LifeWatch ERIC
e-Science European Infrastructure
for Biodiversity and Ecosystem Research

The key role of semantics for data integration and interoperability in biodiversity and ecosystems research

Caterina Bergami, Bachir Balech, Angela Boggero, Fabio Cianferoni, Paolo Colangelo, Stefano De Felici, Nicola Fiore, Alessandro Oggioni, Cataldo Pierri, Leonilde Roselli, Elena Stanca, Paolo Tagliolato, Ilaria Rosati

LifeWatch ITALIA

Conferenza Annuale di LifeWatch Italia 2018
25-27 giugno 2018 Roma

Context & Motivation

Large amounts of relevant data sources…….

…..highly *heterogeneous* (structural and semantic differences)

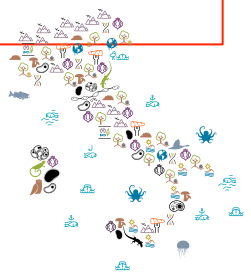
Heterogeneity impedes:

- **Discovery**
- **Integration**
- **Re-usability**

What do we need?

- **Harmonization**
- **Standardization**

for sharing information and revealing its full potential



Solution

Semantic Resources

Semantics enriches resources with additional information attaching machine-processable and readable meaning through…..

Controlled vocabularies

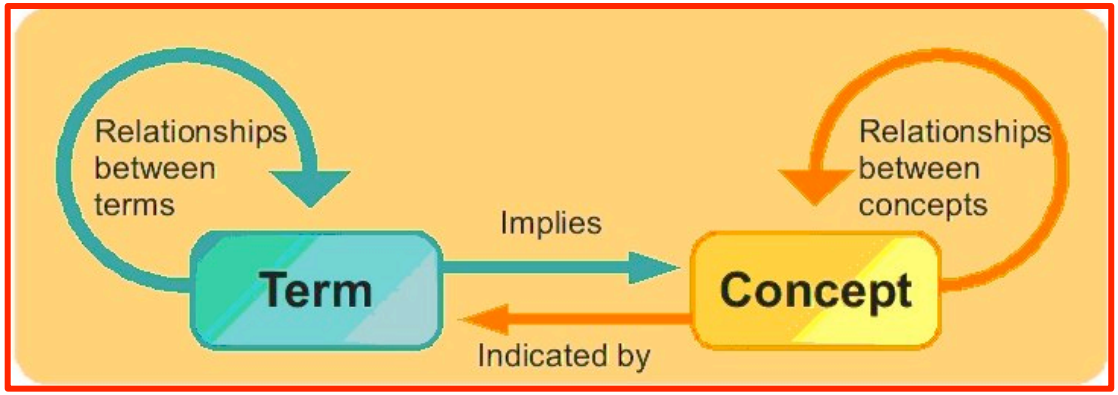
Thesauri

Ontologies

Thesaurus

What is?

…a controlled and structured vocabulary in which concepts are represented by terms, organised so that relationships between concepts are made explicit, (International Organization for Standardization 2011, 2013)

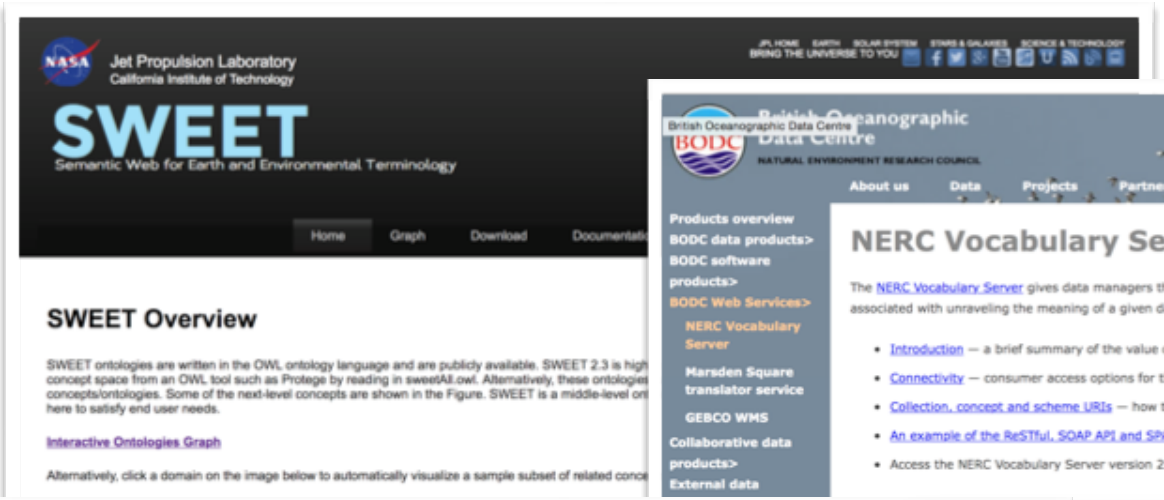


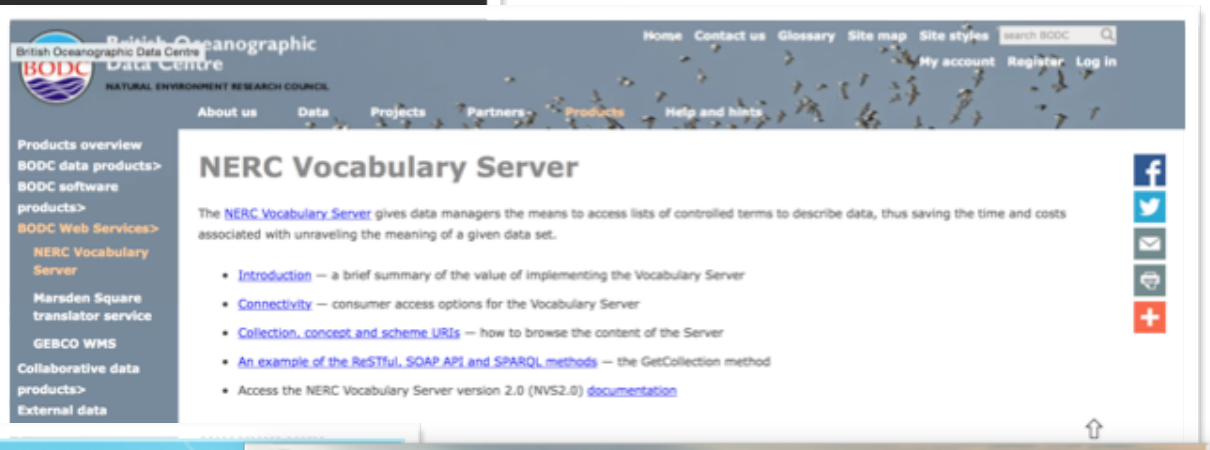
Why is useful?

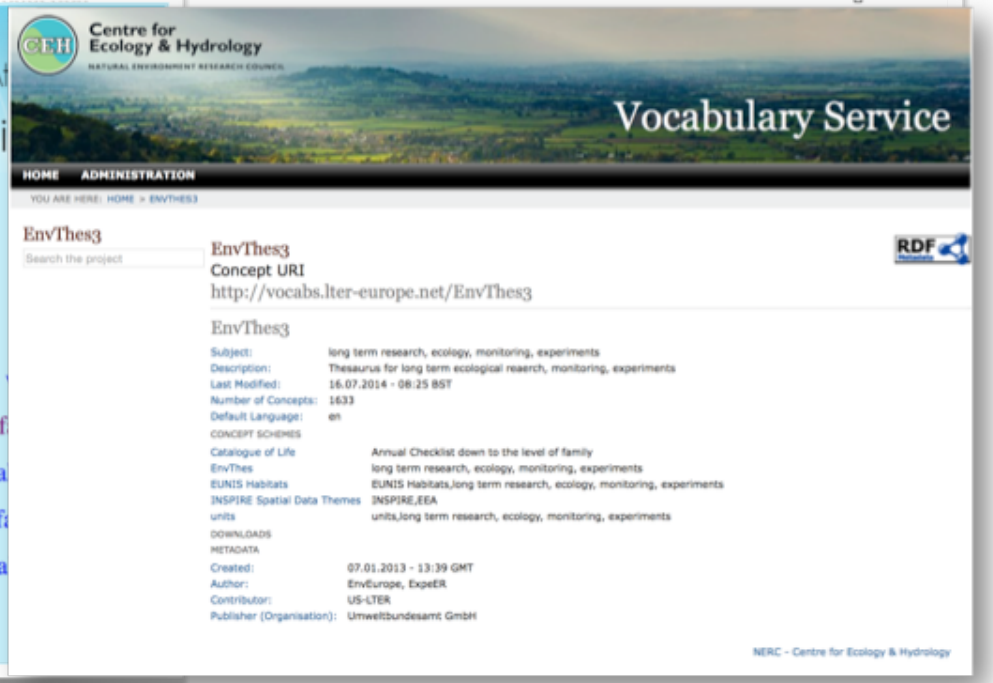
…can set the standards for all information on a particular concept

- allowing for easier search and retrieval of information about a particular domain
- supporting the indexing, retrieval, organisation and navigation of information

Some example…

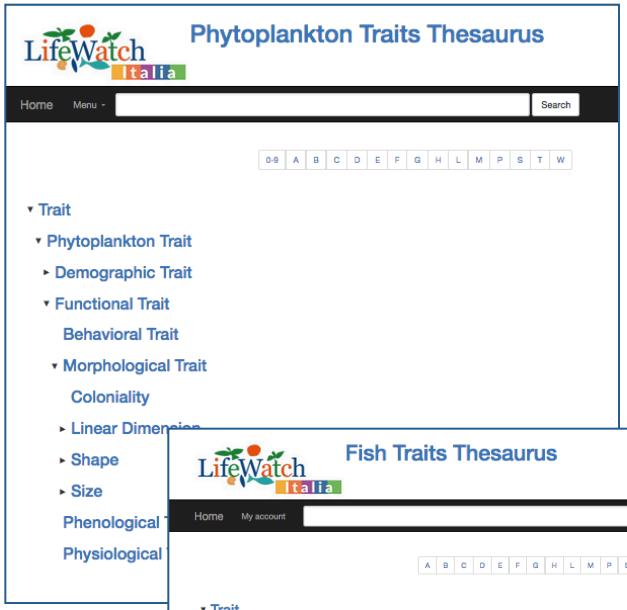


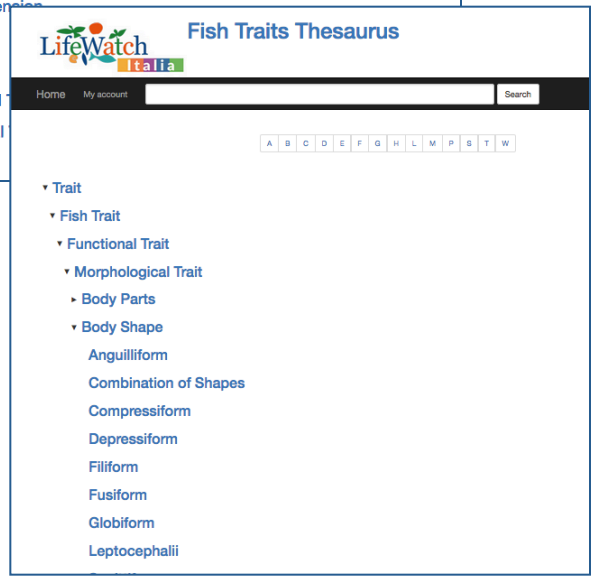




LifeWatch Thesauri

http://thesauri.lifewatchitaly.eu/PhytoTraits/index.php





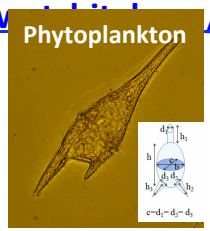
http://thesauri.lifewatchitaly.eu/fishtraits/index.php

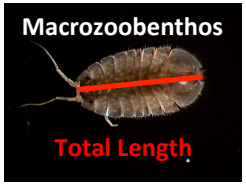
Functional Traits Thesauri

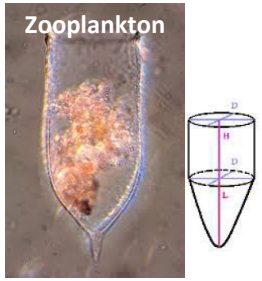
http://thesauri.lifewatchitaly.eu/ZooplanktonTraits/index.php

http://thesauri.lifewatchitaly.eu/macrozoobenthostraits/index.php

http://thesauri.lifev[illegible]/Macroalgae/index.php







Other Thesauri

http://thesauri.lifewatchitaly.eu/AquaticOrganisms/index.php

http://thesauri.lifewatchitaly.eu/alienspecies/index.php

http://thesauri.lifewatchitaly.eu/endemisms/index.php

http://thesauri.lifewatchitaly.eu/GenomicBarcoding/index.php

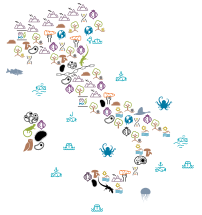
About LifeWatch Thesauri

They are……

❑ A community effort

Developed and managed through a collaborative process involving different working groups with specific roles.

	Planning and Design	Terms research and selection	Formalization phase	Edition phase	Validation phase
 Editors	State of the art review Plan the objectives and features of the thesaurus	Gather terms, choose preferred terms and non preferred terms Provide a first hierarchical structure and a definition	Identify the thesaurus format and the editing tool Provide a first electronic version	Modify and enrich terms, their attributes, hierarchical and equivalence relationships	Submit the thesaurus to the validators Decide to approve or not the validators' annotations Release a new stable version of the thesaurus
 ICT	Research of editing and mapping tools Thesaurus modelling		Identify the thesaurus format and the editing tool Install software and configure thesaurus Provide a first electronic version	Support in technical questions	Support in technical questions
 Validators					Review terms and their structural elements Propose modifications

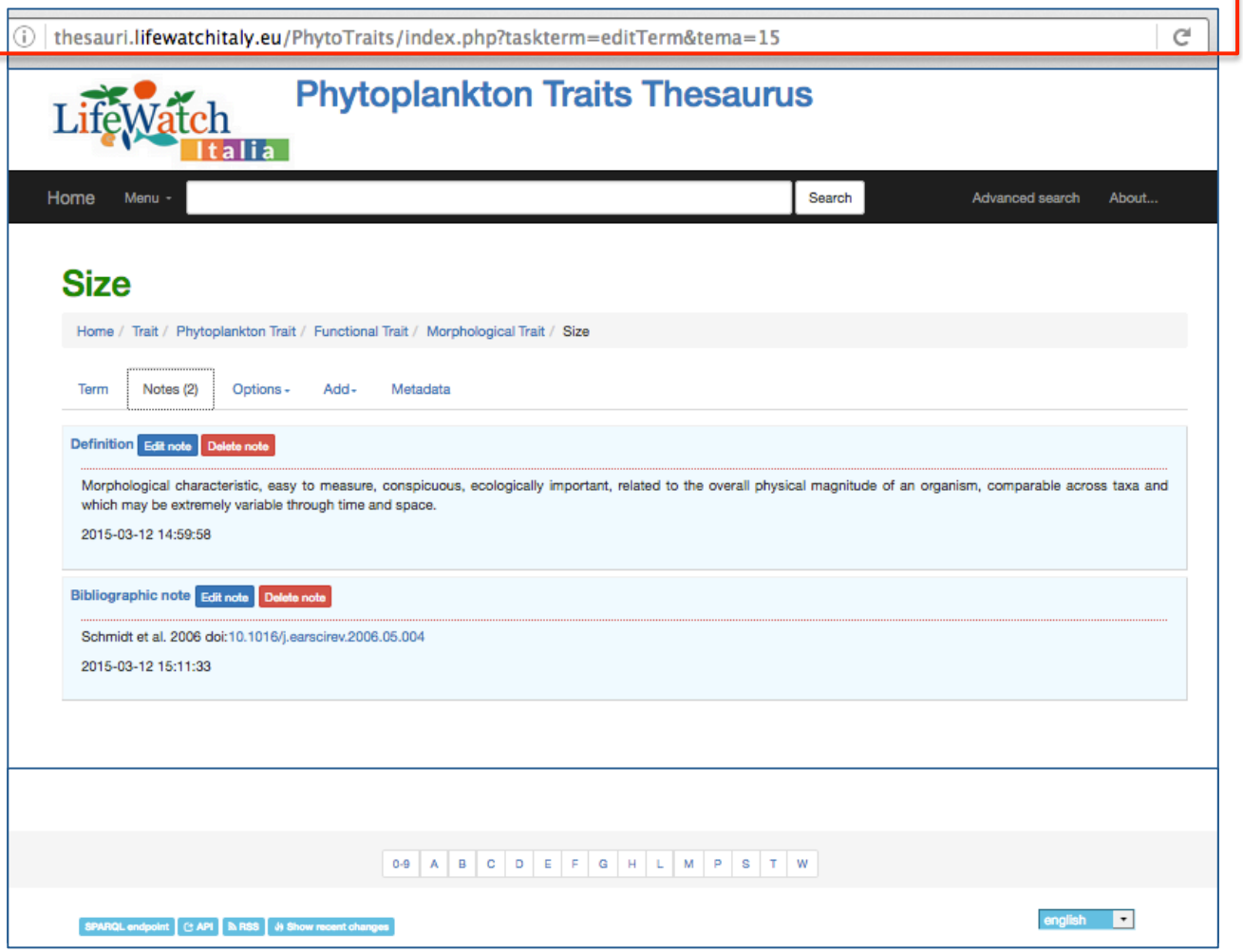


About LifeWatch Thesauri

- **A stable reference resource**

SKOS Data Model **Concept is the fundamental entity** → **Attributes:**

- URI
- Preferred label
- Alternative label
- Definition
- Scope note
- Bibliographic note
- Has broader
- Has narrower
- Has exact match
- Has close match
- …
- Publication date



TemaTres Editor

About LifeWatch Thesauri

- **A stable reference resource**

TemaTres Editor 

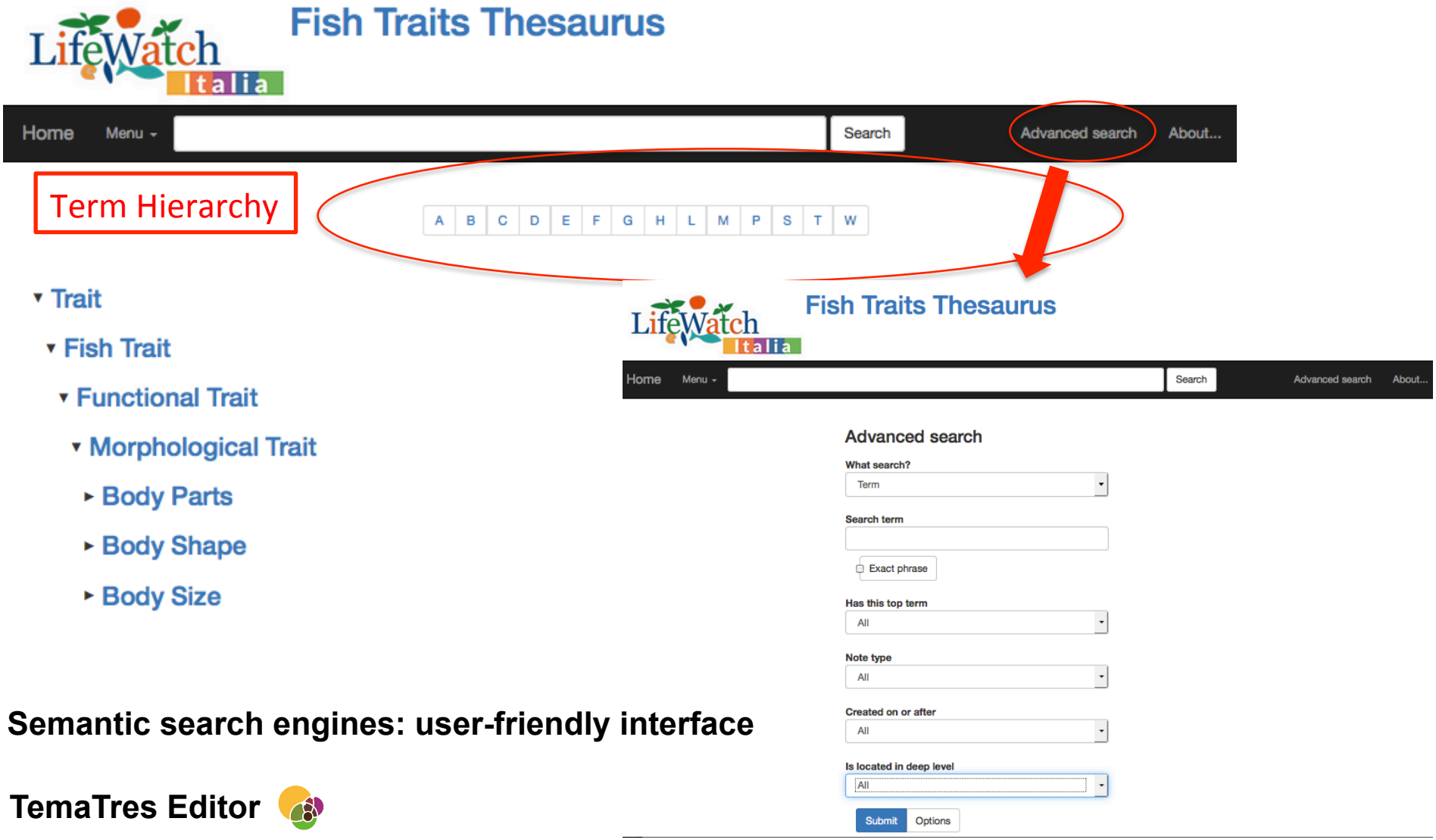
Machine readable **Discoverable** **Interoperable**

```
<rdf:RDF>
<skos:ConceptScheme rdf:about=http://thesauri.lifewatchitaly.eu/PhytoTraits/ >
dc:title>Phytoplankton Traits Thesaurus</dc:title>
<dc:creator>LifeWatch Italy</dc:creator>
<dc:date>2015-03-06</dc:date>
<dct:modified>2016-04-20 12:40:05</dct:modified>
<dc:language>en</dc:language>
<skos:Concept rdf:about=http://thesauri.lifewatchitaly.eu/PhytoTraits/?tema=15 >
<skos:prefLabel xml:lang="en">Size</skos:prefLabel>
<skos:definition xml:lang="en">
Morphological characteristic, easy to measure, conspicuous, ecologically important, related to the overall physical magnitude of an
organism, comparable across taxa and which may be extremely variable through time and space.
</skos:definition>
<skos:note xml:lang="en >
Schmidt et al. 2006 doi:10.1016/j.earscirev.2006.05.004
</skos:note>
<skos:inScheme rdf:resource=http://thesauri.lifewatchitaly.eu/PhytoTraits/>
<skos:broader rdf:resource=http://thesauri.lifewatchitaly.eu/PhytoTraits/?tema=13/>
<skos:narrower rdf:resource=http://thesauri.lifewatchitaly.eu/PhytoTraits/?tema=17/>
<skos:narrower rdf:resource=http://thesauri.lifewatchitaly.eu/PhytoTraits/?tema=20/>
<skos:narrower rdf:resource=http://thesauri.lifewatchitaly.eu/PhytoTraits/?tema=18/>
<dct:created>2015-03-07 07:00:23</dct:created>
</skos:Concept>
</rdf:RDF>
```

About LifeWatch Thesauri

❑**Free and open (http://www.servicecentrelifewatch.eu/catalogue-of-services)**

LifeWatch thesauri are available as a web service for ecological community in order to make data interoperable between different research groups.



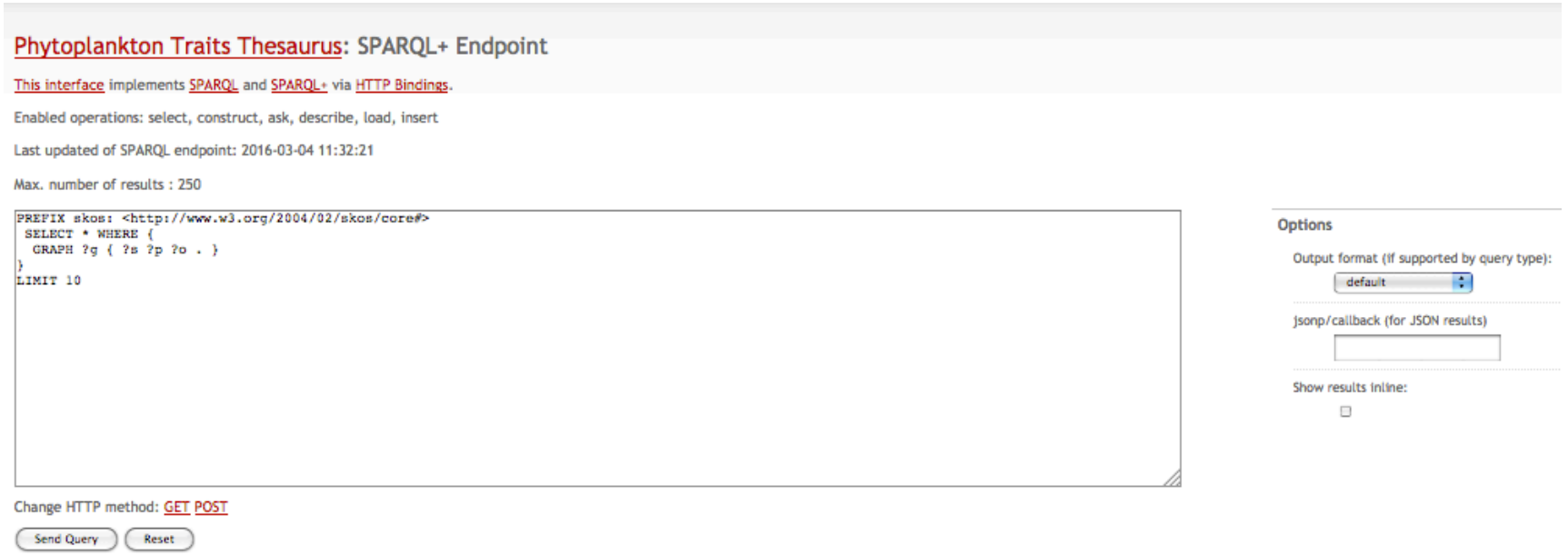
Semantic search engines: user-friendly interface

TemaTres Editor

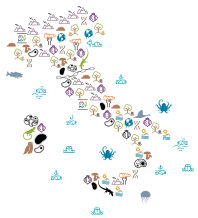
About LifeWatch Thesauri

- Can be queried via web endpoints (SPARQL and API)

http://thesauri.lifewatchitaly.eu/PhytoTraits/sparql.php



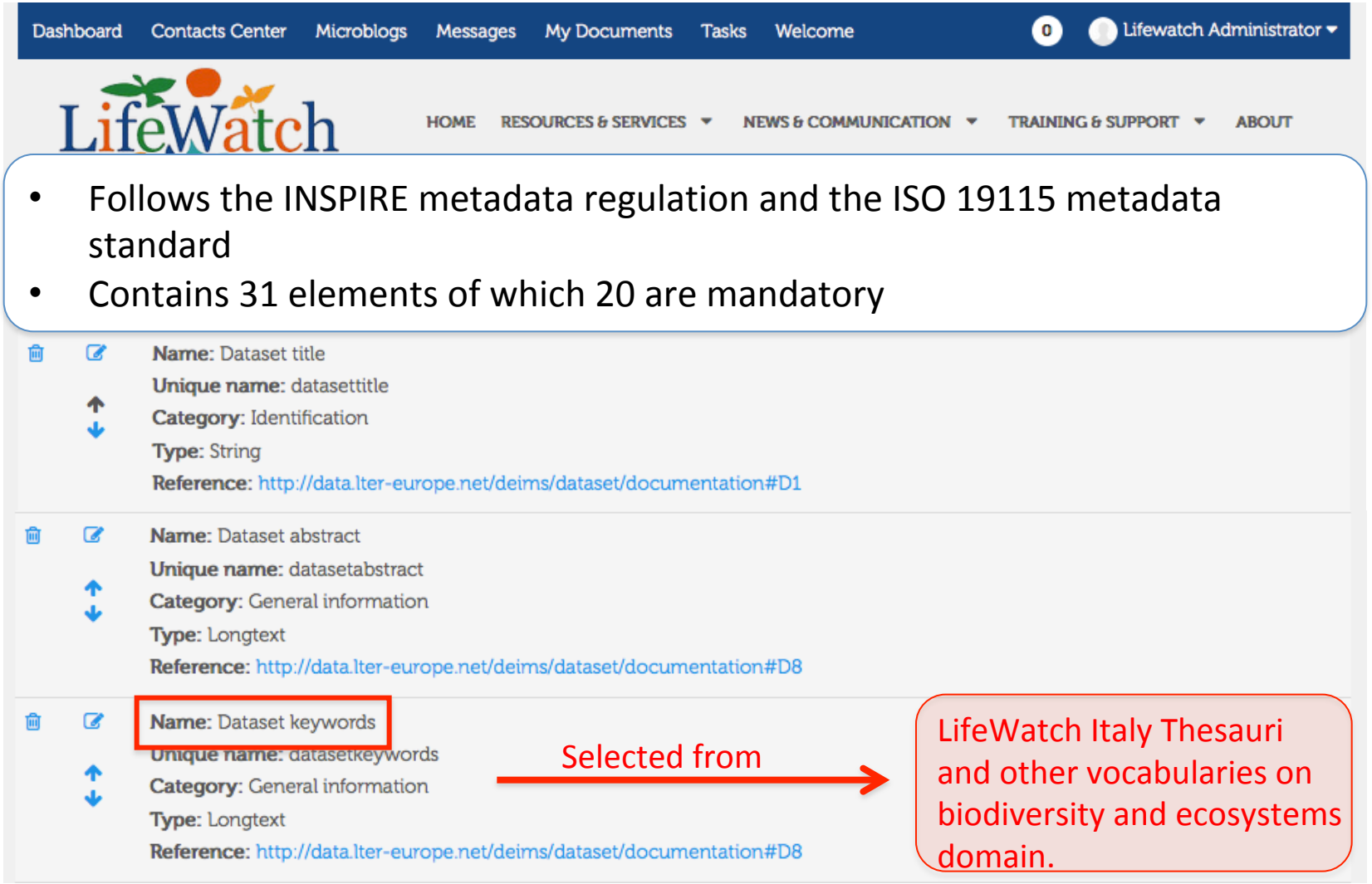
http://thesauri.lifewatchitaly.eu/PhytoTraits/services.php



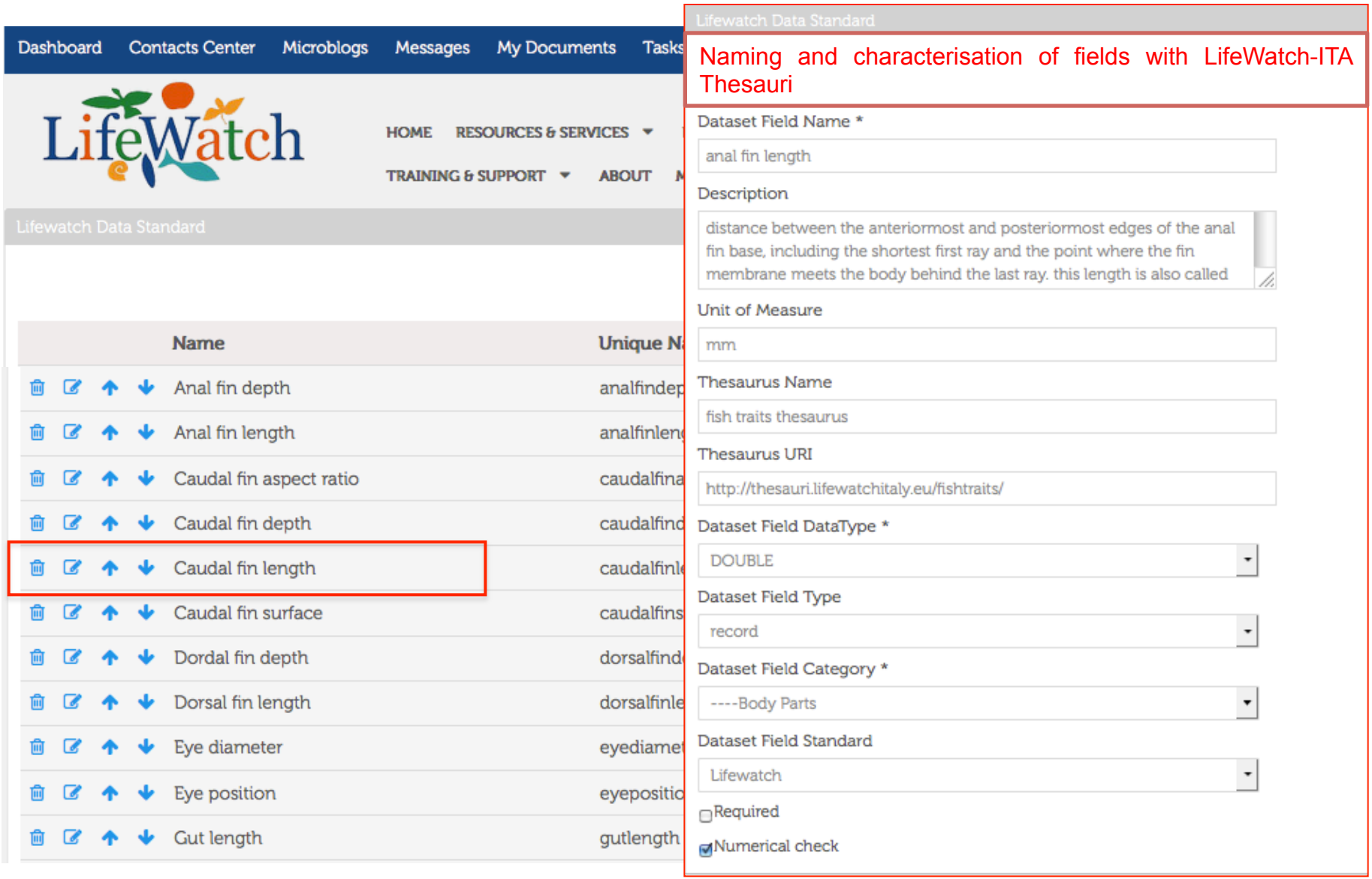
How and Where we use them?

- ✓ Reference in documents (if you enter the URL of a concept you can follow it like a hyperlink)
- ✓ Metadata and data schema definition in the LifeWatch Data Portal (http://www.servicecentrelifewatch.eu/catalogue-of-resources) in order to support semantic interoperability;
- ✓ as building blocks for a LifeWatch core ontology allowing complex searches/analysis on data sources (see Fiore et al. this session)

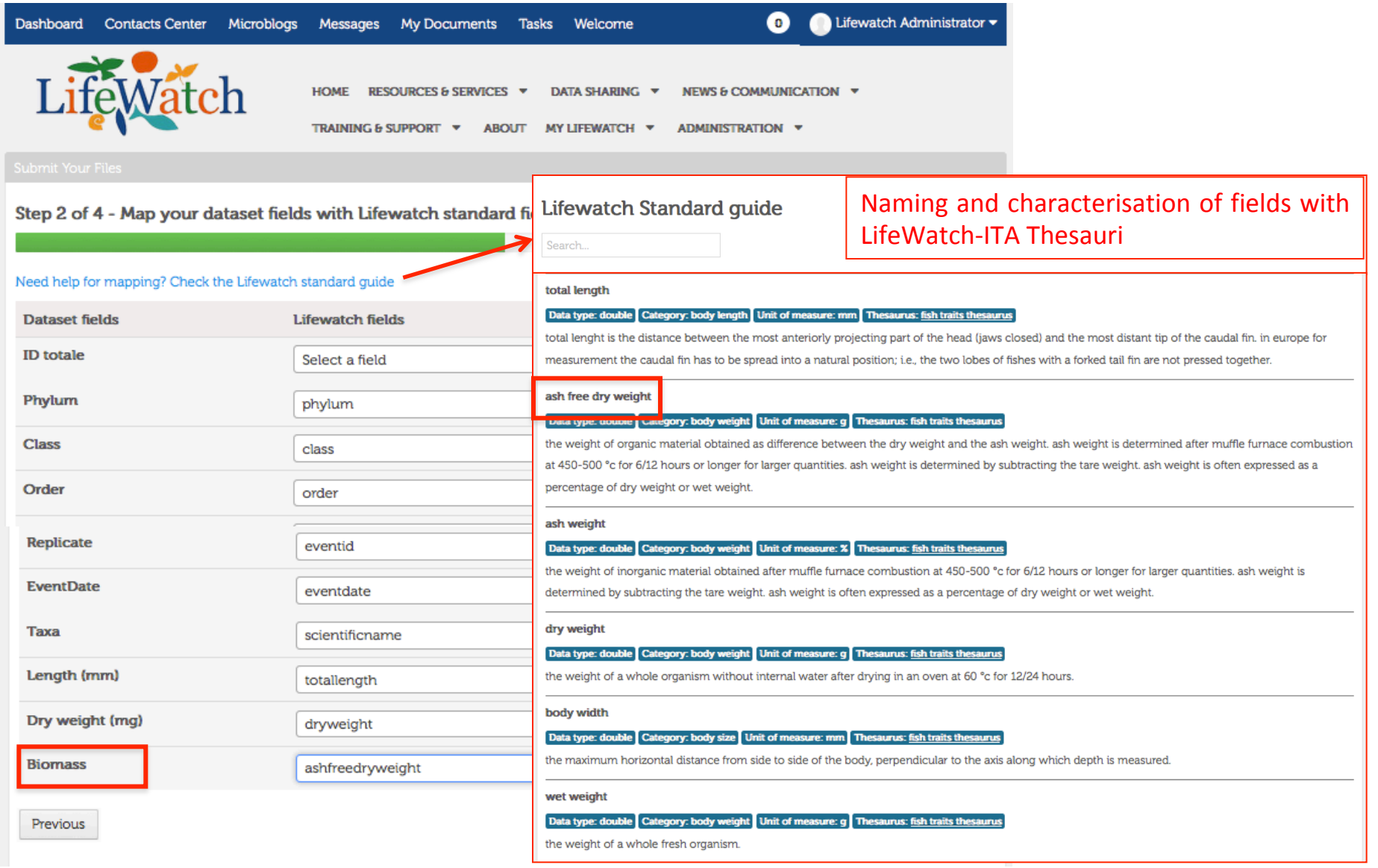
LifeWatch Italy Metadata Schema



LifeWatch Italy Data Schema



LifeWatch Italy Data Schema



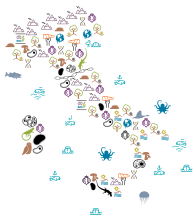
Next steps…

- Development of a platform for managing semantic resources
  - Extension of thesauri (other concepts and multilingualism)
  - Mapping and Alignment with other Thesauri

In this context, experts in the specific domain had a central role in the different phases of thesauri implementation, allowing to produce shared and stable versions of thesauri which could become common standards for data management on biodiversity and ecosystems domain.

Experts from several institution of the LW-ITA JRU have already contribute to the implementation of LW-ITA thesauri.
Big working groups => more stable and shared SR

To collaborate: lifewatchitalia@unisalento.it




grazie
merci
arigatô
danke
gracias
thank you
dankie
shukran
dank u
благодаря
obrigado
hvala
takk
tak
tänan
faleminderit
kiitos
paldies
dhanyavād
xièxiè
grazzi
tack
diolch
köszönöm
terima kasih
mahalo
mh'gōi
dziękuję
ďakujem
děkuji
ačiū
gràcies
баярлалаа
дякую
спасибо
хвала
choukrane
nandri
ngiyabonga
kop khun
teşekkür ederim
mulţumesc
благодарам
a dank