## **Taxonomic Backbone for LifeWatch**

Services and databases for species taxonomy, distribution and traits

Klaas Deneudt, Stefanie Dekeyzer, Bart Vanhoorne, Filip Waumans, Simon Claus, Leen Vandepitte, Francisco Hernandez



Vlaams Instituut voor de Zee vzw Flanders Marine Institute

## Biodiversity research



Van veen grab at x,y,t Observation:

X,Y

Depth: 20m

Sediment: sand

Analysis:

- What species do we expect to find?
- What species are absent?
- What are possibly wrong identifications?
- What are common, what are rare findings?

Species

Amphiodia atra (Stimpson, 1852) Hemipholis elongata (Say, 1825) Chione sp.

Corbula caribea (Orbigny, 1842) Corbula sp.

Cyclinella tenuis (Rècluz, 1852)

Lioberus castaneus (Say, 1822)

Nucula semiornata (Orbigny, 1846)

Periploma compressa (Orbigny, 1846)

Temnoconcha brasiliana (Dall, 1921)

Thysanocardia catharinae (Grube, 1868) Branchiostoma platae (Hubbs, 1922)



#### **Options:**

- Look into existing literature
- Look for existing data
- Check with experts

## What is LifeWatch?



#### **VIRTUAL LABORATORY**

LifeWatch is a distributed **virtual laboratory** and will be used for biodiversity research, for climatological and environmental impact studies, to support the development of ecosystem services and to provide information for policy makers in Europe. This large European research infrastructure will consist of several biodiversity observatories, databases, web services and modeling tools. It will be integrating the existing systems, upgrading them were possible and developing new systems where needed.

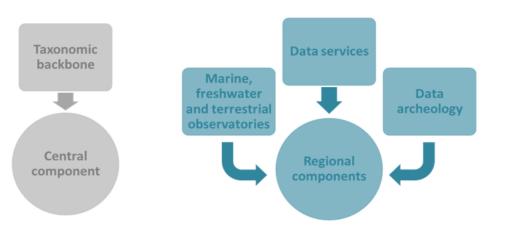
#### PARTICIPATING COUNTRIES

# Construction phase

#### **FLEMISH CONTRIBUTIONS**

The Flemish contributions to LifeWatch are coordinated by the Flanders Marine Institute (VLIZ) and the Research Institute for Nature and Forest (INBO). The Flemish LifeWatch consortium is funded by the Hercules Foundation.





#### LIFEWATCH PHASES

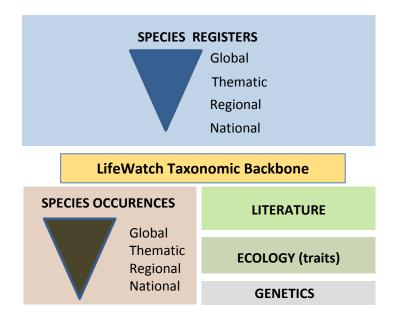


What is the LifeWatch Taxonomic Backbone?



## **Taxonomic Backbone**

- Central contribution to LifeWatch
- Facilitates the standardisation of species information
- (Virtually) brings together different component databases & data systems
- 5 major branches





## Biodiversity research



Van veen grab at x,y,t Observation data:

X,Y

Date

Depth: 20m

Sediment: sand

What species do we expect to find?

What species are absent?

What are possibly wrong identifications?

What are common, what are rare findings?

Species

Amphiodia atra (Stimpson, 1852) Hemipholis elongata (Say, 1825) Chione sp.

Corbula caribea (Orbigny, 1842) Corbula sp.

Cyclinella tenuis (Rècluz, 1852) Lioberus castaneus (Say, 1822) Nucula semiornata (Orbigny, 1846) Periploma compressa (Orbigny, 1846) Temnoconcha brasiliana (Dall, 1921) Thysanocardia catharinae (Grube, 1868)

Branchiostoma platae (Hubbs, 1922)



#### **Options:**

Check with LifeWatch Taxonomic Backbone



Source	Analysis	Result			
Observation	identify species take position	P{} observed species x,y			
Marine regions	region of x,y	Westhinder bank North sea			
OBIS occurence data WoRMS distribution data LW ENM data	species in North sea	S{}			
TB – Habitat Traits	S{Species in benthos}	Sb{} benthic species			
TB – Size Traits	$Sb{X < size < Y}$	Sm{} macrobenthic species			
Observation	take depth describe substrate	20m soft sediment			
TB – Environment Traits	Sm{Species at 20m, soft sediment}	E{} expected species			
	$E\{\} - P\{\}$ $P\{\} - E\{\}$	A{} absent species E{} possible errors U{} unique observations			

## **Taxonomic Backbone**

#### Example questions:

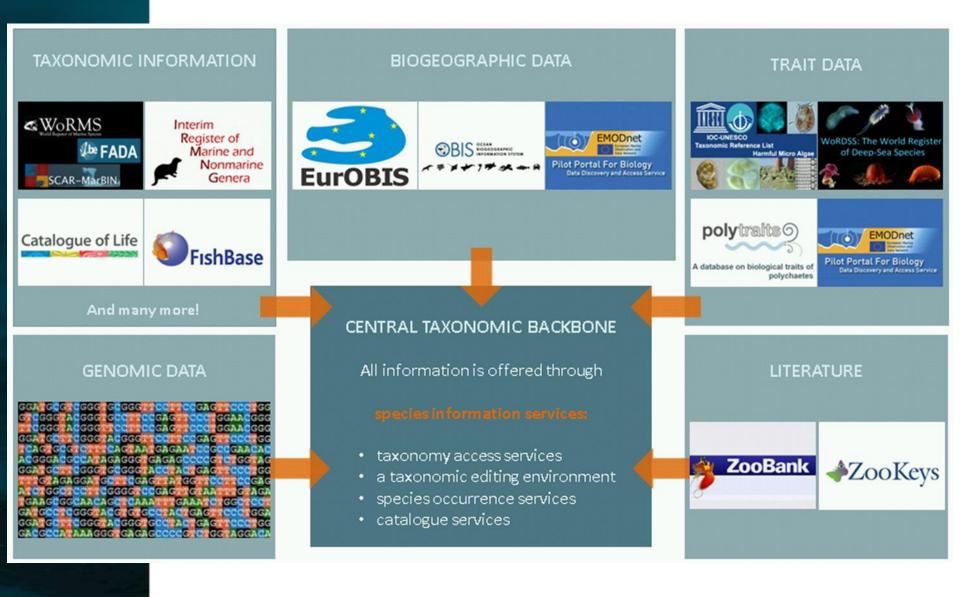
- I'm planning to take a Van Veen benthic grab at a particular location in the North Sea. What species can I expect to capture?
- Which invasive planktonic species are known to occur in the Black Sea?
- Where does species 'X' appear?
- Which species from the Habitat/Bird Directive are on the IUCN Red List?



How do we gather the required information?



## 1) Integrate existing databases

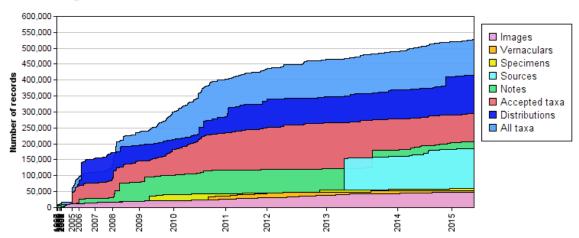


## 2) Complete & update species data



- Support by Data Management Team
- Technical developments

#### Database growth





## 3) Organize & mobilize (taxonomic) experts











## **Taxonomic backbone: Species names**













WoRMS Currently holds: 229,408 accepted species, of which 96% checked (more stats)



















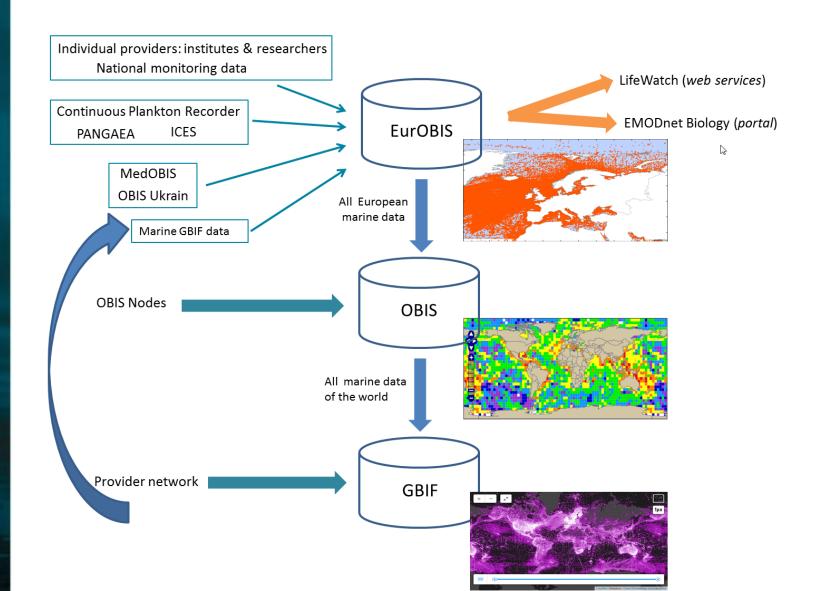








## **Taxonomic backbone: species distribution**





### Taxonomic backbone: species traits



#### fossil range Classificatio Importa Habitat Distribution

**Traits Attribut** 

Enter the attribute you

Valid wildcards are '%'

taxa belonging to:  $\frac{1}{e.c}$ 

Search

Search: F

Taxon rank: In

Limit to

Website

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE RDF>
     <rdf:RDF xmins:property="http://coastalwiki.org/traits/wiki/Special:URIResolver/Property-3A" xmlns:wiki="http://coastalwiki.org/traits/wiki/Special:URIResolver/" xmlns:swivt="http://semantic</pre>
 xmins:owl="http://www.w3.org/2002/07/owl#" xmins:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmins:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
xmins:owl="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
xmins:owl="htt
```

<swivt:creationDate rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2015-05-20T15:15:01+02:00</swivt:creationDate> <owl:imports rdf:resource="http://semantic-mediawiki.org/swivt/1.0"/> </owl:Ontology>

<swivt:Subject rdf:about="http://coastalwiki.org/traits/wiki/Special:URIResolver/Traits-3AAbsent"> <rdfs:label>Traits:Absent/rdfs:label>

<swivt:page rdf:resource="http://www.marinespecies.org/traits/wiki/Traits:Absent"/>

<rdfs:isDefinedBy rdf:resource="http://www.marinespecies.org/traits/wiki/Special:ExportRDF/Traits:Absent"/>
<swivt:wikiNamespace rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">200</swivt:wikiNamespace>

<swivt:wikiPageModificationDate rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2015-05-06T11:25:24Z</swivt:wikiPageModificationDate>

<swivt:specialProperty\_uri rdf:resource="http://www.marinespecies.org/traits/Absent"/> 

<swivt:Subject rdf:about="http://coastalwiki.org/traits/wiki/Special:URIResolver/Traits-3AAbundance">

<rdfs:label>Traits:Abundance</rdfs:label>

<swivt:page rdf:resource="http://www.marinespecies.org/traits/wiki/Traits:Abundance"/> <rdfs:isDefinedBy rdf:resource="http://www.marinespecies.org/traits/wiki/Special:ExportRDF/Traits:Abundance"/>

<swivt:wikiNamespace rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">200</swivt:wikiNamespace>

<swivt:wikiPageModificationDate rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2015-05-06T12:07:34Z</swivt:wikiPageModificationDate</p>

<swivt:specialProperty\_uri rdf:resource="http://www.marinespecies.org/traits/Abundance"/> 

property:Has\_value rdf:resource="http://coastalwiki.org/traits/wiki/Special:URIResolver/Traits-3ACommon\_to\_dominant"/>

property:Has\_value rdf:resource="http://coastalwiki.org/traits/wiki/Special:URIResolver/Traits-3AFluctuating"/> 

</swivt:Subject>

<swivt:Subject rdf:about="http://coastalwiki.org/traits/wiki/Special:URIResolver/Traits-3AAbyssobenthic">

<rdfs:label>Traits:Abyssobenthic</rdfs:label

<swivt:page rdf:resource="http://www.marinespecies.org/traits/wiki/Traits:Abyssobenthic"/ <rd><rdfs:isDefinedBy rdf:resource="http://www.marinespecies.org/traits/wiki/Special:ExportRDF/Traits:Abyssobenthic"/></rd>

<swivt:wikiNamespace rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">200</swivt:wikiNamespace>

<swivt:wikiPageModificationDate rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2015-05-06T14:23:32Z</swivt:wikiPageModificationDate>

<swivt:specialProperty\_uri rdf:resource="http://www.marinespecies.org/traits/Abyssobenthic"/>

"http://www.w3.org/2001/XMLSchema#string">Occupying the ocean floor from ca 4000 - 6000 m depth. Usually a more or less flat plain. (Lincoln et al., 19 

- <swivt:Subject rdf:about="http://coastalwiki.org/traits/wiki/Special:URIResolver/Traits-3AAbyssopelagic">

grazer > interface grazer herbivore

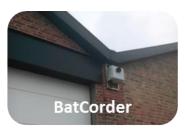


# **Gathering new observation data**-> LW observatory



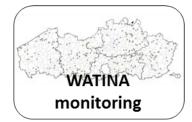
































How can you use the taxonomic backbone?



**Web services** = systems that allow communication between several computers over the web

⇒ allows integration of services and tools into own scripts or applications

Example: <a href="http://geo.vliz.be/geoserver/MarineRegions/wms?service=WMS&">http://geo.vliz.be/geoserver/MarineRegions/wms?service=WMS&</a> version=1.1.0&request=GetMap&layers=MarineRegions:eez&styles=& bbox=-180.000015258789,-85.3755950927734,180.000015258789,87.023948 6694336&width=689&height=330&srs=EPSG:4326&format=application/openlayers

http://geo.vliz.be/geoserver/MarineRegions/wms?service=WMS

version=1.1.0

request=GetMap

layers=MarineRegions:eez

styles=

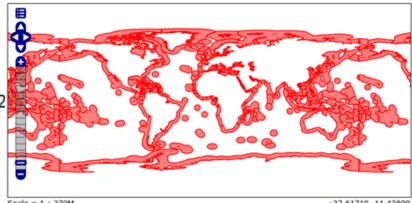
bbox=-180.000015258789,-85.3755950927734,180.0000152

width=689

height=330

srs=EPSG:4326

format=application/openlayers



Click on the map to get feature info

-37.61719, 11.43800

**Workflows** = the concatenated use of the data services

=> output first data service is input next data service





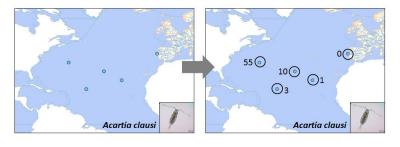
4 example workflows:





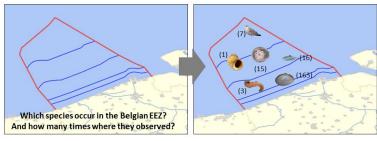
#### **USE CASE 1:**

Marine species observations in a 1000m radius around your own observation points



#### **USE CASE 2:**

Marine species list and number of observations per geographical area



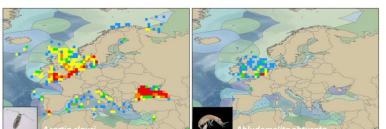
LIFEWATCH



DATA FORMAT CHECK: Missing mandatory fields?

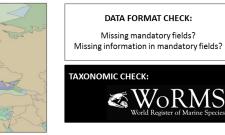
#### **USE CASE 3:**

List of geographical areas per marine species



#### **USE CASE 4:**

**Quality control of biodiversity datasets** 







USE-CASE 2: Extract a list of species for a specific region, including the number of observations for that region

INPUT = place name(s)

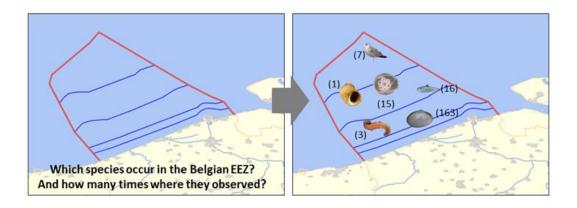
Service 1: "Translation" of place name to coordinates/bounding box

Service 2: Coordinates fed to OBIS

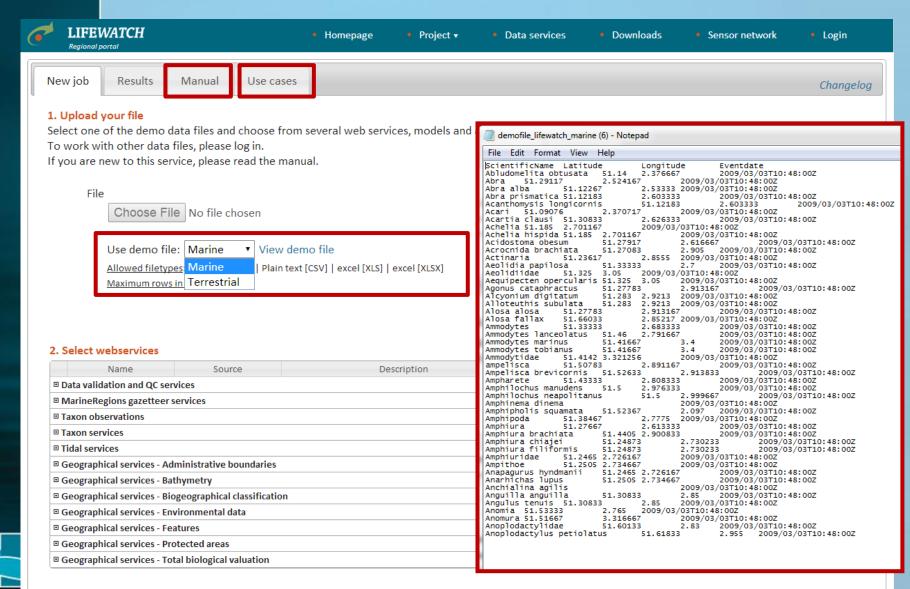
⇒ extract of species list for those coordinates/bounding box

⇒ Calculation of number of observations per species for those coordinates

OUTPUT = list of species with number of available observations in OBIS







#### 3. Verify order, change order if necessary and run

Selected services



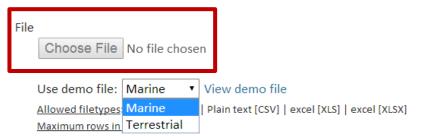


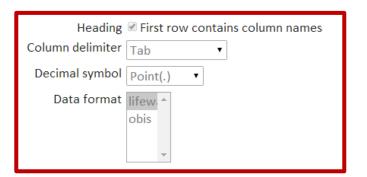
#### 1. Upload your file

Select one of the demo data files and choose from several web services, models and applications to process the data.

To work with other data files, please log in.

If you are new to this service, please read the manual.





#### 2. Select webservices

Name	Source	Description	Environment	Status
■ Data validation and QC serv	vices			
■ MarineRegions gazetteer s	ervices			
■ Taxon observations				
■ Taxon services				
■ Tidal services				
☐ Geographical services - Ada ☐	ministrative boundaries			
☐ Geographical services - Bat ☐	hymetry			
☐ Geographical services - Bio ☐	geographical classificatio	n		
☐ Geographical services - Env ☐	vironmental data			
■ Geographical services - Fea	atures			
☐ Geographical services - Pro ☐	tected areas			
■ Geographical services - Tot	al biological valuation			

#### 3. Verify order, change order if necessary and run

Selected services





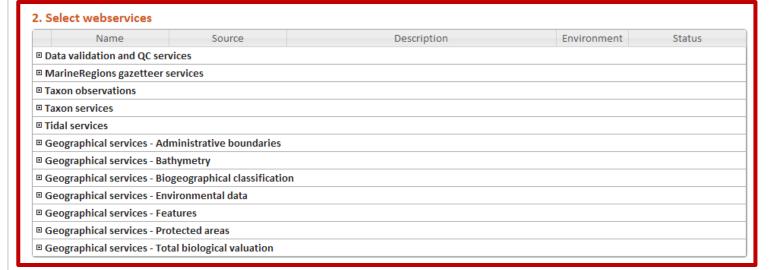
#### 1. Upload your file

Select one of the demo data files and choose from several web services, models and applications to process the data.

To work with other data files, please log in.

If you are new to this service, please read the manual.





#### 3. Verify order, change order if necessary and run

Selected services



3. Verify order, change order if necessary and run

Selected services

Get lat-long by name

Homepage

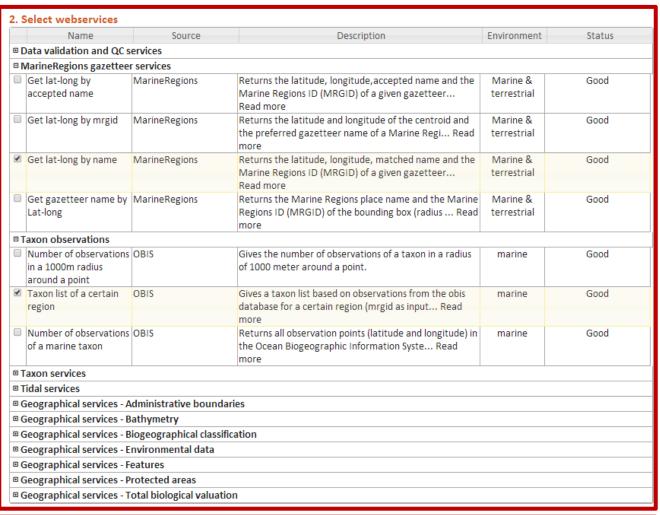
Project ▼

Data services

Downloads

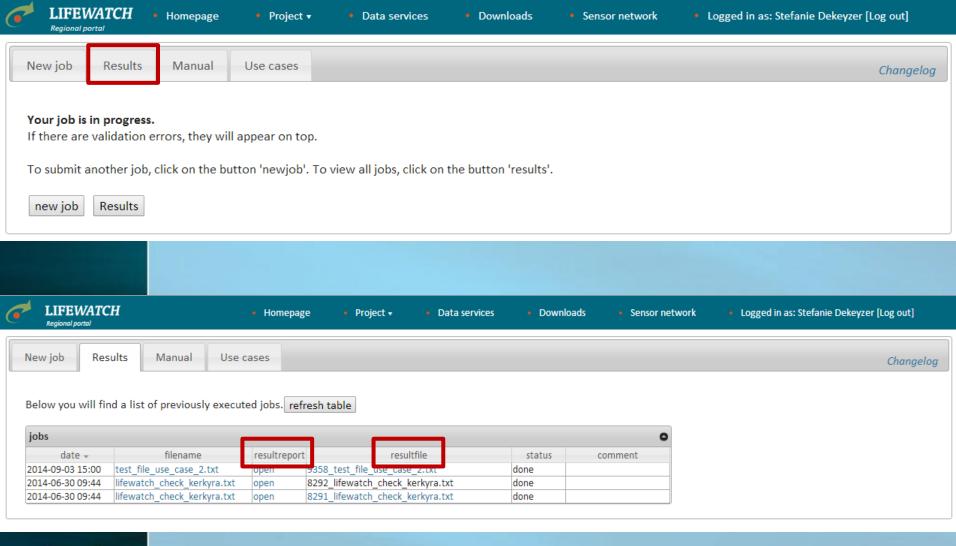
Sensor network

Logged in as: Stefanie Dekeyzer [Log out]





- 1. Get lat-long by name
- 2. Taxon list of a certain region





#### Report

#### http://www.lifewatch.be/data-services

#### Job properties

Name: Dekeyzer Stefanie

Number job: 5816

Comment:

Inputfile: test\_file\_use\_case\_2.txt
Resultfile: 5816\_test\_file\_use\_case\_2.tab

#### Requested Services (2)

- · Get lat-long by name
- Get a taxon list of a certain region (mrgid)

#### Resultfile

(F

#### Results - Get lat-long by name

Number of no exact match(es):0

Number of fuzzy match(es) :0

Number of no match(es) :0

Number of error(s):0 Number of exact match(es):19

#### Results - Get a taxon list of a certain region (mrgid)

Wandelaar: Number of unique species found in OBIS => 0

Belgian Exclusive Economic Zone: Number of unique species found in OBIS => 6805

North Sea: Number of unique species found in OBIS => 5147

Central North Sea: Number of unique species found in OBIS => 15930 Northern North Sea: Number of unique species found in OBIS => 15930

Southern North Sea: Number of unique species found in OBIS => 17025

German part of the North Sea: Number of unique species found in OBIS => 5147

Danish part of the North Sea: Number of unique species found in OBIS => 4487

Dutch part of the North Sea: Number of unique species found in OBIS => 7258

Norwegian part of the North Sea: Number of unique species found in OBIS => 5692

French part of the North Sea: Number of unique species found in OBIS => 2757

United Kingdom part of the North Sea: Number of unique species found in OBIS => 10224

Belgian part of the North Sea: Number of unique species found in OBIS => 6805 North Sea Bottom Current: Number of unique species found in OBIS => 6805

#### Legend - added fields

added\_row\_ws Indicates which service has added the row to the result file

aphiaid Unique identifier within the Aphia database

Latitude\_by\_name Longitude\_by\_name

matched\_name the name of the matching marineregions gazetteer name

mrgid\_by\_name

required\_fields\_check Are the required fields present and completed? (1=yes, 0=no)

Scientific\_name Name of taxon

taxonlist mrgid note Note on taxonlist for mrgid



## List of species for the North Sea

Number of observations for that species in the North Sea

A	В	С	D	Е	F		G	Н	1	J		L	
1 location	required_fields_check	latitude_by_name	longitude_by_name	matched_name	mrgid_by_name	scientific_name		aphiaid	tax	added_row_ws	num_observation		
6812 North Sea	1	5.596.853	5.596.853	North Sea	2350								
6813						Microcystis wesenb	pergii	146561		taxonlist_mrgid	1		
6814						Ampharetinae		152252		taxonlist_mrgid	7		
6815						Culicimorpha		150928		taxonlist_mrgid	5		
6816						Raniceps raninus		126442		taxonlist_mrgid	1		
6817						Munnidae		118263		taxonlist_mrgid	5		
6818						Protohydridae		22788		taxonlist_mrgid	1		
6819						Rhodophyta		852		taxonlist_mrgid	50		
6820						Rhizostoma		135257		taxonlist_mrgid	5		
6821						Talorchestia		101788		taxonlist_mrgid	1		
6822	LIFEWATCH					Gyrodinium spirale		109876		taxonlist_mrgid	1		
6823	Regional portal					Pyramidellinae		224522		taxonlist_mrgid	1		
6824						Eulimidae		135		taxonlist_mrgid	3		
6825	USE CASE 2:					Hirudinea		2041		taxonlist_mrgid	1		
6826		USE (	LASE Z:			Thalassiosira nitzso	chioides	149227		taxonlist_mrgid	1		
6827						Proclea		129711		taxonlist_mrgid	1		
6828	Marir	o chocios	lict and nu	mhor of		Phaxas		138335		taxonlist_mrgid	18		
6829	Marine species list and number of					Clytia		117030	1	taxonlist_mrgid	119		
6830	observations per geographical area					Cellaria salicornia		153836		taxonlist_mrgid	1		
6831						Scopelocheirus hop	ei	102720	1	taxonlist_mrgid	1		
6832						Pasiphaea sivado		107677		taxonlist_mrgid	2		
6833						Laminariaceae		143755		taxonlist_mrgid	2		
6834	^		^			Margelopsis		117151		taxonlist_mrgid	32		
6835			(7)			Halaphanolaimus lo	ongisetosus	121257		taxonlist_mrgid	1		
6836						Cydippida		1251		taxonlist_mrgid	458		
6837			(1)	(16)		Metopa		101764		taxonlist_mrgid	1		
6838			(15)	1		Carpelimus (Taenos	soma) foveolatus	150688		taxonlist_mrgid	2		
6839				(163)	3	Dictyosphaerium pu	ulchellum	178933		taxonlist_mrgid	3		
6840		No.	(3)		10	Lagerheimia genever	nsis	178610	1	taxonlist_mrgid	3		
6841	1			E		Delavalia		347056		taxonlist_mrgid	11		
6842	Which species occur		7. 800	}		Nassarius nitidus		140509		taxonlist_mrgid	1		
6843	And how many times w	vhere they observed?		- " " "	-6	Choniolaimus papil	llatus	120928		taxonlist_mrgid	2		
6844			500			Monhysteridae		2188		taxonlist_mrgid	137		
6845						Chaetonotus (Schizo	ochaetonotus)	371021		taxonlist_mrgid	1		
6846						Microlaimus pingui	is	153418		taxonlist_mrgid	3		
6847						Heterocheila		151022		taxonlist_mrgid	2		
6848						Ironidae		2199		taxonlist_mrgid	121		
6849						Alosa fallax		126415		taxonlist_mrgid	15		
6850						Jania		144012		taxonlist_mrgid	1		
6851						Lymnocardiinae		381865		taxonlist_mrgid	17		
6852						Strombidiidae		1703		taxonlist_mrgid	1		
6853						Tomopteris krampi		131556		taxonlist_mrgid	1		
6854						Antithamnionella te	ernifolia	163275		taxonlist_mrgid	2		
6855						Pherusa flabellata		130110	1	taxonlist_mrgid	1		

# WORKSHOP: Biodiversity data preparation and analysis using LifeWatch virtual labs and web services

- Some of the **skills** that are taught are...
  - Retrieving and working with web services (REST, OGC WMS/WFS, Biodiversity catalogue, GeoNetwork, etc.)
  - Biodiversity data standardization and quality control (taxon matching, geographic quality control, etc.)
  - Building analysis workflows in R, Taverna, LifeWatch elab, etc.
- Organizers: VLIZ-Flanders Marine Institute
- Duration: one and a half two days
- Venue: Oostende
- **Date:** November 26-27 (2015)
- Price: free





## **Conclusions**



## **Conclusions**

- LifeWatch supports biodiversity research and ecosystem studies
- Taxonomic backbone offers online access to species information
- Research & data tools are based on web services and workflows
- Register for the workshop

