occurrenceID	term2	term3
value	value	value

Humboldt extension:
for protocol details

eventID	term2	term3
value	value	value

Measurement of fact extension:
for measures or other facts linked to an event

measurementID	measurementType	measurementValue	measurementUnit
value	value	value	value

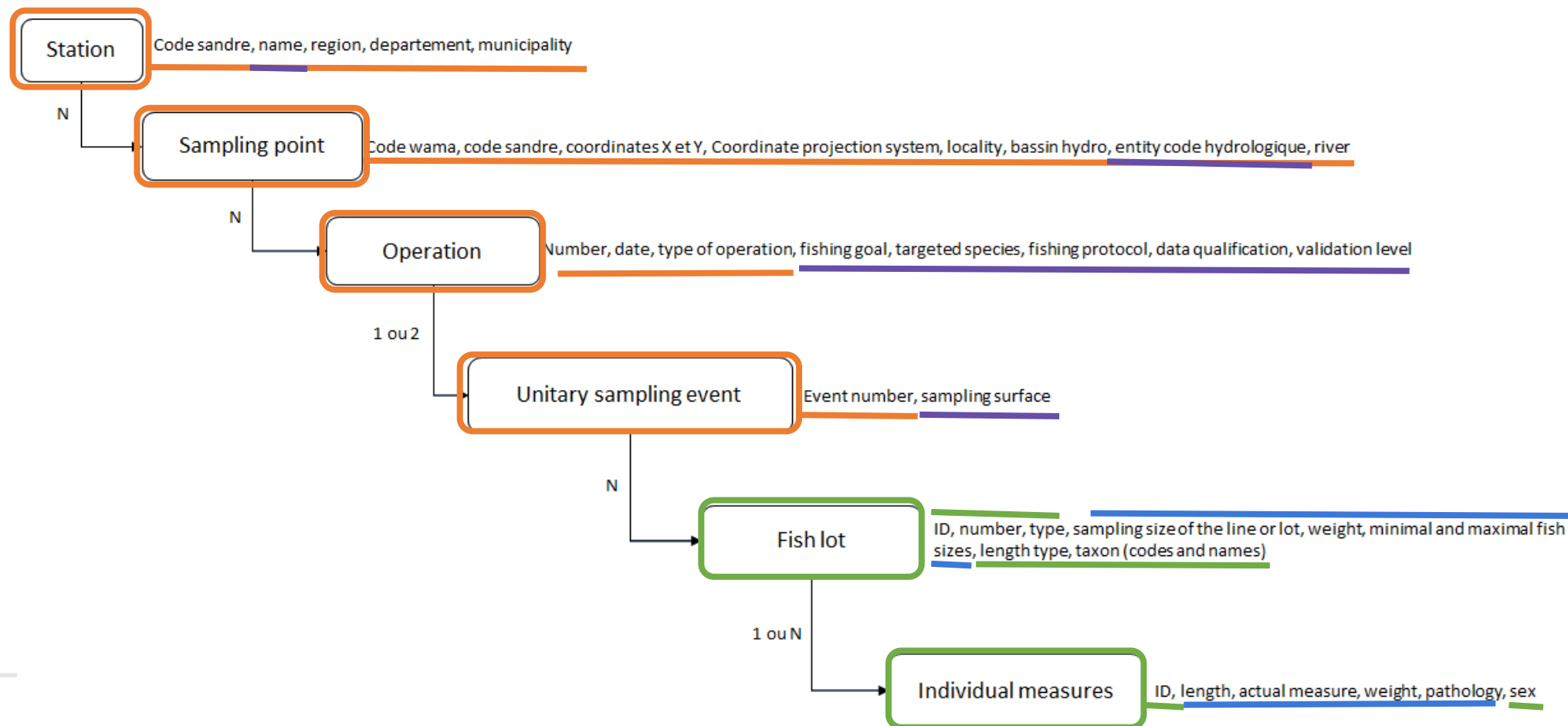
Measurement of fact extension:
for measures or other facts linked to an observation

measurementID	measurementType	measurementValue	measurementUnit
value	value	value	value



Darwin core and monitoring data

Electrofishing protocol





Electrofishing protocol

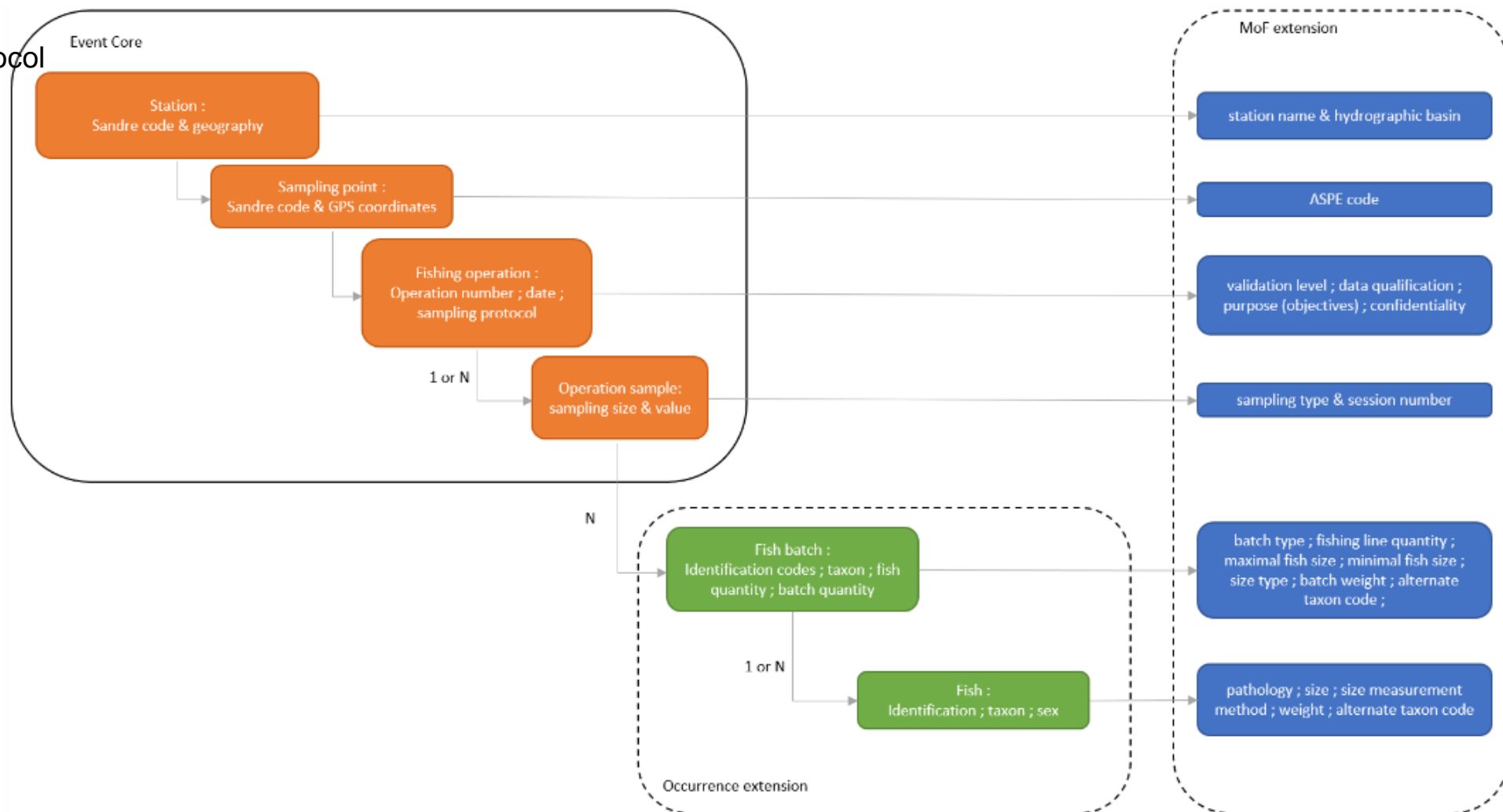


Figure 6: Scheme of variables in electrofishing data as DwC terms



eventID	locationID	waterBody	stateProvince	county	municipality	verbatimLocality
event1	4613007	La Dême	Centre-Val-de-Loire	Indre-et-Loire	Beaumont-La-Ronce	Pont de la D29 – Lieu-dit Moque-Souris

eventID	parentEventID	locationID	decimalLatitude	decimalLongitude	geodeticDatum
event1:1	event1	6	524028.00000000	6728114.00000000	EPSG:2154

eventID	parentEventID	fieldNumber	eventDate	samplingProtocol	sampleSizeValue	sampleSizeUnit
event1:1:1	event1:1	90170	23-10-12	Complete electrofishing with one or more sessions	351.00	square meter

eventID	parentEventID
event1:1:1:1	event1:1:1

measurementID	eventID	measurementType	measurementValue
event1/mea1	event1	Station name	R DÈME A LOUESTAULT
event1/mea2	event1	Hydrographic basin name	Loire-Bretagne
event1/mea3	event1	Hydrographic entity code	M1345800

event1:1/mea1	event1:1	ASPE (Wama) code	
---------------	----------	------------------	--

event1:1:1/mea1	event1:1:1	operation is confidential	No
event1:1:1/mea2	event1:1:1	operation validity level	Validity level 1
event1:1:1/mea3	event1:1:1	data qualification	correct
event1:1:1/mea4	event1:1:1	fishery purpose	RRP – Réseau de Référence Pérenne

event1:1:1:1/mea1	event1:1:1:1	operation sampling type	session
event1:1:1:1/mea2	event1:1:1:1	session number	1

Figure 7: Transcription of the event level as seen in Figure 6



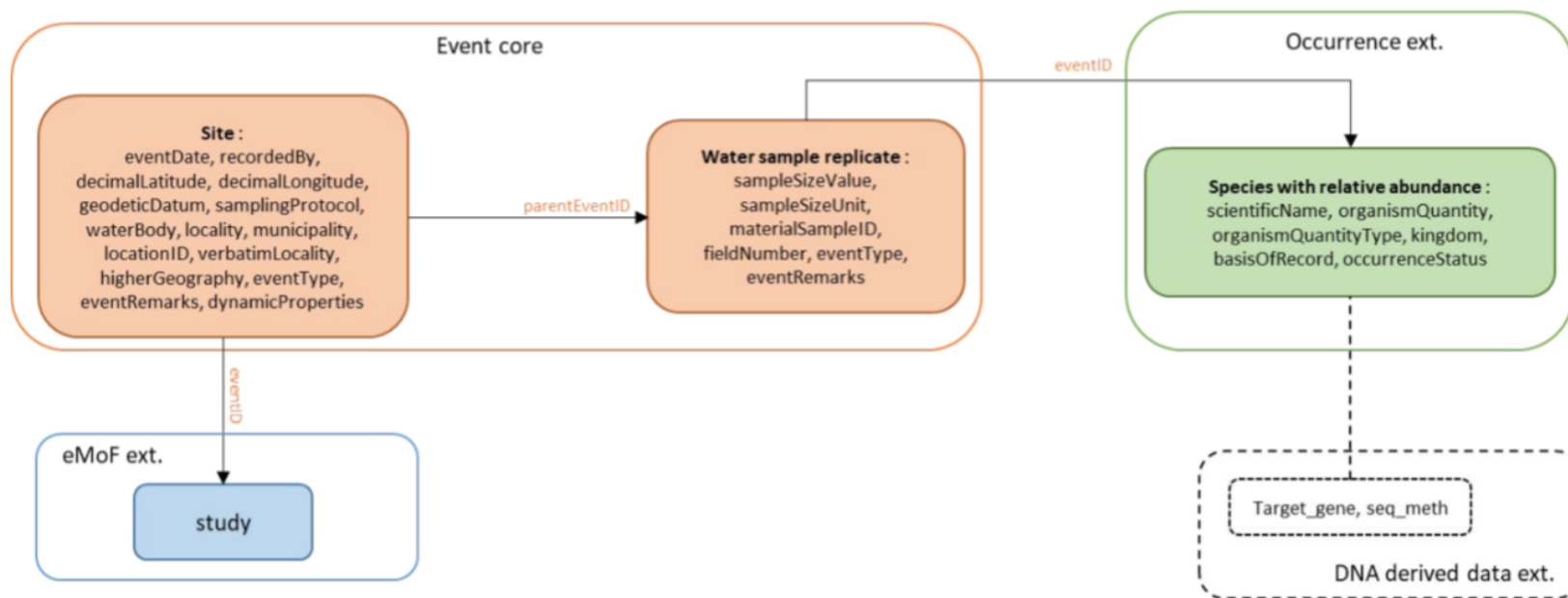
occurrenceID	eventID	catalogNumber	recordNumber	basisOfRecord	occurrenceStatus	scientificName	scientificNameAutorship	taxonID	nameAccordingTo	vernacularName	individualCount
event1:1:1:1/occ1	event1:1:1:1	5648513	1	HumanObservation	present	Salmo trutta fario	Linnaeus, 1758	67778	TaxRef	River trout (riverine ecotype)	1
event1:1:1:1/occ2	event1:1:1:1	5648575	63	HumanObservation	present	Barbatula barbatula	Linnaeus, 1758	67550	TaxRef	Stone loach	3
event1:1:1:1/occ3	event1:1:1:1	5648569	57	HumanObservation	present	Phoxinus phoxinus	Linnaeus, 1758	67404	TaxRef	Common minnow	158

measurementID	occurrenceID	measurementType	measurementValue	measurementUnit
event1:1:1:1/occ1/mea1	event1:1:1:1/occ1	Batch type	N	
event1:1:1:1/occ1/mea2	event1:1:1:1/occ1	Line quantity	1	
event1:1:1:1/occ1/mea3	event1:1:1:1/occ1	Total batch weight	36	g gram
event1:1:1:1/occ1/mea4	event1:1:1:1/occ1	Batch individuals maximal size		mm millimeter
event1:1:1:1/occ1/mea5	event1:1:1:1/occ1	Batch individuals minimal size		mm millimeter
event1:1:1:1/occ1/mea6	event1:1:1:1/occ1	Batch size type	total	
event1:1:1:1/occ1/mea7	event1:1:1:1/occ1	Taxa alternate code	TRF	



occurrenceID	eventID	parentOccurrenceID	individualCount	catalogNumber	basisOfRecord	occurrenceStatus	scientificName	scientificNameAutorship	taxonID	nameAccordingTo	vernacularName	sex
event1:1:1:1/occ1:1	event1:1:1:1	event1:1:1:1/occ1	1	22483116	HumanObservation	present	Salmo trutta fario	Linnaeus, 1758	67778	TaxRef	River trout (riverine ecotype)	
event1:1:1:1/occ2:1	event1:1:1:1	event1:1:1:1/occ2	1	22483352	HumanObservation	present	Barbatula barbatula	Linnaeus, 1758	67550	TaxRef	Stone loach	
event1:1:1:1/occ2:2	event1:1:1:1	event1:1:1:1/occ2	1	22483353	HumanObservation	present	Barbatula	Linnaeus, 1758	67550	TaxRef	Stone loach	
event1:1:1:1/occ2:3	event1:1:1:1	event1:1:1:1/occ2	1	22483354	HumanObservation	present	Barbatula	Linnaeus, 1758	67550	TaxRef	Stone loach	
event1:1:1:1/occ3:1	event1:1:1:1	event1:1:1:1/occ3	1	22483355	HumanObservation	present	Phoxinus	Linnaeus, 1758	67404	TaxRef	Common minnow	
event1:1:1:1/occ3:2	event1:1:1:1	event1:1:1:1/occ3	1	22483356	HumanObservation	present	Phoxinus	Linnaeus, 1758	67404	TaxRef	Common minnow	

measurementID	occurrenceID	measurementType	measurementValue	measurementUnit
event1:1:1:1/occ1:1/mea1	event1:1:1:1/occ1:1	Individual size	162	mm
event1:1:1:1/occ1:1/mea2	event1:1:1:1/occ1:1	Is real size	yes	
event1:1:1:1/occ1:1/mea3	event1:1:1:1/occ1:1	Individual weight		g
event1:1:1:1/occ1:1/mea4	event1:1:1:1/occ1:1	Pathology	no	



eDNA protocole



Citizen science bird count

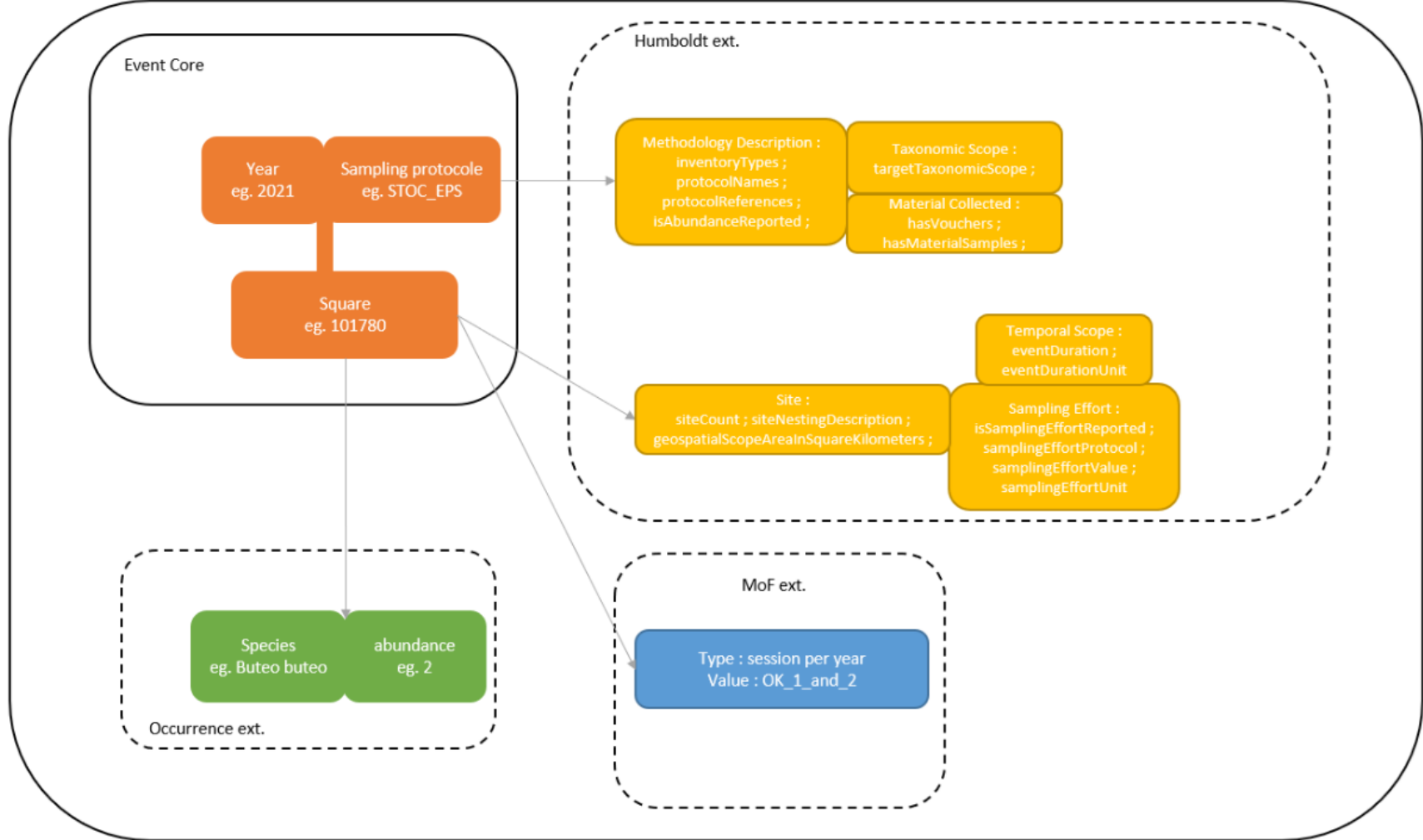


Figure 3 : Scheme of variables in STOC-EPS data as DwC terms

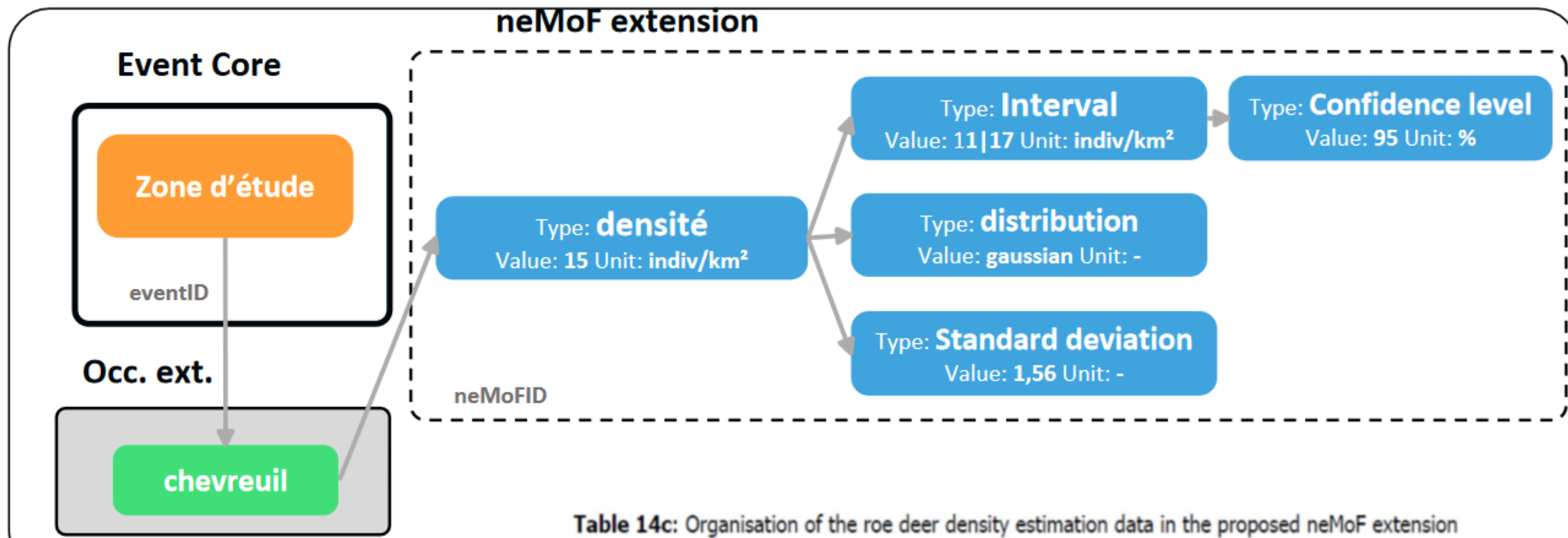
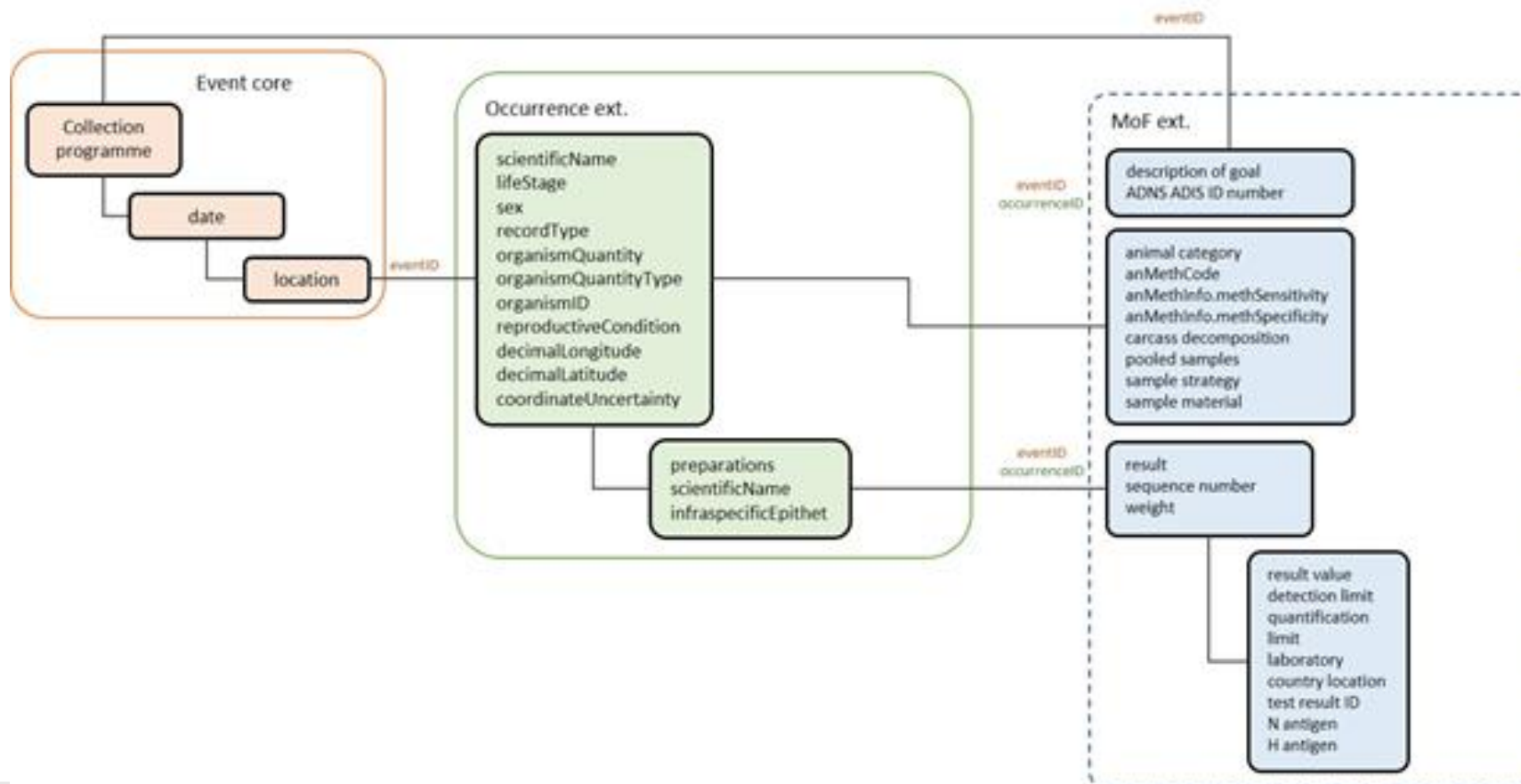


Table 14c: Organisation of the roe deer density estimation data in the proposed neMoF extension

Roe deer density estimate



Epidemiological data: the SIGMA database





Difficulties and advices

It works !! most of the time.

Request recursivity in occurrences and in measurements,

(not yet included in Darwin Core official, nor in GBIF/OBIS, but new GBIF model should allows it)

“EventType” would be useful to clearly state the kind of event we are talking about

Humbolt core to precise/ajust while trying.

Still some issues:

Habitat/Ecosystem approach not so well developed. Darwin core remains taxon oriented

Non biological observation (boats, pollution) not included, but should be easy to develop

Which event first : location or time or sampling => no rules, but easier while following the protocol logic

Graphical metadata useful, and color code useful as well to ease the understanding



More documentation necessary, and guideline for monitoring

Reference:

- De Pooler D. et al. (2017) Toward a new data standard for combined marine biological and environmental datasets - expanding OBIS beyond species occurrences. Biodiversity Data Journal 5: e10989 <https://bdj.pensoft.net/articles.php?id=10989>
- Enetwild conApplying the Darwin core standard to the monitoring of wildlife species, their management and estimated records: <https://doi.org/10.2903/sp.efsa.2020.EN-1841>
- Applying the Darwin Core data standard to wildlife disease – advancements toward a new data model <https://doi.org/10.2903/sp.efsa.2022.EN-7667>
- Darwin core terms: <https://dwc.tdwg.org/terms/>

Tools to develop:

try this Shiny app (and improve it): <https://github.com/fja062/WLDM.standardisation>



Thank you

sophie.pamerlon@mnhn.fr
guillaume.body@ofb.gouv.fr



biodiversa+
European Biodiversity Partnership

**Biodiversa+ Capacity building workshop
on Darwin Core standard**

7th of June from 9:00 to 12:40 CEST

Use of DwC standard in the context of a national participation in GBIF

Rui Figueira
Nó Português do GBIF



INSTITUTO
SUPERIOR DE
AGRONOMIA
Universidade de Lisboa



nó português do **GBIF**

GBIF – Global Biodiversity Information Facility, created in 2001



2.6 Bi shared records

GBIF – Global Biodiversity Information Facility, created in 2001

GBIF—the Global Biodiversity Information Facility—is an international network and data infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth.



2.6 Bi shared records

What is GBIF? - example

INVASIVE SPECIES:

Yellow-legged Asian hornet (*Vespa velutina*) expansion in Europe

2005



What is GBIF? - example

INVASIVE SPECIES:

Yellow-legged Asian hornet (*Vespa velutina*) expansion in Europe

2010

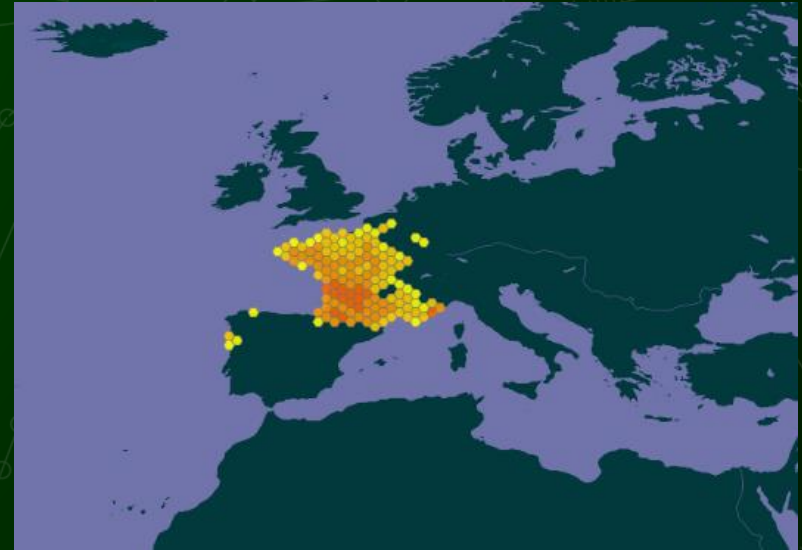


What is GBIF? - example

INVASIVE SPECIES:

Yellow-legged Asian hornet (*Vespa velutina*) expansion in Europe

2015

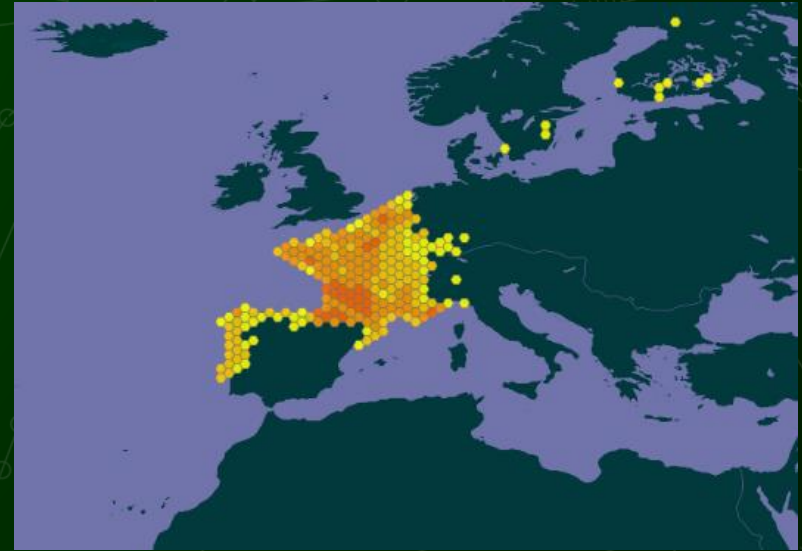


What is GBIF? - example

INVASIVE SPECIES:

Yellow-legged Asian hornet (*Vespa velutina*) expansion in Europe

2020

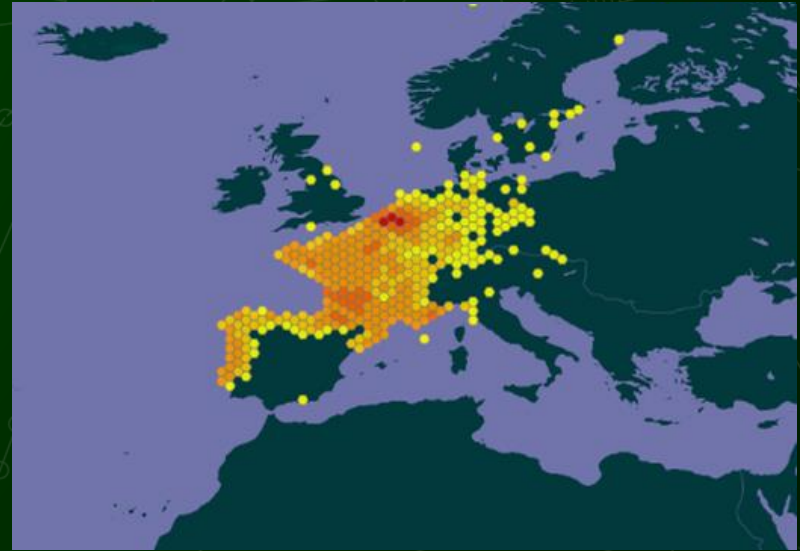


What is GBIF? - example

INVASIVE SPECIES:

Yellow-legged Asian hornet (*Vespa velutina*) expansion in Europe

2024



Datasets ●
91,826

● Hosted portals
19

Country
Participants ●
62

● Peer-review papers
using data
9,923

Organizational
Participants ●
43

● Average records
downloaded per month (2023)
173.8 billion

Publishers ●
2,170

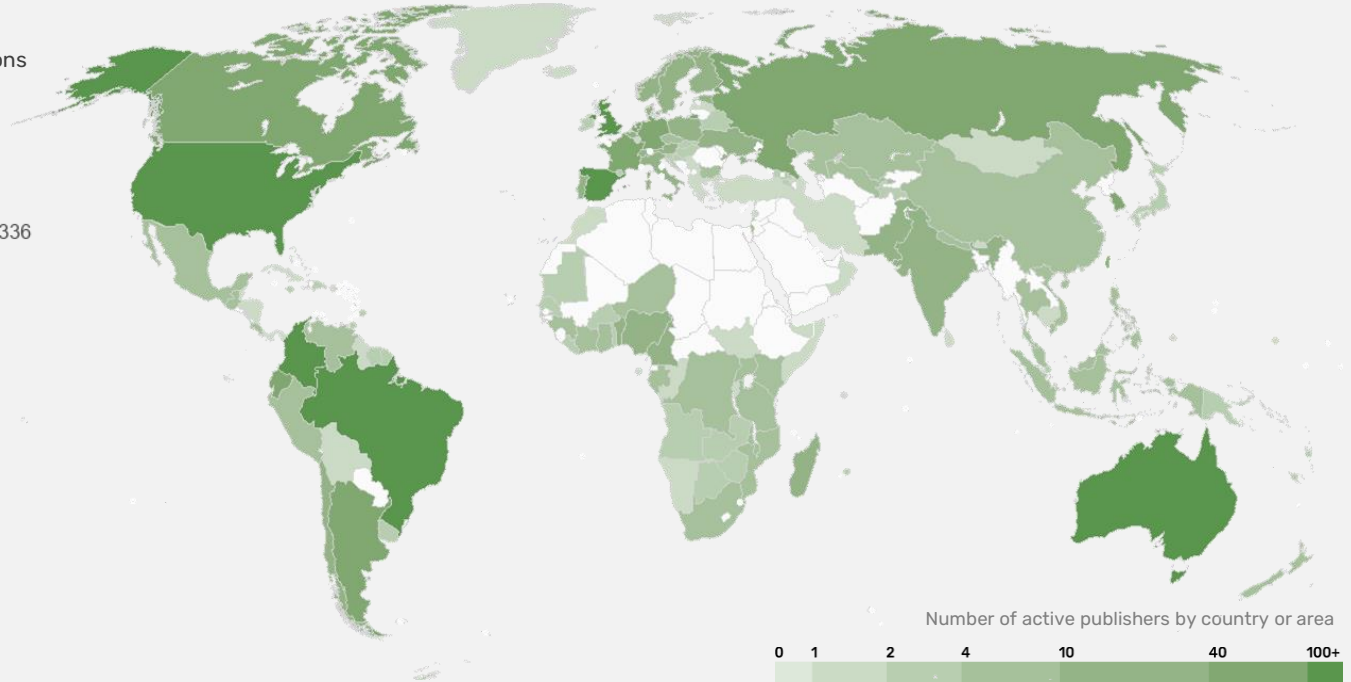
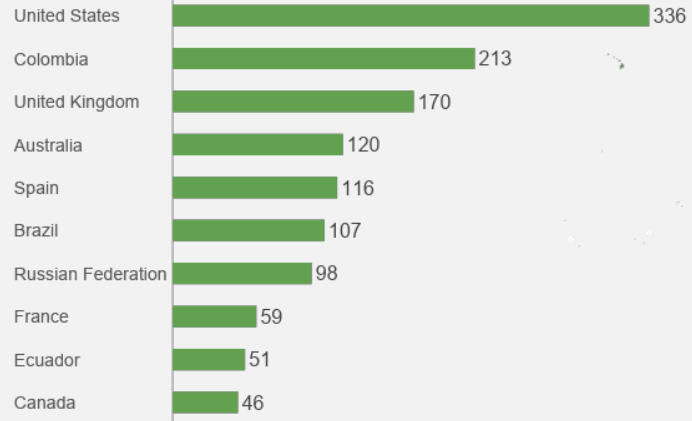
Species
● occurrence records
2,627,518,860



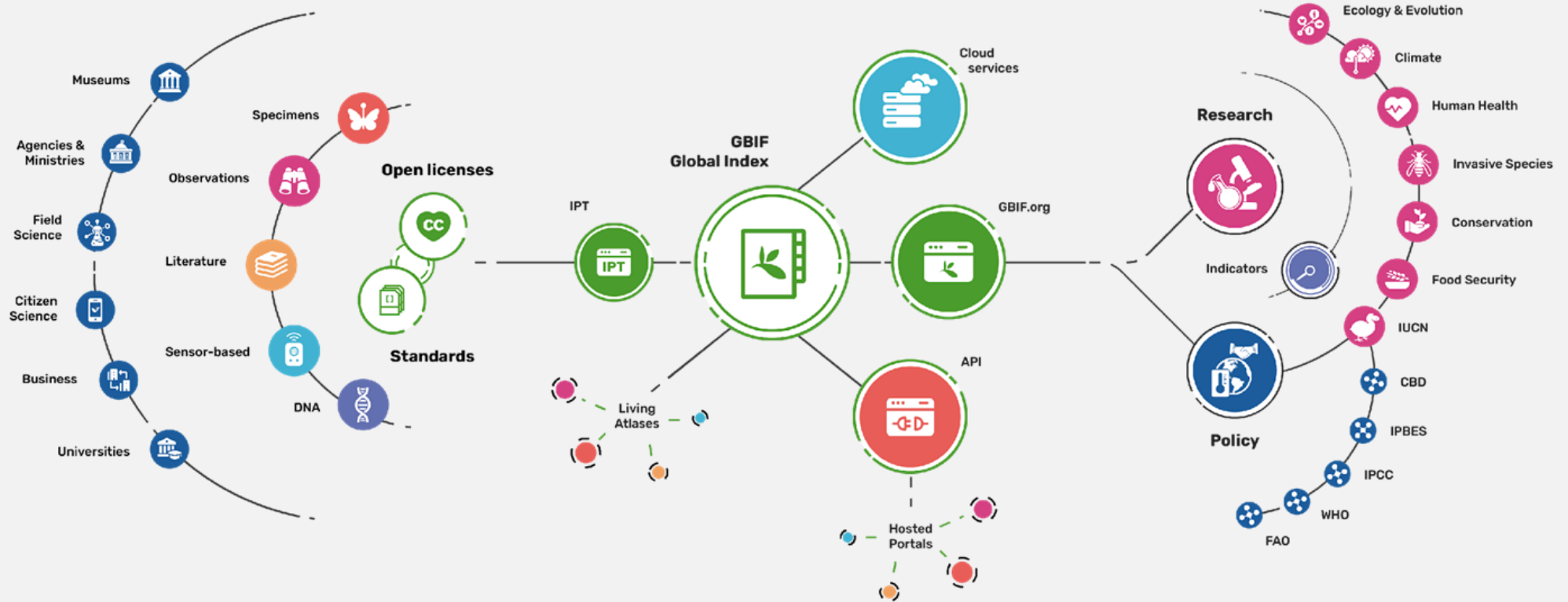
GBIF network of data publishing institutions

142

countries/areas with institutions
sharing data through GBIF



Providing biodiversity evidence for research and policy



GBIF IN PORTUGAL

Imagem: César García



Information dimensions

What

Taxonomic/nomenclatural
data

Where

Spatial data

Who

Occurrence
data

When

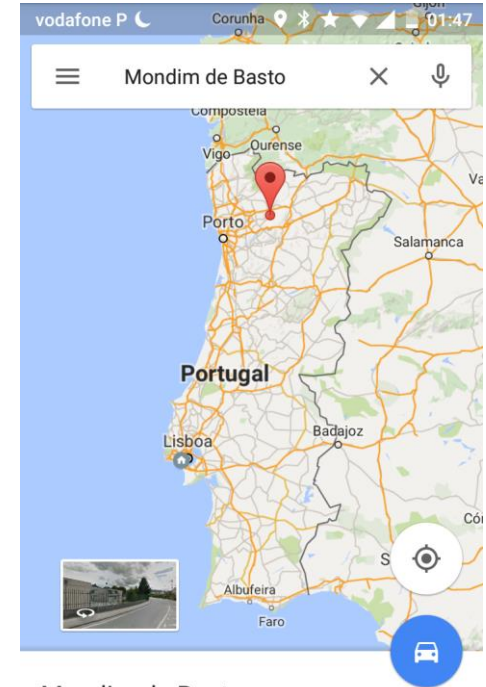
Sampling data

What

Descriptive data



@CésarGarcia, 2010/11/23
Amanita muscaria



Mondim de Basto

Information dimensions

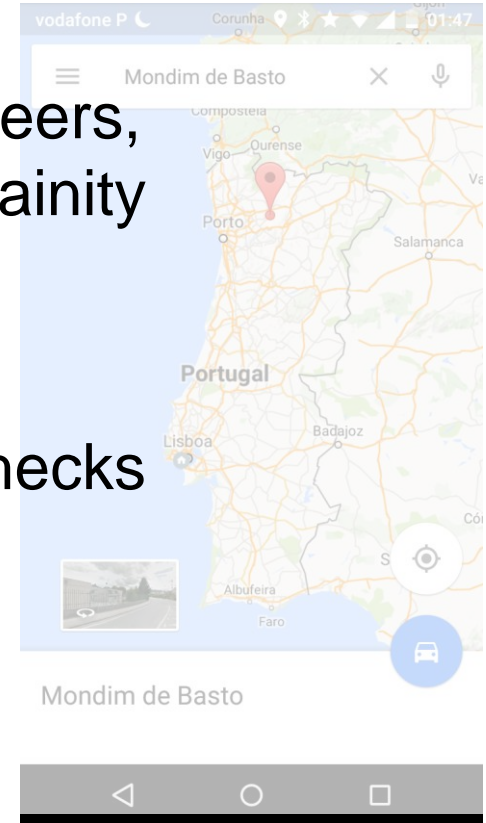
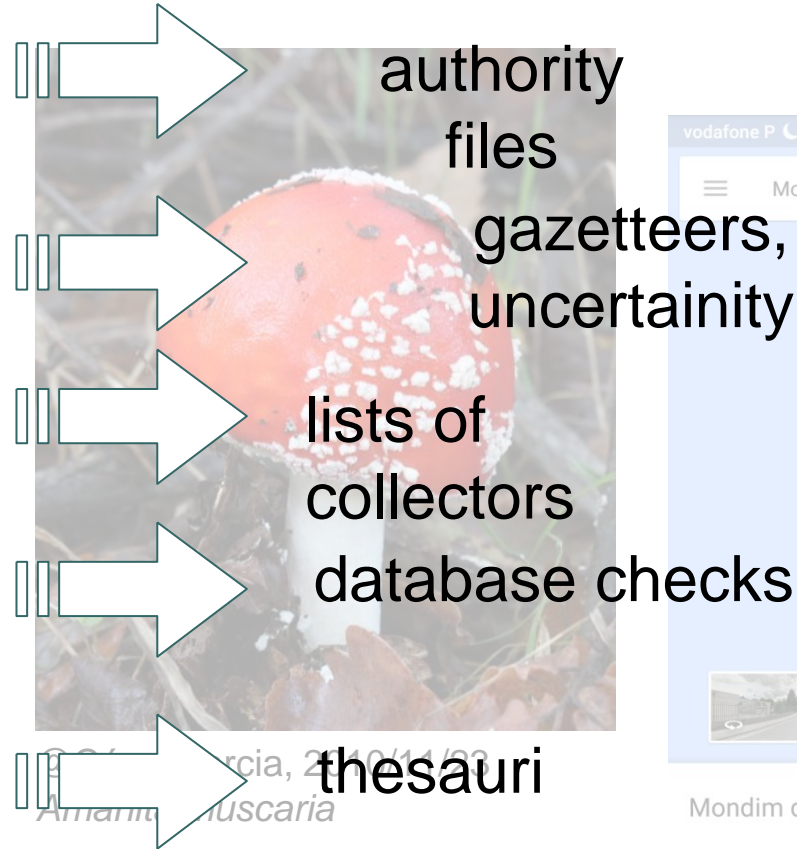
What
Taxonomic/nomenclatural
data

Where
Spatial data

Who
Occurrence
data

When
Sampling data

What
Descriptive data





Biodiversity Information Standards (TDWG)

We are a non-profit organization and a community
dedicated to developing biodiversity information
standards.

Image by Jennifer Latuperisa-Andresen



Darwin Core

Darwin Core is a standard maintained by the Darwin Core maintenance group. It includes a glossary of terms intended to facilitate the sharing of information about biological diversity by providing identifiers, labels, and definitions. Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, samples, and related information.

[Website](#)[GitHub !\[\]\(a870788d6ed9b8fd294b7654a8c8526b_img.jpg\)](#)

Image by Alex Guillaume



TDWG

Darwin Core

Record-level

Occurrence

Organism

MaterialEntity

MaterialSample

Event

Location

GeologicalContext

Identification

Taxon

MeasurementOrFact

ResourceRelationship

UseWithIRI

LivingSpecimen

PreservedSpecimen

FossilSpecimen

MaterialCitation

HumanObservation

MachineObservation

Cite Darwin Core

Record-level

This category contains terms that are generic in that they might apply to any type of record in a dataset.

type

modified

language

license

rightsHolder

accessRights

bibliographicCitation

references

institutionID

collectionID

datasetID

institutionCode

collectionCode

datasetName

ownerInstitutionCode

basisOfRecord

informationWithheld

dataGeneralizations

dynamicProperties

Occurrence

occurrenceID

catalogNumber

recordNumber

recordedBy

recordedByID

individualCount

organismQuantity

organismQuantityType

sex

lifeStage

reproductiveCondition

caste

behavior

vitality

establishmentMeans

degreeOfEstablishment

pathway

georeferenceVerificationStatus

occurrenceStatus

associatedMedia

associatedOccurrences

associatedReferences

associatedTaxa

otherCatalogNumbers

occurrenceRemarks

Event

eventID

parentEventID

eventType

fieldNumber

eventDate

eventTime

startDayOfYear

endDayOfYear

year

month

day

verbatimEventDate

habitat

samplingProtocol

sampleSizeValue

sampleSizeUnit

samplingEffort

fieldNotes

eventRemarks



TDWG

Darwin Core

<https://dwc.tdwg.org/terms/>

Record-level



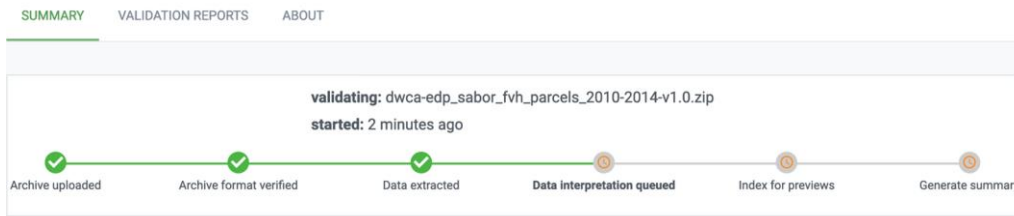
basisOfRecord		Property
Identifier	http://rs.tdwg.org/dwc/terms/basisOfRecord	
Definition	The specific nature of the data record.	
Comments	Recommended best practice is to use the standard label of one of the Darwin Core classes.	
Examples	PreservedSpecimen , FossilSpecimen , LivingSpecimen , MaterialSample , Event , HumanObservation , MachineObservation , Taxon , Occurrence	

Darwin Core - training

Yearly training workshops in one of the topics:

- data publication through GBIF
- data quality / data cleaning
- georeferencing
- Specify - management of biological collections

Extensive use of GBIF's and other resources, e.g., GBIF Data Validator



<https://www.gbif.org/tools/data-validator/89a13f02-e55e-4632-9b35-7b82f1b3f08e>

Darwin Core – core



occurrence.txt



event.txt



taxon.txt

Occurrences

<http://tools.gbif.org/dwca-validator/extension.do?id=dwc:Occurrence>

Events

<http://tools.gbif.org/dwca-validator/extension.do?id=dwc:Event>

Taxa

<http://tools.gbif.org/dwca-validator/extension.do?id=dwc:Taxon>

• Darwin Core – core

Occurrences

<http://tools.gbif.org/dwca-validator/extension.do?id=dwc:Occurrence>



occurrence.txt



event.txt



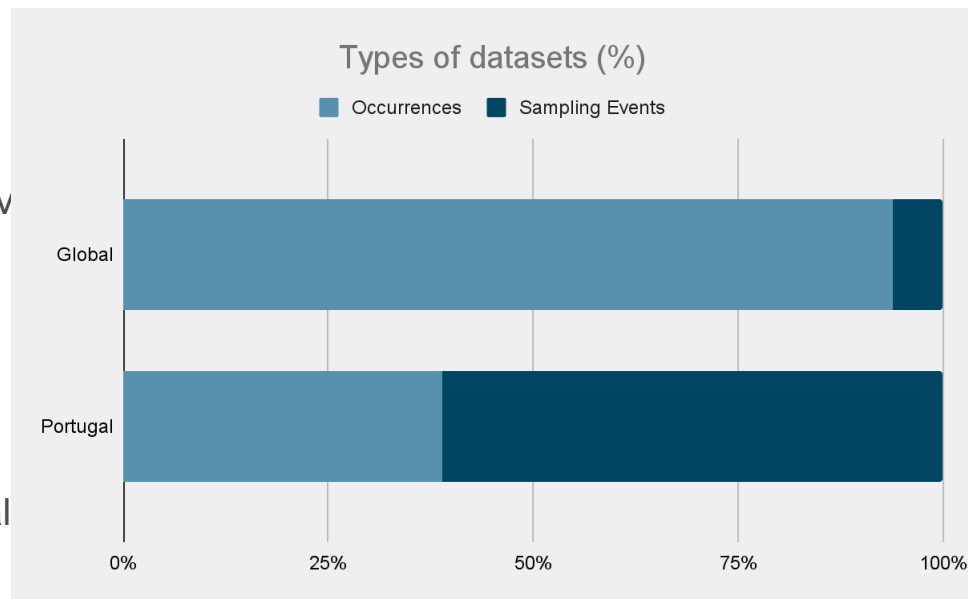
taxon.txt

Events

<http://tools.gbif.org/dwca-v>

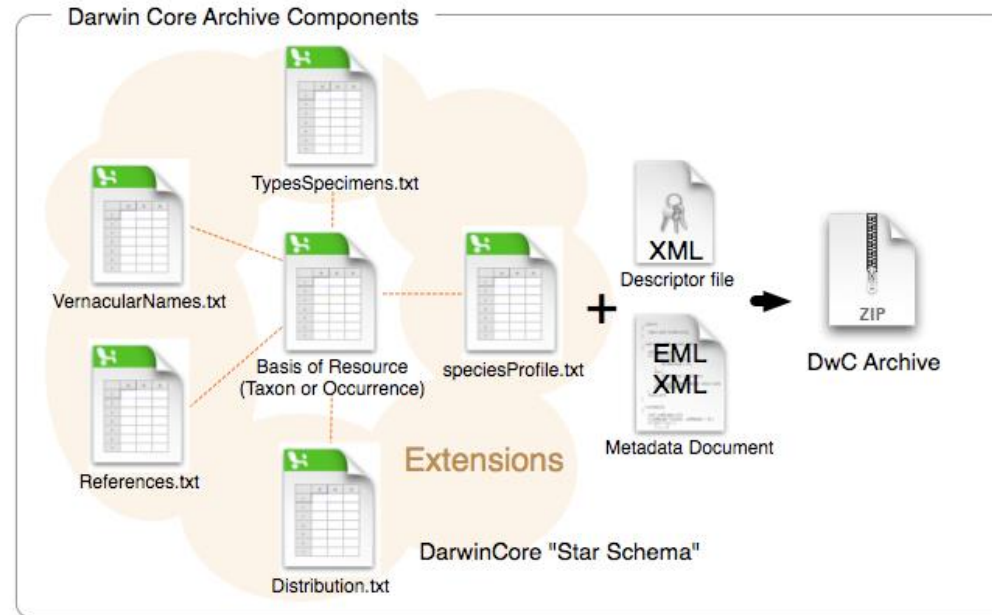
Taxa

<http://tools.gbif.org/dwca-val>



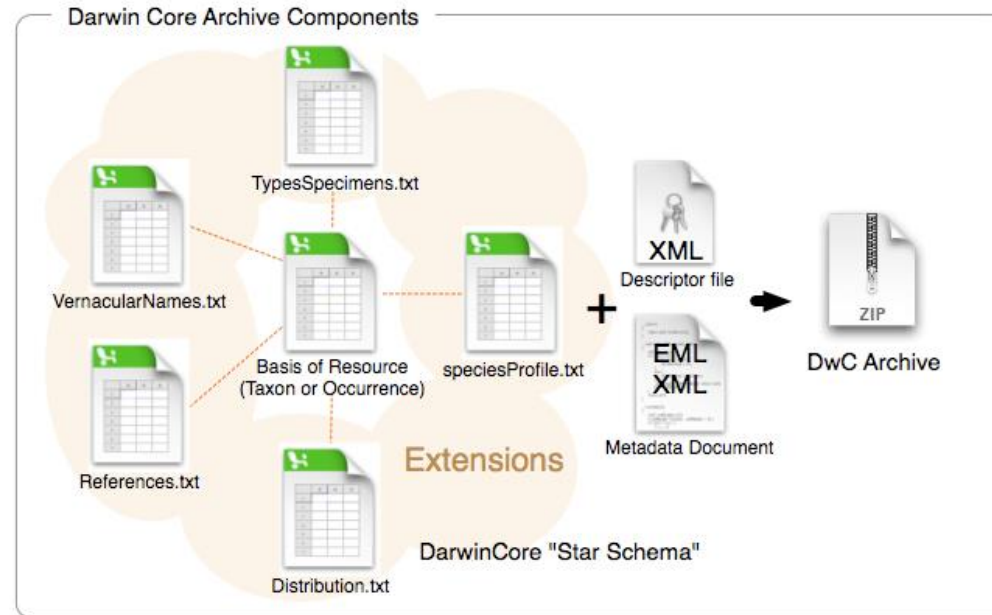
Darwin Core – Star schema, supporting extensions for several types of data

- germplasm
- multiple identifications
- types and specimens
- common names
- alternative identifiers
- species profile
- references in literature
- taxon description
- traits
- images, audio, videos
- others...



Darwin Core – Star schema, supporting extensions for several types of data,
but a constrain to more richer datasets...

- germplasm
- multiple identifications
- types and specimens
- common names
- alternative identifiers
- species profile
- references in literature
- taxon description
- traits
- images, audio, videos
- others...



Darwin Core - MoF

MeasurementOrFact

measurementID	parentMeasurementID	measurementType	measurementValue	measurementAccuracy
measurementUnit	measurementDeterminedBy	measurementDeterminedDate	measurementMethod	
measurementRemarks				

MeasurementOrFact Class	
Identifier	http://rs.tdwg.org/dwc/terms/MeasurementOrFact
Definition	A measurement of or fact about an <code>rdfs:Resource</code> (http://www.w3.org/2000/01/rdf-schema#Resource).
Comments	Resources can be thought of as identifiable records or instances of classes and may include, but need not be limited to instances of <code>dwc:Occurrence</code> , <code>dwc:Organism</code> , <code>dwc:MaterialEntity</code> , <code>dwc:Event</code> , <code>dcterms:Location</code> , <code>dwc:GeologicalContext</code> , <code>dwc:Identification</code> , or <code>dwc:Taxon</code> .
Examples	<p>the weight of a <code>dwc:Organism</code> in grams</p> <p>the number of placental scars</p> <p>surface water temperature in Celsius</p>

Darwin Core – dynamicProperties

Record

Basis of record	Human observation
Collection code	edp_sabor_fvh_parcelas
Dataset name	EDP Baixo Sabor: Construction and Reservoir Filling Phases - Vascular Flora and Habitats: species from priority habitats/communities - [2010-2014]
Dynamic properties	{ <code>"lithology": "schists"</code> , <code>"% rock": 35</code> , <code>"exposureInCardinalDirections": "W"</code> , <code>"slopeInDegrees": 30</code> }
Institution code	EDP
Institution ID	e5150835-f502-424c-b470-24dd496b1b18
Owner institution code	EDP



GBIF Data Model Work

Tim Robertson
John Wiecezorek

<https://www.gbif.org>

 **GBIF** | Global Biodiversity
Information Facility

< 1 > ⋮

Google Slides

<https://globalnodes.gbif.org/en/gnm/data-model#meeting>

GBIF 20-year review

“...biodiversity data is **more complicated than 'just' the occurrence of species** in time and space; organisms interact, co-occur, move and evolve.

This implies a need for **richer and more varied types of data** than GBIF has thus far supported”

IV SUMMARY OF RECOMMENDATIONS

The high expectations on GBIF as well as the technical and scientific challenges ahead strongly suggest that, while GBIF is well positioned to meet them, a scenario of “business as usual” may fail to preserve GBIF’s relevance. Strong growth should be considered seriously, with implications for GBIF’s technology, services, organization, and funding.

Our specific recommendations are generally ordered in a sequence that reflects a plausible phasing of related activities. Thus, we present our recommendations in three sub-sections. The first one sets forth generic guidelines mainly formulated to maintain a focus on what we have identified as the main factors in GBIF’s success and reputation. Specific short-term

recommendations in the second section aim to encourage progress on the most urgent challenges, which could be achieved at the current funding level and structure, and without revising the organization’s strategy. The third section provides recommendations related to all of the challenges, but which probably can only be approached by a longer-term activity or by a substantial expansion of funding in the next two funding cycles (of five years, each). That third set of recommendations thus might require an explicit expression in strategy, organization, or even governance. The recommendations that follow are cross-referenced to pages in the FULL REPORT.

GENERAL RECOMMENDATIONS

In order to maintain and strengthen GBIF’s relevance and standing, we recommend the consideration of growth paths in a number of dimensions. Such growth will enable GBIF to support biodiversity research broadly in the future. To do so, the organization will probably also need to participate in and contribute to the rapid development of its scientific data methods and standards.

R1a, Data Quality and Quantity (p.121): GBIF needs to build on and maintain its reputation as the most comprehensive source of openly available global occurrence data. This means that it should continue the trajectory of growth of the data that it mediates, in quantity as in quality, by extending and deepening relations within its network, and by supporting non-traditional types of biodiversity data.

R1b, Technology and standards (p.121): GBIF needs to maintain or attain leadership in essential technological and standardization areas related to biodiversity informatics. In order to do so, it should continue to work actively with other stakeholders in the “landscape”, such as the researchers at the forefront of such fields as metagenomics, remote sensing, and observation and cloud technologies, to keep abreast of developing data sources, standards, and technologies. GBIF should add a long-term focus on IT security.
















R2, Networking (p.121): GBIF should maintain and even strengthen its capabilities to network its stakeholders and to lead them to cooperate and build consensus. GBIF, and particularly the Secretariat, should continue to be seen as a neutral broker.

CODATA, the Committee on Data of the International Science Council, Pfeiffenberger, Hans. Uhler, Paul, & Hodson, Simon. (2020). Twenty-Year Review of GBIF. <https://doi.org/10.35035/ctzm-hz97>



<https://globalnodes.gbif.org/en/gnm/data-model#meeting>

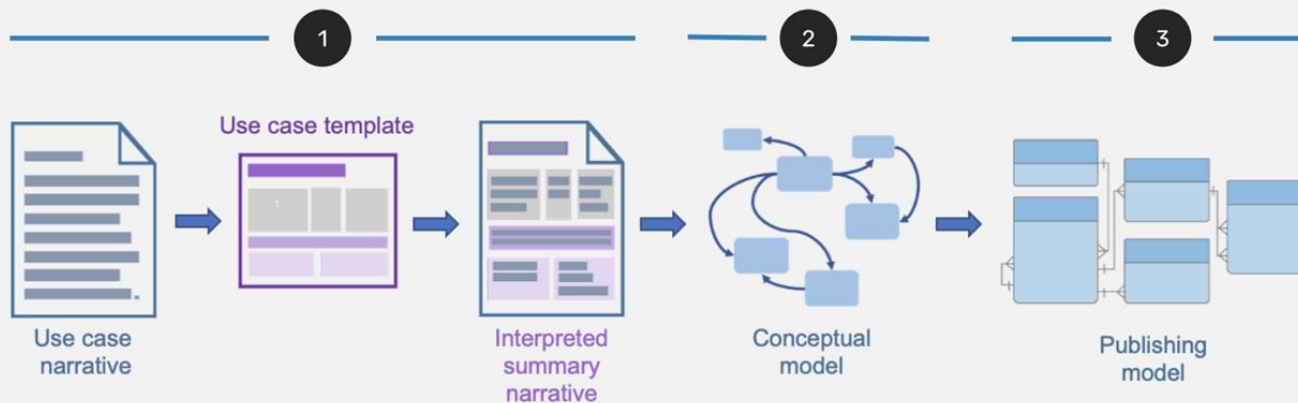
Case studies

				
Introduction	eDNA metabarcoding	Camera trap	Tissue samples	Automatic moth trap
				
Global malaise programme	iNaturalist	Specimens with media	Environmental and community measurements	Taxonomic treatments in literature
				
Malaise trapping for reference barcode collection	Taxonomic checklist	Ecological survey data exchange specification	Biotic interactions	Recording absence data

[gbif.org/new-data-model](https://globalnodes.gbif.org/en/gnm/data-model#meeting) 

<https://globalnodes.gbif.org/en/gnm/data-model#meeting>

Working through a case study

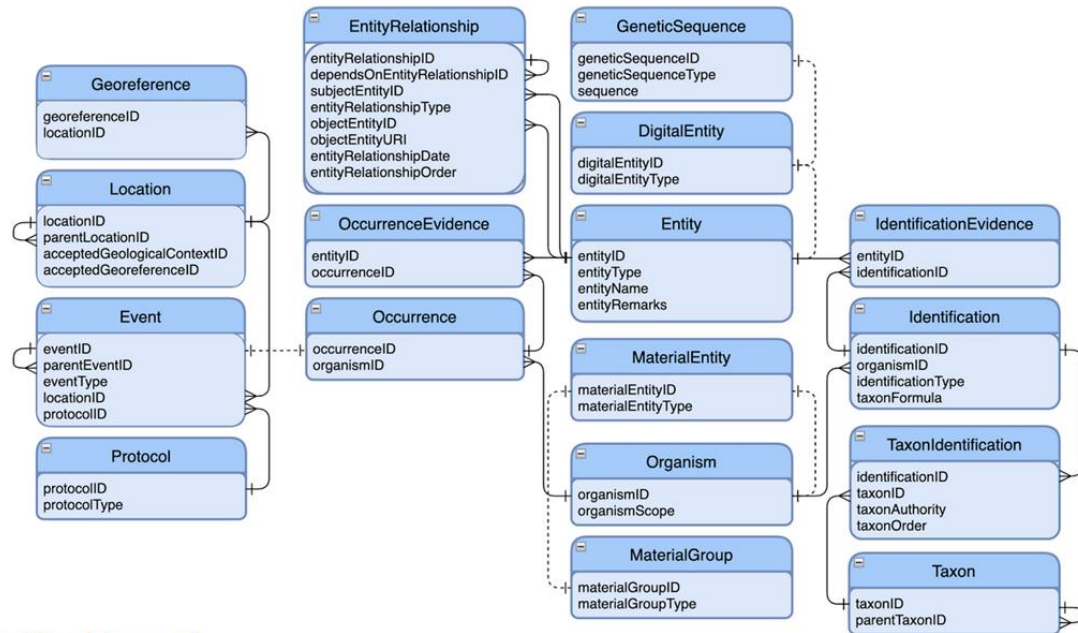


gbif.org/new-data-model

Google Slides

<https://globalnodes.gbif.org/en/gnm/data-model#meeting>

Instruction: “Map your data to this, please”

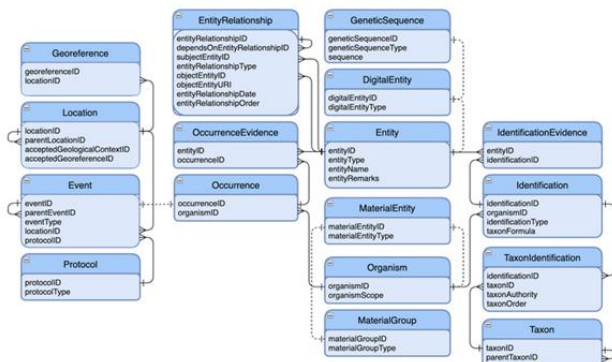


<http://github.com/gbif/model-material>



<https://globalnodes.gbif.org/en/gnm/data-model#meeting>

Instruction: “Map your data to this, please”



<http://github.com/gbif/model-material>

Psilorhynchus pseudocheneis

Indrawati River at bridge crossing through Metamachi Town

Catalogue item

Preparations: EIDH - 24 Tissue - 1
 Catalogue number: 40562
 Received by: Bentley, Andrew G.
 Institution: KUS
 Collection code: KUS

Collecting event

Field number: R.01.08.46
 Event date: 10/08/08
 Event remarks: KUSC 08.21

Occurrence

Organism quantity: 25
 Occurrence status: PRESERVED
 Higher geography: Asia, Nepal, Bagmati, Sindhupalchok
 Country: Nepal
 County: Sindhupalchok
 Locality: Indrawati River at bridge crossing through Metamachi Town

Identification

Scientific name: *Psilorhynchus pseudocheneis*
 Classification: *Psilorhynchus pseudocheneis* Minson & Collis, 1984
 Annotated: *Psilorhynchus pseudocheneis* Minson & Collis, 1984
 Identified by: Bentley, Andrew G.
 Verified: KUSC 08.21
 Version identification: *Psilorhynchus pseudocheneis*
 Nature of ID: unknown
 Date: December 16, 2008

Sequences

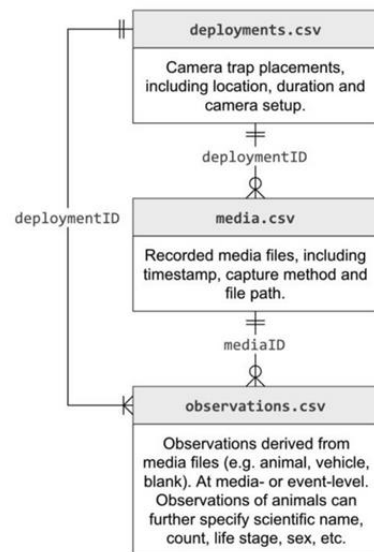
Object entity in: <https://www.ncbi.nlm.nih.gov/nucleotide/40562.1>
 Object entity in: <https://www.ncbi.nlm.nih.gov/nucleotide/40562.2>
 Object entity in: <https://www.ncbi.nlm.nih.gov/nucleotide/40562.3>
 Object entity in: <https://www.ncbi.nlm.nih.gov/nucleotide/40562.4>
 Object entity in: <https://www.ncbi.nlm.nih.gov/nucleotide/40562.5>
 Object entity in: <https://www.ncbi.nlm.nih.gov/nucleotide/40562.6>

Roles

Role	Name	Type
cataloguer	Bentley, Andrew G.	person
annotator	Conrady, Kevin W.	person
collector	Minson, Richard L.	person
collector	Conrady, Kevin W.	person

<https://globalnodes.gbif.org/en/gnm/data-model#meeting>

Publishing model: Camtrap DP



Credit: Tom Rickman, CC BY-NC
<https://www.inaturalist.org/observations/148735438>

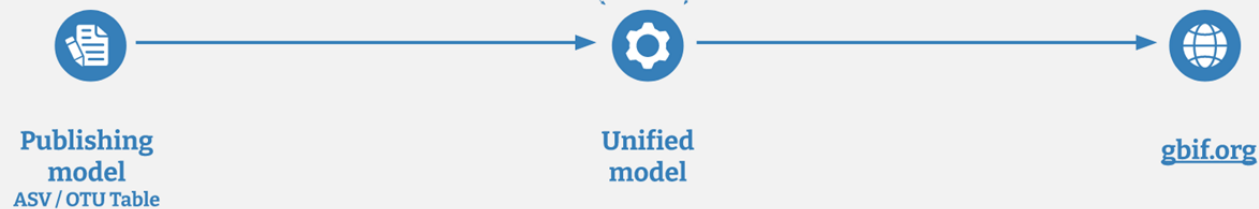
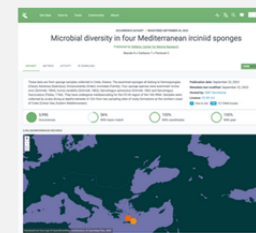
<https://bit.ly/camtrap-dp-tdwg-2023>



gbif.org/dna



Annotation

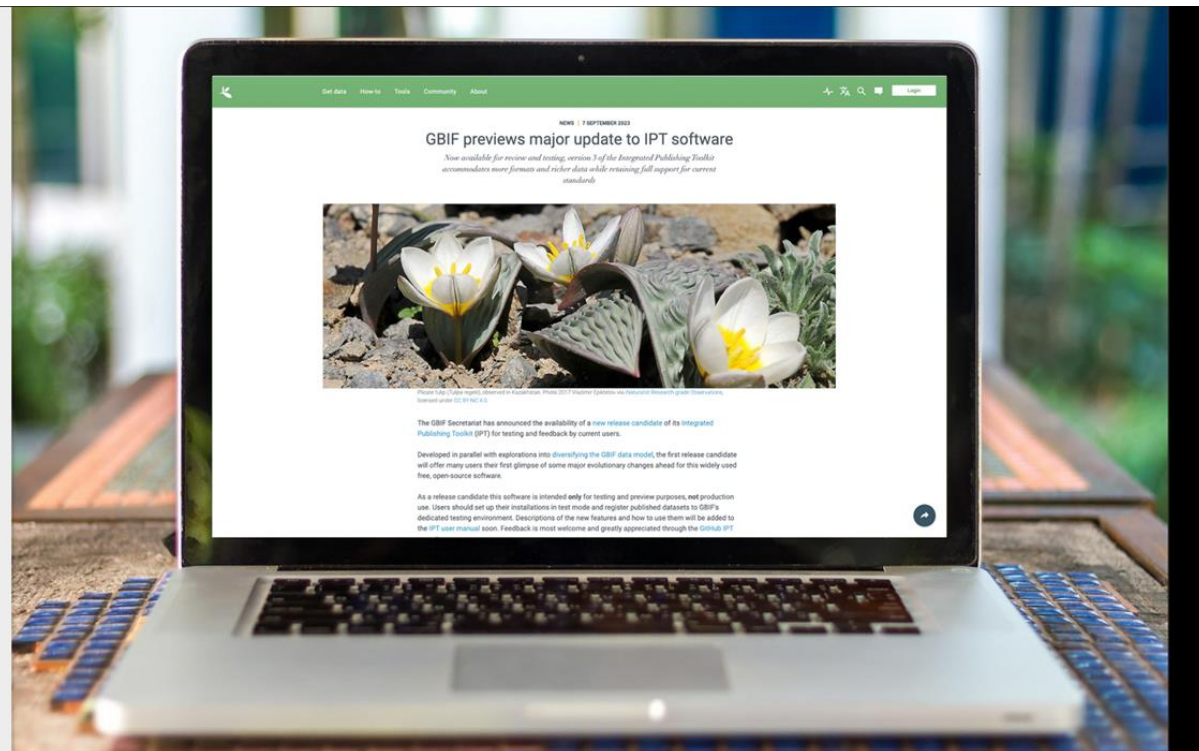
[illegible]

<https://globalnodes.gbif.org/en/gnm/data-model#meeting>

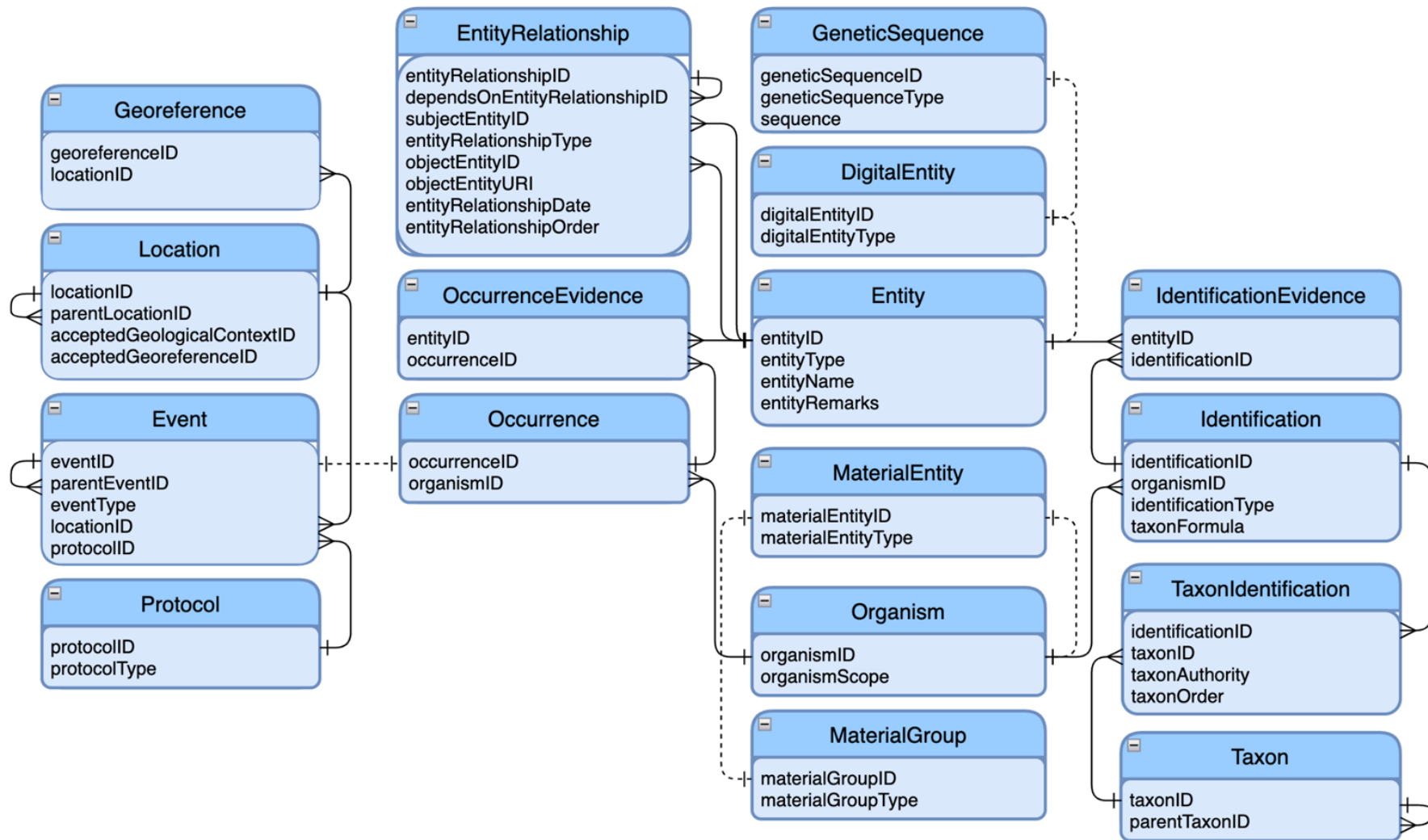
Integrated Publishing Toolkit

- Version 3.0 nearing release
- Supports mapping to Frictionless Data standards
- Critical to the support of new publishing models

<https://gbif.org/ipt>



<https://globalnodes.gbif.org/en/gnm/data-model#meeting>

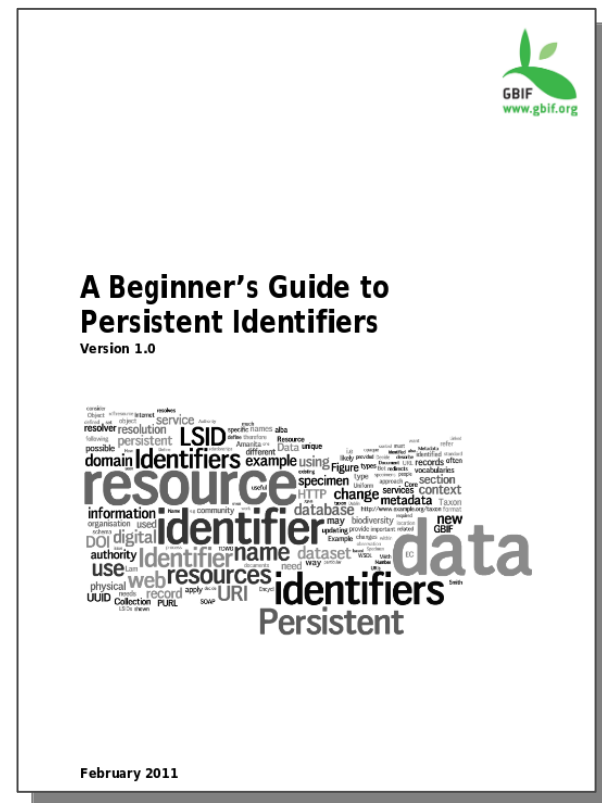


Persistent identifiers

Why do we find so many xxxxIDs under the DwC standard?
What do they mean and what are they for?

A Beginner's Guide to Persistent Identifiers

<https://www.gbif.org/document/80575>



Persistent identifiers - relation to FAIR

Why do we find so many xxxxIDs under the DwC standard?
What do they mean and what are they for?

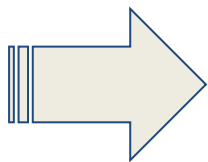
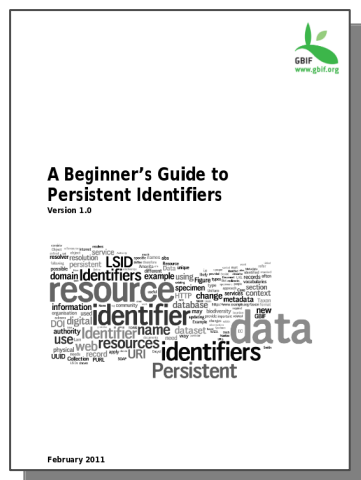


Table 9: Overall summary of the FAIRness assessment of the EGI-ACE Data Spaces.

Data Space	Findability (F)	Accessible (A)	Interoperable (I)	Reusability (R)
GBIF	100% Compliant	100% Complaint	86% Compliant	100% Compliant

<https://documents.egi.eu/document/3815>

Thank you!

RUI FIGUEIRA
NÓ PORTUGUÊS DO GBIF
INSTITUTO SUPERIOR DE AGRONOMIA
ruifigueira@isa.ulisboa.pt

The Portuguese Node is hosted by ISA-ULisboa with the support of



This content is available under an open access license
<https://creativecommons.org/licenses/by-sa/4.0/>



**Biodiversa+ Capacity building workshop
on Darwin Core standard**

7th of June from 9:00 to 12:40 CEST

The use of Darwin Core as standard terminology to harmonise biodiversity data in LifeWatch Italy

Andrea Tarallo

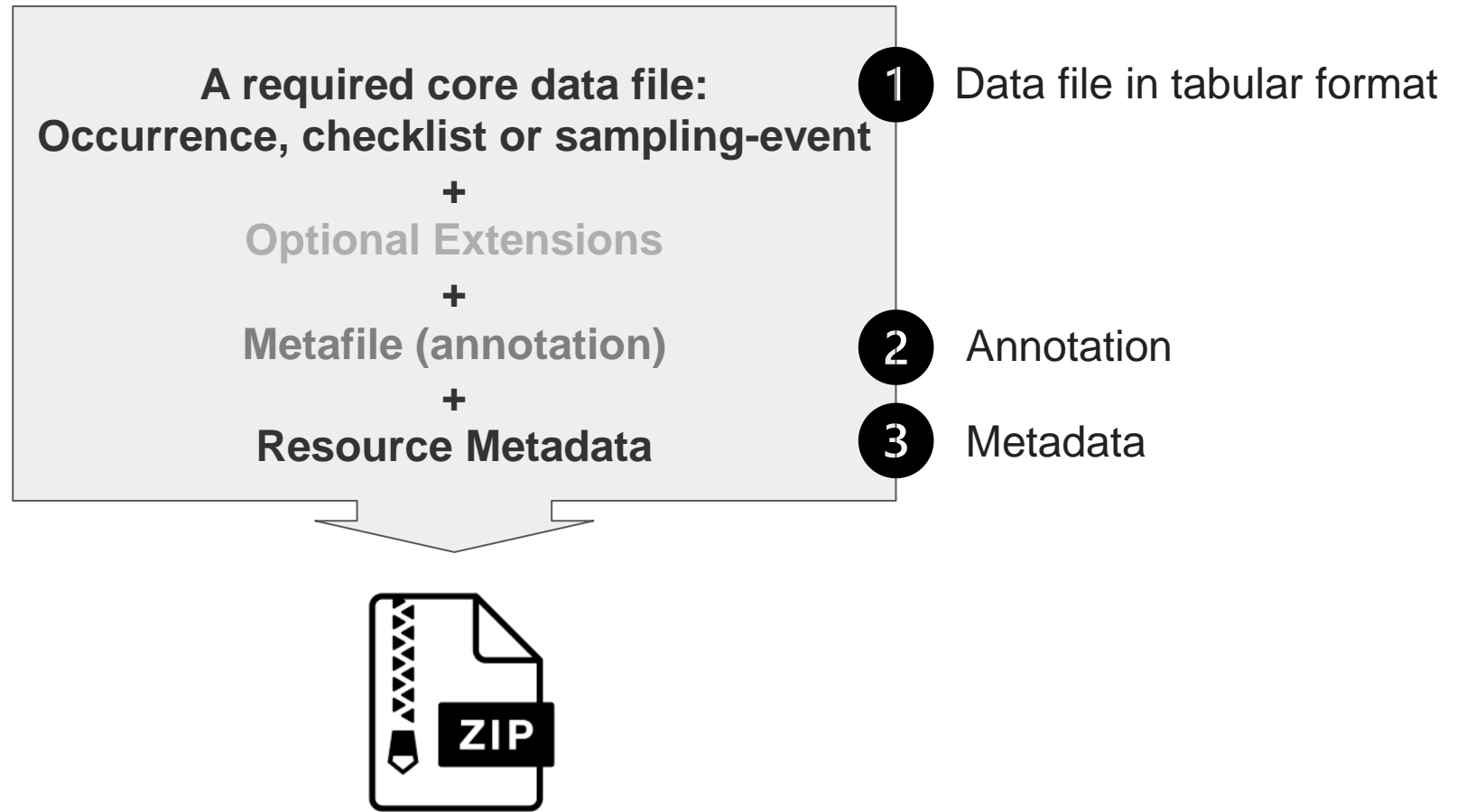
Institute of Research on Terrestrial Ecosystems (IRET)
National Research Council (CNR), Lecce, Italy



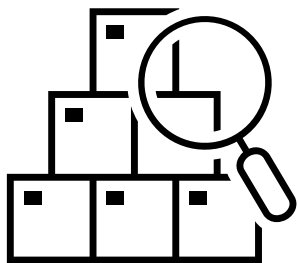
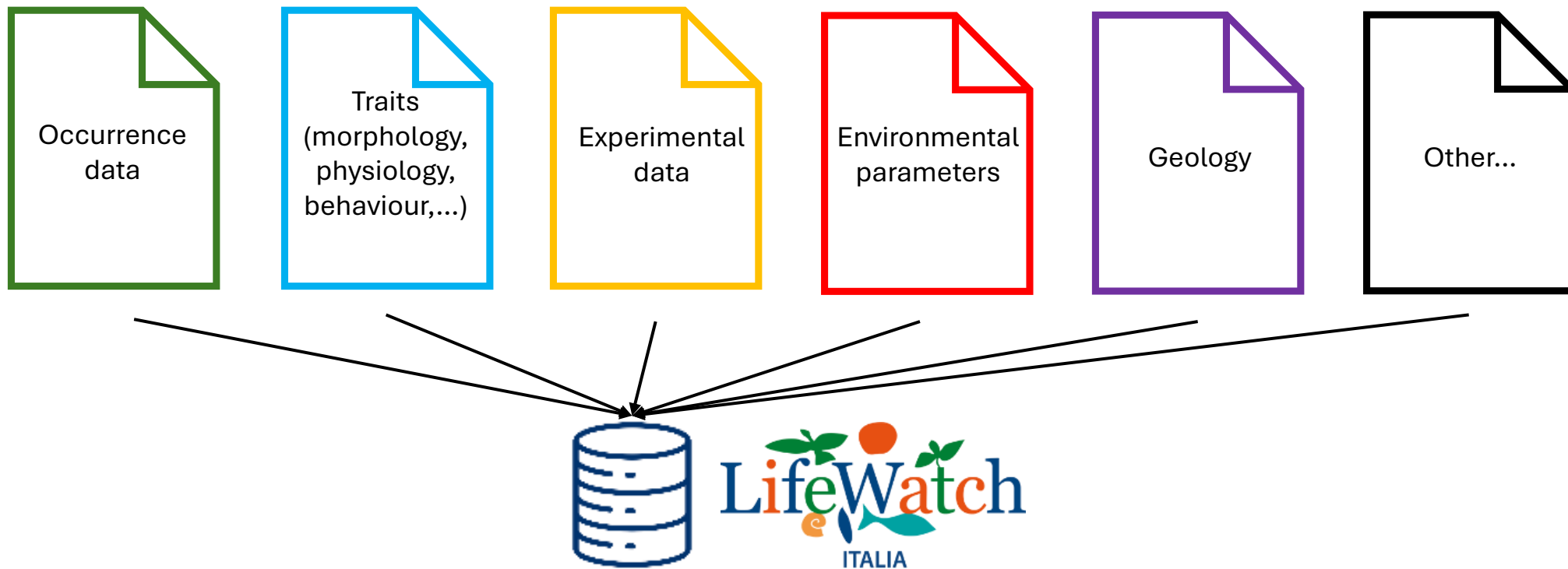
LifeWatch Italy Joint Research Unit

LifeWatch ITA has been established as a **Joint Research Unit (JRU)** in 2010 with 9 founding members **led by CNR**. It has now **35 members** including research institutes, universities, associations and private companies distributed, seeking to reinforce integrated scientific research on biodiversity and ecosystems.





The Darwin Core Archive is an effective way to mobilise and exchange specimen occurrence and observational records

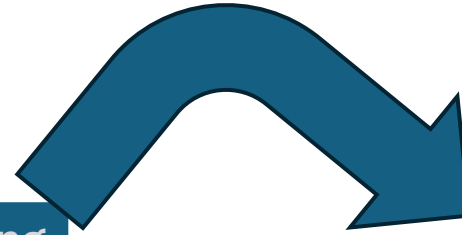


Flexible solution but that allow complex queries

Map variables of CSV files with standard vocabularies and thesauri

XLS files

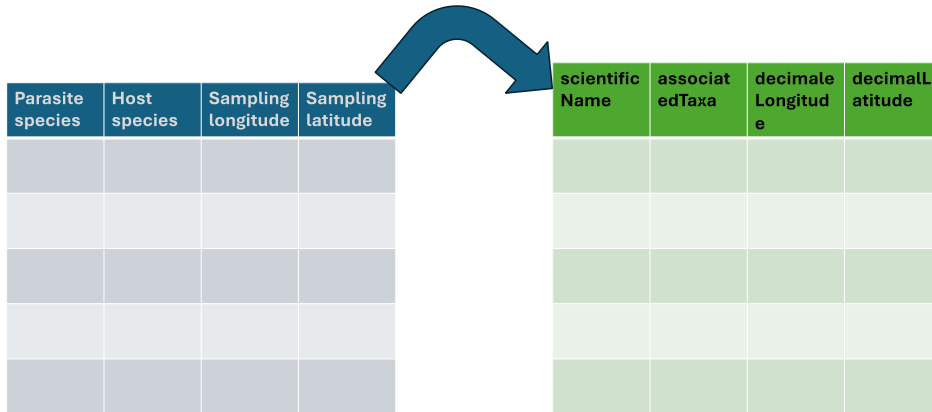
Parasite species	Host species	Sampling longitude	Sampling latitude



CSV UTF-8 files

scientific Name	associatedTaxa	decimal Longitude	decimal Latitude

Map variables of CSV files with standard vocabularies and thesauri



- Darwin Core
- EcoPortal (LifeWatch Italy thesauri and others)
- NERC Vocabulary Server
- BioPortal and other OntoPortal Alliance repositories
- Others



DATA PROVIDER FILE

Organism				
Id	Species	Taxonomy	Life.stage	SEX
1021315821	Carabus auratus	Coleoptera	adult	female
1134943631	Carabus auratus	Coleoptera	adult	male
1591440928	Oniscus asellus	Isopoda; Oniscidae; Oniscus	NA	female
633288236	Oniscus asellus	Isopoda; Oniscidae; Oniscus	NA	male
409193800	Oniscus asellus	Isopoda; Oniscidae; Oniscus	NA	unsexed
1050793085	Oreochromis mossambicus	fish	Fry	NA
1050793084	Oreochromis mossambicus	fish	Fry	NA
1050793083	Oniscus asellus	Oniscus asellus	Juvenile	unsexed

Metabolism							
Harmonized_MR_watts	Originalg_unit_MR	Harmonized_Body_mass_grams	Dry.weight	Measurment_Temperature_C	Coordinates	Ecosystem	
3,57E+09	Watts	11.9	4.95	5	NA	Terrestrial	
3,25E+09	ml/h	13.4	0.00191	10	NA	Terrestrial	
2,19E+09	ml/h	0.512	0.1617	15	67 30' N, 26 40' E	Terrestrial	
2,03E-02	μLO2 h- 1	12.9	0.00253	20	67 30' N, 26 40' E	Terrestrial	
1,05E+09	μLO2 h- 1	12.9	0.00672	25	NA	Terrestrial	
0.000288972	μLO2 h- 1	43.7	0.0078	30	41°30N; 145°50E-44°00N; 155°00E	Aquatic	
0.009355331	μLO2 h- 1	21.3	0.0309	20	41 46 N 65 28' W	Aquatic	
0.009585727	J h-1	4.65	0.0138	22	NA	Terrestrial	

REF			EXP		
Openess.of.data..reusable.policy	Review.Doi	Original.Doi	Thermoregulatory	Salinity	Laboratory.Condition
https://www.journals.uchicago.edu/t-and-c paragraph 3.1.7	https://doi.org/10.1086/505997•	10.1016/0198-0149(90)90104-4	ectotherm	NA	NA
CC-BY	https://doi.org/10.1086/505997•	10.1016/0198-0149(90)90104-4	ectotherm	NA	No
https://onlinelibrary.wiley.com/library-info/resources/text-and-datamining	https://doi.org/10.1086/505997•	10.1016/0300-9629(73)90241-7	ectotherm	NA	NA
https://onlinelibrary.wiley.com/library-info/resources/text-and-datamining	https://doi.org/10.1086/505997•	10.1016/0300-9629(73)90241-7	ectotherm	NA	NA
https://onlinelibrary.wiley.com/library-info/resources/text-and-datamining	https://doi.org/10.21203/rs.3.rs-1092818/v1	10.1016/0198-0149(90)90104-4	ectotherm	10	none specified
Creative Commons Attribution 4	https://doi.org/10.21203/rs.3.rs-1092818/v1	https://doi.org/10.1016/0300-9629(75)90146-2	ectotherm	30	Yes
Creative Commons Attribution 4	https://doi.org/10.21203/rs.3.rs-1092818/v1	https://doi.org/10.1016/0300-9629(75)90146-2	ectotherm	20	Yes
Creative Commons Attribution 4.0	https://doi.org/10.21203/rs.3.rs-1092818/v1	https://doi.org/10.1016/0300-9629(75)90146-2	ectotherm	30	Yes

MAPPING OF COLUMNS HEADER

Original label	Label	Terminology used	ID
ID	catalogNumber	DwC	http://rs.tdwg.org/dwc/terms/catalogNumber
Kingdom	kingdom	DwC	http://rs.tdwg.org/dwc/terms/kingdom
Phylum	phylum	DwC	http://rs.tdwg.org/dwc/terms/phylum
Class	class	DwC	http://rs.tdwg.org/dwc/terms/class
Order	order	DwC	http://rs.tdwg.org/dwc/terms/order
Family	family	DwC	http://rs.tdwg.org/dwc/terms/family
Genus	genus	DwC	http://rs.tdwg.org/dwc/terms/genus
Species	scientificName	DwC	http://rs.tdwg.org/dwc/terms/scientificName
Life.stage	lifeStage	DwC	http://rs.tdwg.org/dwc/terms/lifeStage
SEX	sex	DwC	http://rs.tdwg.org/dwc/terms/sex
Harmonized_MR_watts	metabolic rate	Environmental Thesaurus	http://vocabs.lter-europe.net/EnvThes/21316
Originalg_unit_MR	Original Result Unit	NCI Thesaurus OBO Edition	http://purl.obolibrary.org/obo/NCIT_C82586
Harmonized_Body_mass_grams	body mass	Environmental Thesaurus	http://vocabs.lter-europe.net/EnvThes/21364
Dry.weight	Dry Mass	zooplankontraits	https://kos.lifewatch.eu/thesauri/zooplankontraits/c_19
Measurment_Temperature_C	temperatureValue	ABCD	http://rs.tdwg.org/abcd/terms/temperatureValue
Coordinates	verbatimCoordinates	DwC	http://rs.tdwg.org/dwc/terms/verbatimCoordinates
N/A	decimalLatitude	DwC	http://rs.tdwg.org/dwc/terms/decimalLatitude
N/A	decimalLongitude	DwC	http://rs.tdwg.org/dwc/terms/decimalLongitude
Ecosystem	habitat	Environmental Thesaurus	http://vocabs.lter-europe.net/EnvThes/21894
Openess.of.data..reusable.policy	license	DwC	http://purl.org/dc/terms/license
Review.Doi	Is Referenced By	DC	http://purl.org/dc/terms/isReferencedBy
Original.Doi	references	DC	http://purl.org/dc/terms/references
Thermoregulatory	body temperature regulation trait	Vertebrate trait ontology	http://purl.obolibrary.org/obo/VT_0001777
Salinity	Water body salinity	EMODnet Chemistry aggregated parameter names	http://vocab.nerc.ac.uk/collection/P35/current/EPC00001/
Laboratory.Condition	laboratory study	Radiation Biology Ontology	http://purl.obolibrary.org/obo/RBO_00002024


catalogNumber	kingdom	phylum	class	order	family	genus	scientificName	Full taxonomy		
1	Animalia	Arthropoda	Insecta	Coleoptera	Carabidae	Carabus	Carabus auratus			
2	Animalia	Arthropoda	Insecta	Coleoptera	Carabidae	Carabus	Carabus auratus			
3	Animalia	Arthropoda	Malacostraca	Isopoda	Oniscidae	Oniscus	Oniscus asellus			
4	Animalia	Arthropoda	Malacostraca	Isopoda	Oniscidae	Oniscus		Use of curated resources to describe variables (primary Darwin Core)		
5	Animalia	Arthropoda	Malacostraca	Isopoda	Oniscidae	Oniscus				
lifeStage	sex	metabolic rate	Original Result Unit	body mass	Dry Mass	temperatureValue	country	verbatimlocality		
adult	female	0,02617798	Watts	0,875	0,06365	5	NA	NA		
adult	male	0,07338301	ml-hour	0,564	0,06365	10	NA	NA		
NA	female	0,072122288	ml-hour	0,784	0,02038	15	Malaysia	Danum Valley Conservation Area, south-east Sabah		
NA	male	0,178103254	μLO2-hour	0,684	0,03758	20	Malaysia	Danum Valley Conservation Area, south-east Sabah		
NA	unsexed	0,016082891	μL O2-hour	0,967	0,06365	25	NA	NA		
fry	NA	0,009523393	verbatimCoordinates		decimalLatitude	decimalLongitude	habitat	license		
fry	NA	0,012314669	NA		NA	NA	Terrestrial	https://www.journals.uchicago.edu/t-and-c paragraph 3.1.7		
juvenile	unsexed	0,031881146	NA		NA	NA	Terrestrial	CC-BY 4.0		
			67 30' N, 26 40' E		67.53	26.66	Terrestrial	https://onlinelibrary.wiley.com/library-info/resources/text-and-datamining		
			67 30' N, 26 40' E		67.53	26.66	Terrestrial	https://onlinelibrary.wiley.com/library-info/resources/text-and-datamining		
			NA		NA	NA	Terrestrial	https://onlinelibrary.wiley.com/library-info/resources/text-and-datamining		
			41°30N; 145°50E-44°00N; 155°00E		41.76	-65.46	Aquatic	CC-BY 4.0		
			41 46 N 65 28' W		41.76	-65.46	Aquatic	CC-BY 4.0		
NA			Is Referenced By			references		body temperature regulation trait	Water body salinity	laboratory study
			https://doi.org/10.1086/505997			10.1016/0198-0149(90)90104-4		ectotherm	NA	NA
			https://doi.org/10.1086/505997			10.1016/0198-0149(90)90104-4		ectotherm	NA	No
			https://doi.org/10.1086/505997			10.1016/0300-9629(73)90241-7		ectotherm	NA	NA
			https://doi.org/10.1086/505997			10.1016/0300-9629(73)90241-7		ectotherm	NA	NA
						98-0149(90)90104-4		ectotherm	10	none specified
			https://doi.org/10.1016/0300-9629(75)90146-2			https://doi.org/10.1016/0300-9629(75)90146-2		ectotherm	30	Yes
			https://doi.org/10.21203/rs.3.rs-1092818/v1			https://doi.org/10.1016/0300-9629(75)90146-2		ectotherm	20	Yes
			https://doi.org/10.21203/rs.3.rs-1092818/v1			https://doi.org/10.1016/0300-9629(75)90146-2		ectotherm	30	Yes


1


Data file in tabular format

How we manage it now: semiautomatic annotation on data portal

UNIONE EUROPEA
Fondo Sociale Europeo
Fondo Europeo di Sviluppo Regionale


Ministero dell'Università
e della Ricerca


PON
RICERCA
E INNOVAZIONE
2014 - 2020


LifeWatch
ITALIA

Attribute Information

Variables

class

family

order

genus

providedscientificname

scientificname

totallength

ashfreedryweight

ashweight

Name

samplingprotocol

original description not harmonised

Label

samplingProtocol

mapped label

Definition

The names of, references to, or descriptions of the methods or protocols used during a dwc:Event.

Storage Type

Measurement Type

nominal

Measurement Domain

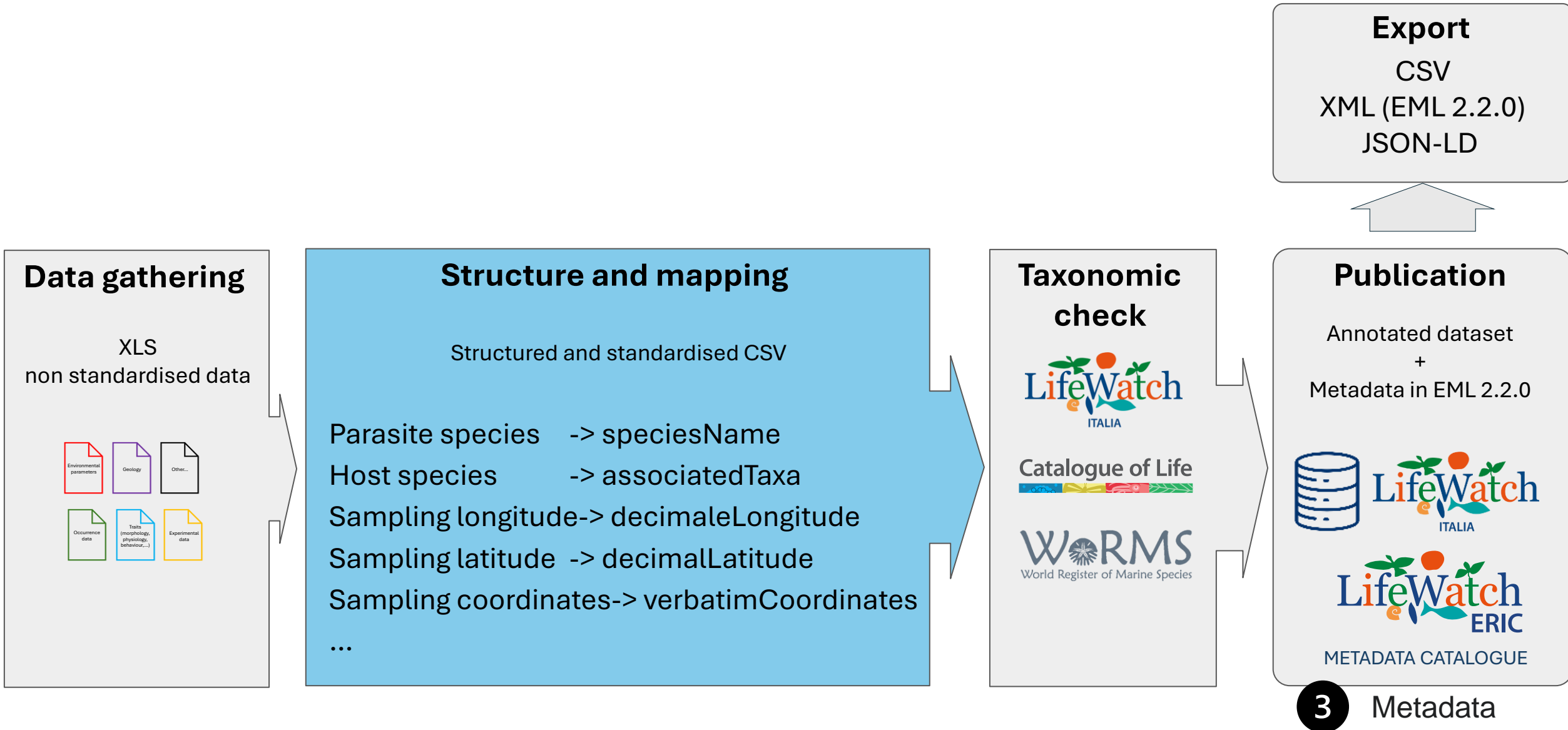
Definition	Pattern	Source
Any Text		

Annotation

Property Name	Property URI	Value	Value URI
http://ecoinformatics.org/oboe/oboe.1.2/oboe-core.owl#containsMeasurementsOfType	OBOE The Extensible Observation Ontology::contains measurements of type	http://rs.tdwg.org/dwc/terms/catalogNumber	catalogNumber

URI

Data harmonisation pipeline



Advantages of our approach



Integration with
semantic platform
and Data Labs



Extensible as user
like



Interoperable with
other systems



Export as Darwin
Core Archive



**Work in
progress**

Thank you 😊

Biodiversa+: Data management and capacity building on Darwin Core Standard workshop

Building Capacity: developing skill and expertise & creating community of practice



Cosimo Vallo, LifeWatch ERIC Training Officer

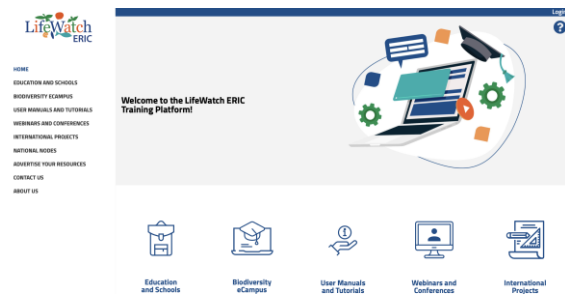
LW ERIC Training Architecture

Training Catalogue



Training,
Learning
and
Education

Training Platform

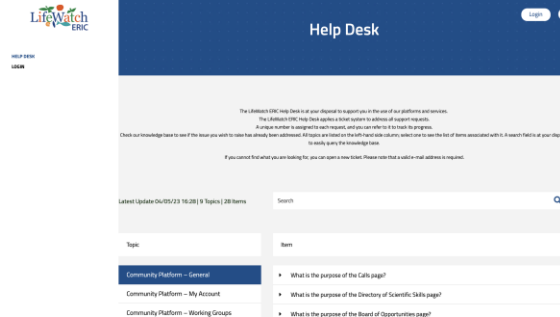


Community



Users
and
Community
Support

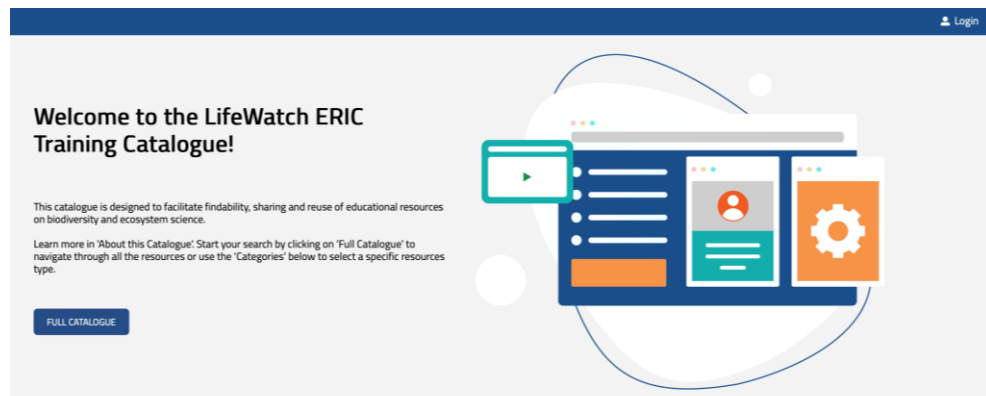
Helpdesk



Training Catalogue

LW ERIC Training Catalogue hosts the metadata of learning resources

- **Enhances FAIRness:** ensures that resources can be shared, searched, discovered, accessed and reused
- **Increases visibility:** offers learning resources produced internally in LW ERIC and hosted on its Training Platform, as well as externally by partner institutions/communities
- **Allows contributors** to publish learning resources



Categories



Education and
Games



Biodiversity
eCampus



Services
Tutorials



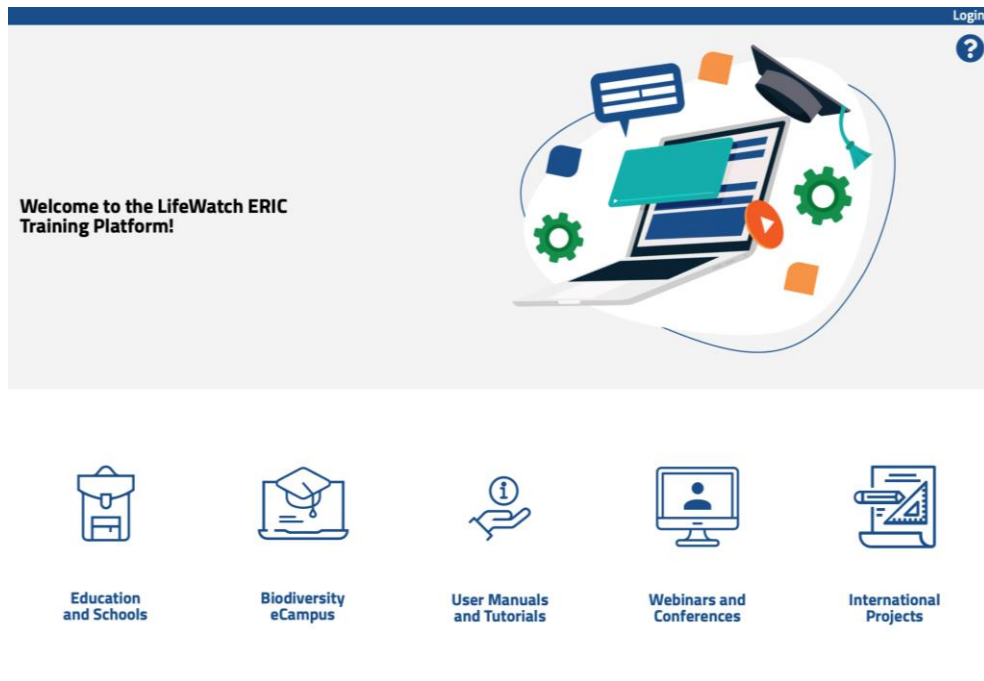
Webinars and
Conferences

<https://trainingcatalogue.lifewatch.eu/home/>

Training Platform

LW ERIC Training Platform addresses training, learning and education needs of the biodiversity research community

- **Training contents** constantly added and curated
- **User experience** is optimized with a user-friendly interface
- **Dedicated section(s)** can be developed to host project-specific contents (ex. EU projects)



<https://training.lifewatch.eu/>

Community

LW ERIC Community space dedicated to members of the wider scientific community

- **Meet, interact and collaborate:** interactive space for collaborative working and networking
- **Create Working Groups:** join researchers and experts working in the field of biodiversity and ecosystem research
- **Share, discover, find and promote** targeted opportunities and calls



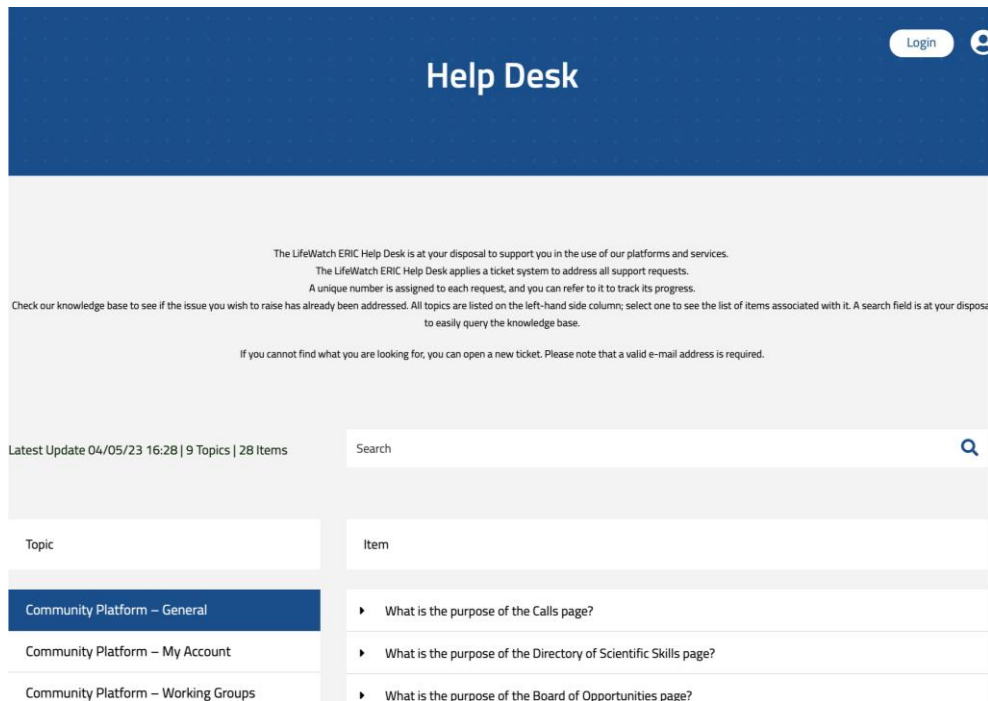
Working Groups

<https://community.lifewatch.eu>

Helpdesk

LW ERIC Helpdesk supports the use of our platforms and services

- **Ticket system** to address all support requests. Unique number to refer and track progress of your request
- **Open new ticket** to get support and solve issues as well as to suggest improvements to our products
- **Knowledge base** of all the issues raised with a search field to easily query the knowledge base



The screenshot shows the LifeWatch ERIC Help Desk interface. At the top, there is a dark blue header with the text "Help Desk" and a "Login" button. Below the header, a light gray box contains the following text:

The LifeWatch ERIC Help Desk is at your disposal to support you in the use of our platforms and services.
 The LifeWatch ERIC Help Desk applies a ticket system to address all support requests.
 A unique number is assigned to each request, and you can refer to it to track its progress.

Check our knowledge base to see if the issue you wish to raise has already been addressed. All topics are listed on the left-hand side column; select one to see the list of items associated with it. A search field is at your disposal to easily query the knowledge base.

If you cannot find what you are looking for, you can open a new ticket. Please note that a valid e-mail address is required.

Below this text, there is a search bar with the placeholder text "Search" and a magnifying glass icon. To the left of the search bar, it says "Latest Update 04/05/23 16:28 | 9 Topics | 28 Items".

The interface is divided into two main sections: "Topic" on the left and "Item" on the right. Under "Topic", there are three options: "Community Platform – General" (highlighted in dark blue), "Community Platform – My Account", and "Community Platform – Working Groups". Under "Item", there are three entries, each with a right-pointing arrow and a question:

- What is the purpose of the Calls page?
- What is the purpose of the Directory of Scientific Skills page?
- What is the purpose of the Board of Opportunities page?

<https://helpdesk.lifewatch.eu>

Meet the Team

Training Team at LifeWatch Service Center in Lecce



Cosimo Vallo

Training Officer



Eleonora Parisi

EU Projects Training Officer

Landing page: <https://www.lifewatch.eu/training-and-education/>

Training Action Coordination

LifeWatch ERIC created a **Training Working Group (TWG)** to collaborate among different offices and coordinate its training activities.

- **14 training focal points** from all its Common Facilities and National Distributed Centres that hold monthly coordination meetings
- **share, discuss, plan and design training** initiatives, activities and products
- members' engagement and information sharing contributes to **strengthen ERIC position in training within scientific community**
- Achievements: collaboratively developed Training Strategy, establishment of procedures and quality standards



Contact us

Get in touch with us at

training@lifewatch.eu



Biodiversa+: Data management and capacity building on Darwin Core Standard workshop

Building Capacity: developing skill and expertise & creating community of practice



Thank you for the attention!