Joint European Research Infrastructure network for Coastal Observatory – Novel European eXpertise for coastal observaTories - **JERICO-NEXT**

Deliverable title	Document describing the biological data
Work Package Title	WP5: Data Management
Task title	5.2. Integration of biological data
Deliverable number	D5.5
Description	This document summarises the biological data management efforts during the project and gives an overview of the available biological data gathered by JERICO-Next.
Lead beneficiary	VLIZ – Flanders Marine Institute
Lead Authors	P. Oset
Contributors	
Submitted by	
Revision number	V3
Revision Date	30.08.2019
Security	Public

The JERICO-NEXT project is funded by the European Commission's H2020 Framework Programme under grant agreement No. 654410

History			
Revision	**Date**	**Modification**	**Author**
V1	17/07/2019	First draft	PO
V2	23/08/2018	Added annex	PO
V3	30/08/2019	Incorporated latest feedback from data providers	PO

Approvals				
	Name	**Organisation**	**Date**	**Visa**
Coordinator	Patrick Farcy	Ifremer	10/09/2019	PF
WP Leaders	Leonidas Perivoliotis	HCMR	05.09.2019	LP
WP Leaders	Patrick Gorringe	SMHI	10.09.2019	PG

PROPRIETARY RIGHTS STATEMENT

Table of contents

1. Executive summary

Integration of biological data (Task 5.2) in coastal observatories has been one of the key objectives of the JERICO-NEXT project. Biological data collection has taken place in the framework of JRAP1 and JRAP2 (planktonic and benthic communities) using both traditional and novel observing methodologies. The objective of this task is to provide an operational link between EMODnet Biology and JERICO-NEXT, and to make the JRAP1 and JRAP2 data discoverable by users and increase its FAIRness.

All the biological data in the project has been described and is findable in the EMODnet Biology catalogue and through a map viewer interface. To ensure its reusability, dataset descriptions contain machine-readable licenses that have been agreed with the data providers. A considerable amount of these data is now accessible via EMODnet Biology through different mechanisms: a) links to local data systems or repositories; b) a download link to data files in the Marine Data Archive; c) direct download from EMODnet Biology or integrated in the EMODnet Biology download toolbox. The method chosen depends on the degree or the standardisation possibilities that the data outputs offer.

Making the biological data from novel sensors interoperable remains a challenge due to a lack of mature standards for data types that are still in earlier stage of development. Significant improvements have been achieved in the framework of the Sea Data Cloud project to develop controlled vocabularies for Flowcytometry (FCM) data. During the length of the project, EMODnet Biology and OBIS have transitioned to a new data schema (Darwin Core OBIS-ENV) which brings more flexibility and allows the inclusion of new data types (both biotic and abiotic). The upgraded data schema has allowed for the integration of tests FCM datasets in EMODnet Biology which, together with new developments in the download toolbox, have made JERICO-NEXT data discoverable at the record level in the EMODnet Biology portal.

Despite these achievements, significant efforts are needed in terms of harmonisation throughout all the data management phases, definition of best practices and development of semantic standards for these new data types. This is the only way to ensure that we can exploit the monitoring capabilities of these novel observation methodologies with the appropriate certainty. These should be priorities during the next phase of JERICO-RI. However, it also needs to be recognised that putting data management into practice and opening the data is a resource consuming activity, both at the end-side of the chain (i.e. data aggregators or repositories), but also for data collectors and scientists at the initial steps. Data management needs to be appropriately funded throughout all these different stages to avoid potential bottlenecks, which might hinder the actual publication of data.

2. Introduction

A key objective of JERICO-NEXT has been to promote a stronger integration of biological data within coastal observatories. This has been done through specific Tasks in WP3 to implement innovative measuring techniques to study planktonic and benthic communities, and through its valorization through applied joint research in WP4 (JRAP1 and JRAP2). The task 5.2. Integration of biological data aimed at providing an operational link between JERICO-NEXT and EMODnet Biology, making the data generated by JRAP1 and JRAP2 findable and increasing its accessibility and interoperability.

The first deliverable of this task (D5.4) consisted on a report that described the general landscape and data management practices that are applied in the field of marine biodiversity. JRAP1 and JRAP2 partners were consulted; an initial inventory of the types of data and parameters generated by the project was made to evaluate their potential to fit within the current data schemas. This document provides more insight in the actions undertaken during the following phases of the project. To ensure findability, all the biological data collected in JERICO-NEXT was described and ISO19115 metadata compliant records were created for each datasets. The partners were requested to provide citations and to choose from (machine-readable) CC license to increase the re-usability of their data.

An entry point to the metadata records was provided by connecting the JERICO-NEXT collection in EMODnet Biology catalogue to the JERICO-NEXT website. A bespoke map viewer interface was created to provide a visual access to the metadata records too. Access to datasets has been provided in three different ways either with links to external data repositories, direct predefined download links or through the EMODnet Biology download toolbox for fully harmonized datasets. The option chosen depends on the degree or possibilities of standardization of each specific dataset. Throughout the different sections of this document, extended details on each of these steps are provided. Section 5 includes a status summary and brief evaluation of possible further actions of each JRAP1 and JRAP2 dataset. The final chapter gives additional recommendations for future biological data management practices in the JERICO RI.

3. Description of biological data and metadata catalogue

3.1. Inventory of data collected within JRAP1 and JRAP2

Each of the JRAP1 and JRAP2 partners were contacted during different phases of the project in order to create a digital inventory of the biological data resources that would be originated in the framework of JERICO-NEXT. The partners were asked to describe their datasets following the ISO19115 metadata standard using a distributed template or an online metadata submission form (http://emodnet-biology.eu/contribute). This metadata standard is supported by the VLIZ-managed Integrated Marine Information System (IMIS), a central metadata system with interlinked modules that can be plugged into different data portals, such as EMODnet Biology, using thematic collections.

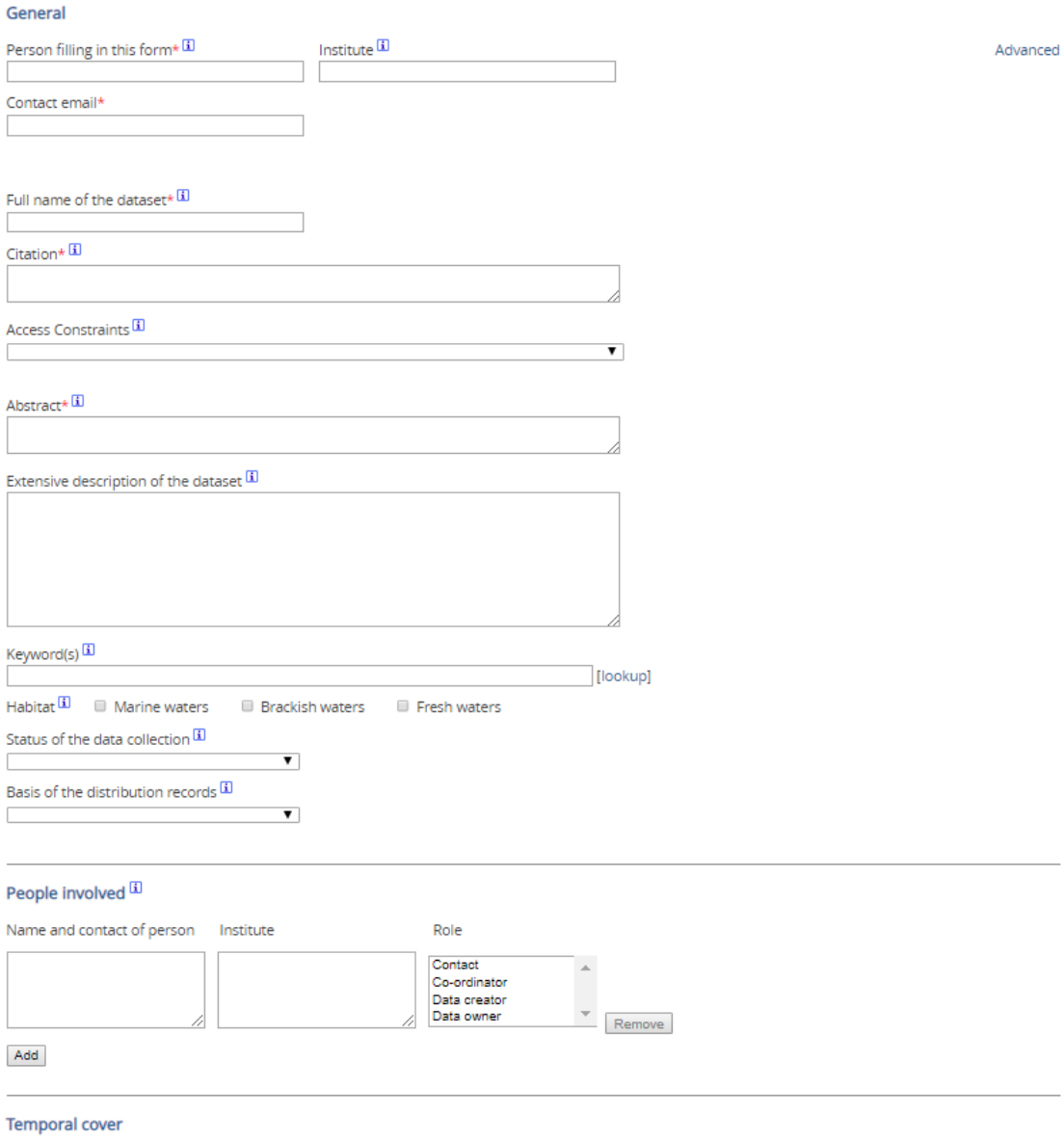


Figure 1. Screen capture of the metadata submission form from EMODnet Biology. A template with equivalent metadata fields was distributed among JRAP1 and JRAP2 partners.

The data required deals with the thematic, temporal and geographical scope of the data, persons and institutes associated with the data collection and processing, as well as basic publishing information (e.g. title, citation, and terms of use). An overview of the mandatory and optional fields from the metadata template is provided in Table 1. The information received by the partners was curated at the VLIZ data centre to enhance the completeness of the metadata records and ensure all the required fields were covered.

Table 1. Overview of mandatory (M) and optional (O) fields in the metadata template distributed amongst the JRAP1 and JRAP2 partners.

Field type	Field name	Key or description
M	Person providing the metadata	Person who is responsible for the content of this metadata record.
M	Dataset title English	Title for the dataset in the original language if it exists.
M	Contact person	Name of the person that should be contacted by a USER to obtain more information about the dataset.
O	Data Creator	The name of a person who is responsible for the creation / maintanance of the dataset. This person will be mentioned in the citation.
M	(Associated) Institute name	1) Either the insitute to that is responsible for the creation / maintanance of the dataset (in this case the field person is not filled out) Or 2) the institute to which the Data Creator person is associated
O	Associated person	The name of a person who is involved in the creation / maintanance of the dataset. This person will NOT be mentioned in the citation.
O	(Associated) Institute name	
M	Dataset Citation	The citation of the dataset. The Citation needs to include as minimum: 1 Data Creator (person and/or insitute); Publicationyear; Dataset title. For Example: Bio-environmental research group; Institute of Agricultural and Fisheries research (ILVO), Belgium; (2015): Epibenthos and demersal fish monitoring at long-term monitoring stations in the Belgian part of the North Sea https://doi.org/10.14284/54 / Vestbo, Stine; Obst, Matthias; Quevedo Fernandez, Francisco; Intanai, Itsara; Funch, Peter; (2017): Horseshoe crab distribution records from East America and Asia https://doi.org/10.14284/293
M	Licence	Creative Commons licence. EMODnet recomended licences are CC-0 https://creativecommons.org/choose/zero/, CC-by https://creativecommons.org/licenses/by/4.0/, CC-by-nc https://creativecommons.org/licenses/by-nc/4.0/
M	Abstract	Provide a short (max 1000 characters) abstract aout the content of the dataset.
O	Description	A larger description about the context of the dataset. May include project information, sampling protocols, ...
M	Geographical Coverage	Indication of the geographical coverage. Use Marineregions.org
M	Temporal Coverage	Minimum to maximum date of data collection. If there is a large temporal gap (+ 2 years) this should be indicated.
M	Taxonomic coverage	Which taxon groups are covered in the dataset. Please consider the appropriated taxonimic level. Guideline: provide 5 taxongroups.
M	Functional groups	List the focus functional group of the dataset

O	Keywords	Any other keyword not yet included elsewhere
O	Websites	Websites which either link to a location where the data can be downloaded, or to a metadata page
O	Publications	A list of all publications which describe the dataset, on which the dataset is based or where the data from the dataset is used.

3.2. Metadata catalogue: JERICO-NEXT collection

The curated information was made available via the IMIS-based EMODnet Biology catalogue using a JERICO-NEXT special collection. The catalogue was pluged in the JERICO-NEXT website and is accessible under the section Data access / Biology catalogue, or directly at the following URL: http://www.emodnet-biology.eu/data-catalog?module=dataset&show=search&spcolid=910.

In total, 29 metadata records were created: 24 for JRAP1 activities and 5 for JRAP 2 activities. The different is related to the number of systems and campaigns within the two JRAPs, but also to the granularity chosen by the partners when describing their data. Most of the JRAP2 partners grouped all the datasets within the same theme and region under the same metadata record, providing a longer temporal extent. The full list of datasets is provided in **Erreur ! Source du renvoi introuvable.**.

A*MIDEX CHROME: Western Mediterranean automated flow cytometry surface sample from Ships of O/P crossing Tunis-Marseille and Tunis-Genova between October 2016-January 2017	JRAP1
Baltic sea UVP5 image data set (July 2017)	JRAP1
Bio-optical measurements from Huvudskär oceanographic buoy 2017	JRAP1
Bio-optical measurements in Utö Baltic Sea, spring 2017-spring 2018	JRAP1
Data from Utö Atmospheric and Marine Research Station	JRAP1
Data from the SMILE buoy in the Bay of Seine	JRAP1
CAMANOC 2014: Ferry Box CAMANOC 2014 cruise data (English Channel, 2014)	JRAP1
Flow cytometer data during Poseidon survey in the Celtic Sea in 2016	JRAP1
Flow cytometer data from RWS during Nephrops survey in the North Sea in 2016	JRAP1
IFCB110-SMHI: Imaging flow cytometry from SMHI in Tangesund 2016	JRAP1
MAREL Carnot: MAREL Carnot data and metadata from Coriolis Data Centre (English Channel, since 2004)	JRAP1
Microscopic analysis of water samples from the Aranda cruise in July 2017	JRAP1
Phycoerythrin fluorescence and cell counts of phycoeryhtrin containing species, ferry Finnmaid Helsinki - Travemunde, summer 2016	JRAP1
Phytoplankton biodiversity data from a North Sea Cruise with R/V Endeavour in June 2016	JRAP1
Phytoplankton biodiversity data from a North Sea Cruise with R/V Simon Stevin in May 2016	JRAP1
Phytoplankton biodiversity data from a North Sea Cruise with R/V Zirfaea in April 2016	JRAP1
Phytoplankton biodiversity data from a North Sea Cruise with R/V Zirfaea in June 2016	JRAP1
Phytoplankton biodiversity data from a Western English Channel Cruise with R/V Endeavour in 2017	JRAP1
Phytoplankton data from the Etoile cruise in the Bay of Biscay in July - August 2017	JRAP1
Phytoplankton data from the PELRAD cruise in the Strait of Dover in July 2017	JRAP1
Phytoplankton data from the PHYCO cruise in the Strait of Dover in April 2017	JRAP1
Plankton biodiversity data from a North Sea Cruise with R/V Simon Stevin in May 2017	JRAP1

Pocket Ferry Box (PFB) and Algae Online Analyser (AOA) data from the IBTS 2015, 2016, 2017 cruises, English Channel and North Sea	JRAP1
Silja Serenade ferrybox measurements between Helsinki and Stockholm, spring 2017-spring 2018	JRAP1
Tavastland ferrybox measurements between Lübeck and Oulu, 2016-2017	JRAP1
Pagure-Next-2016: Benthic biodiversity cruise in the Bay of Brest in October 2016	JRAP2
JERICOBENT: Impact study of Gironde input on benthic ecosystems of the West-Gironde mud-patch (2016 -2018)	JRAP2
Macrobenthos collected in the Cretan Sea between 2016 and 2017	JRAP2
Microbial diversity of the Cretan Sea between 2016 and 2017	JRAP2
Physicochemical characterization of the Cretan Sea between 2016 and 2017	JRAP2

Table 2. List of metadata records available through the JERICO-NEXT collection in the EMODnet Biology catalogue.

3.3. Map interface to the metadata records

In addition to the metadata catalogue, a map interface was developed to allow the user to visualize the geographical extent of the data collected and to access the metadata records of the region of interest directly from the map. Three different layers were created for areas, transects or ship trajectories and points for stations or buoys. By clicking on a feature, an attribute table prompts with basic information about the campaign or dataset: title, temporal scope, JRAP and parameters covered. The table gives access to the metadata record, where more detailed information is available. The map interface allows to filter the displayed features by JRAP and year.

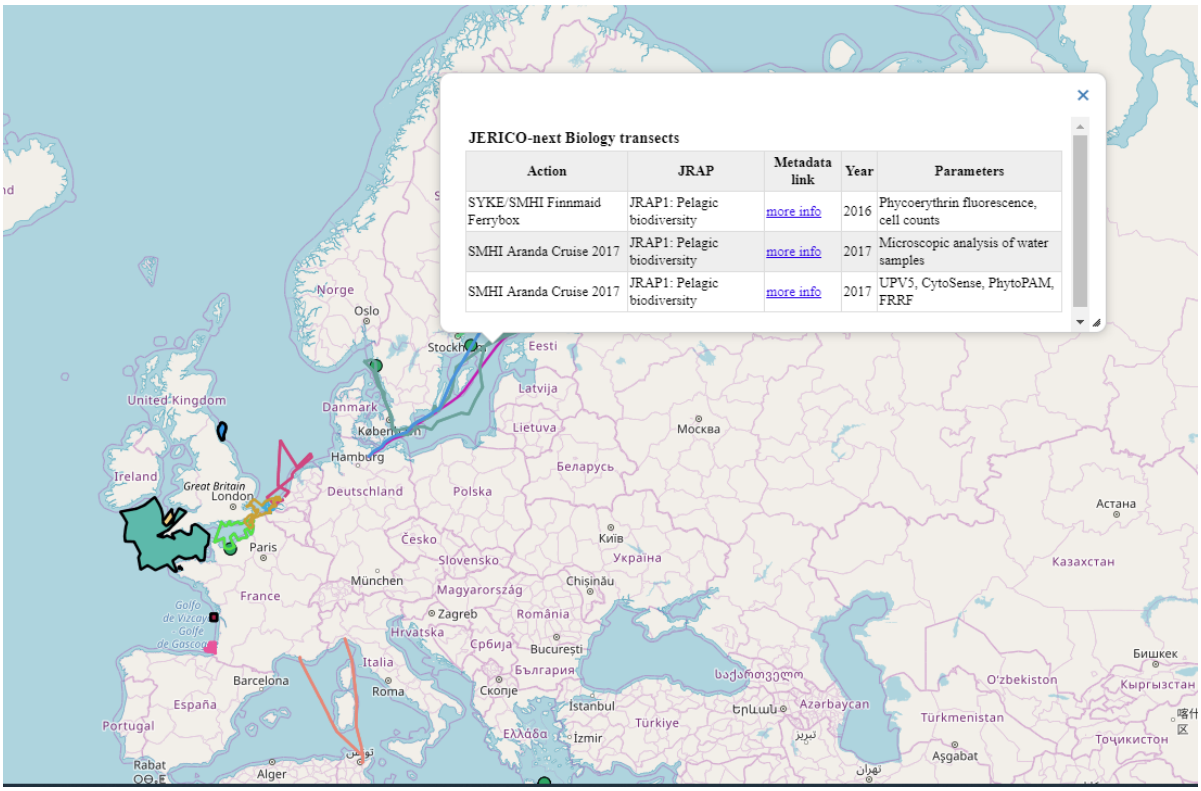


Figure 2. Screen capture of the map interface to access JERICO-NEXT metadata records.

4. Access to biological data

Biological data generated by the project can be made available in different ways throughout the EMODnet Biology portal depending on the level of standardization and integration in the underlying EurOBIS data system. There are two main access to data in the EMODnet Biology portal: with a direct access in the metadata record and integrated in the data download toolbox. The data access provided in the metadata record allows, in turn, for different possibilities.

4.1. Data access via the metadata record

4.1.1. Archived data

Any kind of data in any degree of formatting or standardization can be archived in a documented manner at the Marine Data Archive (MDA). The Marine Data Archive has a public and private section and can be used by individuals or institutions as a back-up or open repository system. A moratorium for publication can be assigned to archived files in the system The MDA is a recommended repository for ocean sciences by Nature *Scientific Data* open access journal.

There are currently five datasets from JERICO-NEXT available in the Marine Data Archive (listed in Table 3).

1	Phytoplankton biodiversity data from a North Sea Cruise with R/V Simon Stevin in May 2016	Rijkswaterstaat (RWD), Flanders Marine Institute (VLIZ), University of Gent (UGent) and Centre National de la Recherche Scientifique (CNRS) were on board of these joint cruises to collect data on phyto- and zooplankton diversity using different methods: fluoroprobe, flow cytometer (FCM), phytopam, FRRF, pigment analysis and zooscan. Different data files associated to this cruise can be accessed on the Marine Data Archive via the metadata record: meteorological and navigation data, CTD, as well as intercomparison FCM data from VLIZ and RWS.
2	Plankton biodiversity data from a North Sea Cruise with R/V Simon Stevin in May 2017	
3	Plankton biodiversity data from a North Sea Cruise with R/V Simon Stevin in April 2018	
4	Pocket Ferry Box and Algae Online Analyzer data from the PHYCO 2017 cruise, English Channel and southern bight of the North Sea	The objective of the PHYCO cruise was to analyse the phytoplankton spring bloom in the English Channel and Southern Bight of the North Sea by combining conventional low frequency and innovative high frequency sampling strategies. Biogeochemical, bio-optical and biological data characteristics were assessed using a pocket Ferry coupled coupled to an algae online analyzer, a pulse shape-recording flow cytometer, an image in flow analysis, a Pulse Amplitude Fluorometer (PAM) and FRRF.
5	Pocket Ferry Box (PFB) and Algae Online Analyser (AOA) data from the IBTS 2015, 2016, 2017 cruises, English Channel and North Sea	This RV Thalassa cruise sampled the entire English Channel (EC) and part of the North Sea (NS) with the objective to describe the species composition of biological assemblages (fish, plankton, benthos), to characterize their habitats and spatial distributions, to understand the food web structure and to establish a set of indicators related to the ecological state and the descriptors from the MSFD. To do so, the EC-NS area was sampled by complementary techniques: hydrological probe, Niskin bottle, high frequency measurements systems such as Pocket Ferry Box, LOPC, plankton nets, GOV trawl, pelagic trawl, grab, dredge, ROV for sub-marine video, multibeam echosounders and visual observations.

Table 3. List and short description of JERICO-NEXT datasets available in the Marine Data Archive.

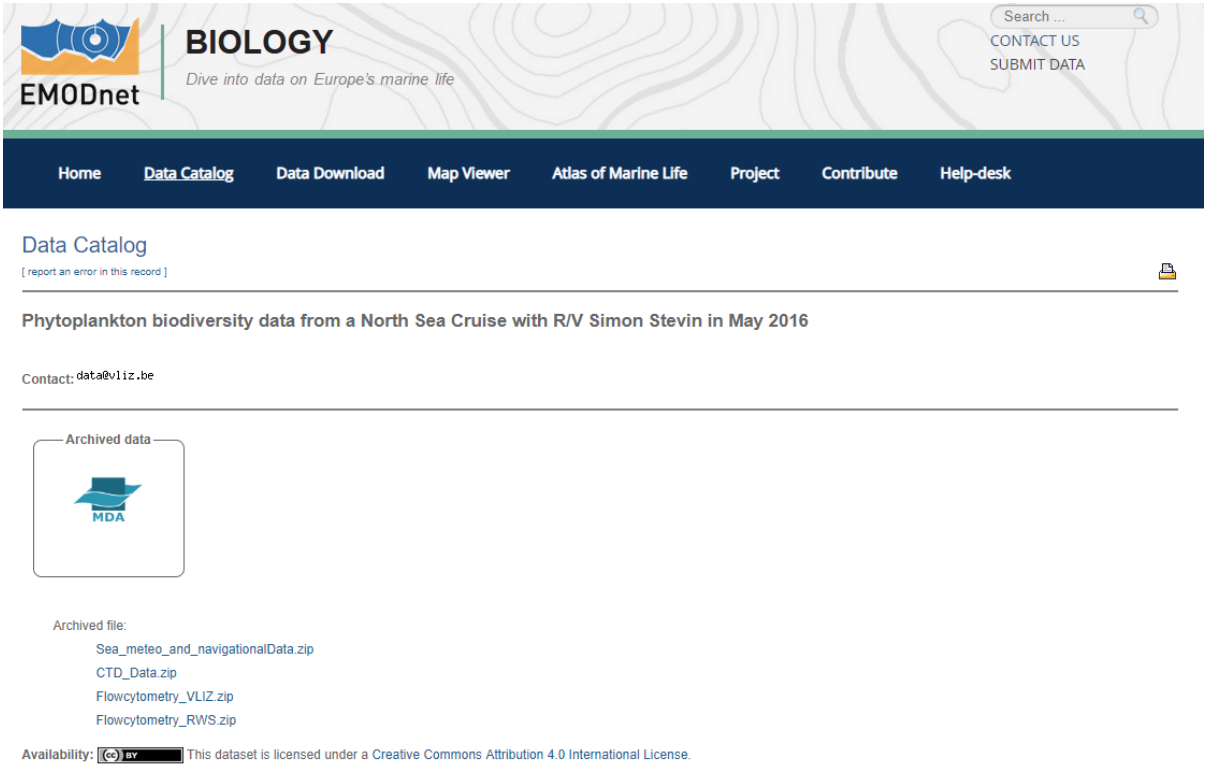


Figure 3. Screen capture of a metadata record in EMODnet Biology providing access to documented archived data in MDA.

4.1.2. *Direct link to in-house system or data repository*

In the cases in which local online data access and visualisation portals exist for a particular dataset, a web link to the existing interfaces can be provided in within the dataset description. There are currently nine JERICO-NEXT datasets with ad-hoc data interfaces (Table 4).

1	Phytoplankton biodiversity data from a North Sea Cruise with R/V Endeavour in June 2016	Biomass and diversity of phytoplankton communities were determined using FerryBox and on line flow cytometer at 4 meters. Flow cytometry analyses were performed every hour from the same inlet water as the FerryBox during the survey. Discrete samples were taken using CTD rosette at 4m to calibrate the chlorophyll fluorescence. The analysis of the phytoplankton functional types were done manually (according to the knowledge of the operator), and the discrete samples were analysed by HPLC (DHI, Denmark). All data from the FerryBox (fluorimetry, turbidity, salinity, temperature, oxygen) as well as HPLC and flow cytometry followed specific quality controled procedures before being available.The Cefas tool, called "Phyto OPS", visualizes data collected during surveys on board the RV Cefas Endeavour, focusing on chlorophyll and phytoplankton functional types. It also gives information on additional parameters (e.g. Ferrybox temperature and salinity).
2	Phytoplankton biodiversity data from a Western English Channel Cruise with R/V Endeavour in 2017	The cruise was organized to perform an integrated monitoring survey of the Celtic Sea, south-western approaches and the western Channel using a random stratified survey design for the purposes of providing fish stock assessment data and associated ecosystem information, such as, amongst

		others chlorophyll, oxygen, salinity temperature, nutrient samples, plankton from water samples. To continuously log sub-surface (4m) salinity, temperature, fluorometry and other environmental data using the 'Ferrybox'. Additionally, a Flow-cytometer (phytoplankton) and Plankton Image Analyser (PIA) (zooplankton) may be run in conjunction with the Ferrybox. These data can be visualized in the Phyto OPS tool.
3	Data from Utö Atmospheric and Marine Research Station	This dataset encompasses bio-optical measurements carried out at Utö Atmospheric and Marine research station, located on Utö island at the outer edge of the Finnish archipelago. It includes continuous measurements of temperature, salinity, O2, fluorescence for Chlorophyll, phycocyanin, phycoerythrin and CDOM, spectral fluorescence and FRRF fluorometry from a flow-through system located at a depth of 5 m. The link gives access to an interface with real-time data measured at the Utö Station.
4	MAREL Carnot: MAREL Carnot data and metadata from Coriolis Data Centre (English Channel, since 2004)	The MAREL Carnot system was developed and implemented by Ifremer in 2004 and is located in the Boulogne-sur-Mer harbor (eastern English Channel) influenced both by marine coastal and fresh waters. It is a moored buoy equipped with physico-chemical and biological measuring devices working in continuous and autonomous conditions, and with near real time data transmission. MAREL Carnot records the following parameters with a high frequency resolution (20 minutes): estimated sea level, gust wind speed, wind from direction relative true north, horizontal wind speed, relative humidity, light irradiance surface PAR, sea temperature, practical salinity, ph, dissolved oxygen, fluorescence, turbidity. Nitrate + nitrite, phosphate and silicate are sampled with 12 hour frequency. The dataset is made available via the SEANOE repository with a temporal coverage ranging from 2004-03-24 to 2015-06-15.
5	Baltic sea UVP5 image data set (July 2017)	The Underwater Vision Profiler (UVP) was used in the Baltic sea during the Aranda cruise in July 2017 in the frame of JERICONEXT to quantify the vertical distribution of macroscopic particles >100µm and plankton > 700 µm in size (incuding large cynaobacterial filaments). The smaller size limit is fixed by optical resolution, whereas the larger size limit is determined by the volume of water illuminated per image. Images are recorded at a frequency up to 20 Hz and the recorded volume per image is 1.02 L. Images with annotated taxonomy are available at Ecotaxa.
6	Data from the SMILE buoy in the Bay of Seine	SMILE Buoy is a moored buoy equipped with physico-chemical and biological measuring devices working in continuous and autonomous conditions. The system is located 2 nautical miles north of Luc-sur-mer (Basse normandie, Central English Channel) influenced both by marine coastal and fresh waters. The measuring station is equipped with high performance systems for seawater analysis and near real time data transmission. This system records with a high frequency resolution (20 minutes), the following parameters, gust wind speed, wind from direction relative true north, horizontal wind speed, relative humidity, atmospheric pressure at sea level, air temperature in dry bulb, light irradiance surface PAR, sea temperature, practical salinity, dissolved oxygen, fluorescence, turbidity. MAREL SMILE is a component of the coastal observing network COAST-HF developed and implemented by Ifremer & the University of Caen Normandy in 2015. Data from the SMILE buoy is available through the Seanoe repository.

7	Tavastland ferrybox measurements between Lübeck and Oulu, 2016-2017	This dataset contains data from continuous measurements along the Tavastland (formerly TransPaper) Serenade ferry line between Oulu-Kemi and Lübeck. At the depth of three meter under the water surface, sea water is constantly pumped into an on-board mini-laboratory. Parameters collected include chloropyll a, phycocyanin and CDOM fluorescence and turbidity.
8	Bio-optical measurements from Huvudskär oceanographic buoy 2017	Hourly measurements using the Huvudskör E. buoy. Several abiotic and bio-optical parameters are collected. For phytoplankton, chlorophyll and phycocyanin fluorescence are the most relevant parameters.
9	Silja Serenade ferrybox measurements between Helsinki and Stockholm, spring 2017-spring 2018	Dataset contains ferrybox (ferry Silja Serenade) measurements between Helsinki and Stockholm from spring 2017 to spring 2018. Dataset includes continuous measurements of temperature, salinity, O2, fluorescence for Chlorophyll, phycocyanin, and CDOM from a flow-through system (sampling depth -5 m). Data set also includes monthly/weekly measurements of extracted Chlorophyll. Data from the flow-through system are made available in CMEMS (previous registration). Data from water samples is available in EMODnet Chemistry.

Table 4. List and short summary of JERICO-NEXT datasets with a dedicated data access interface.

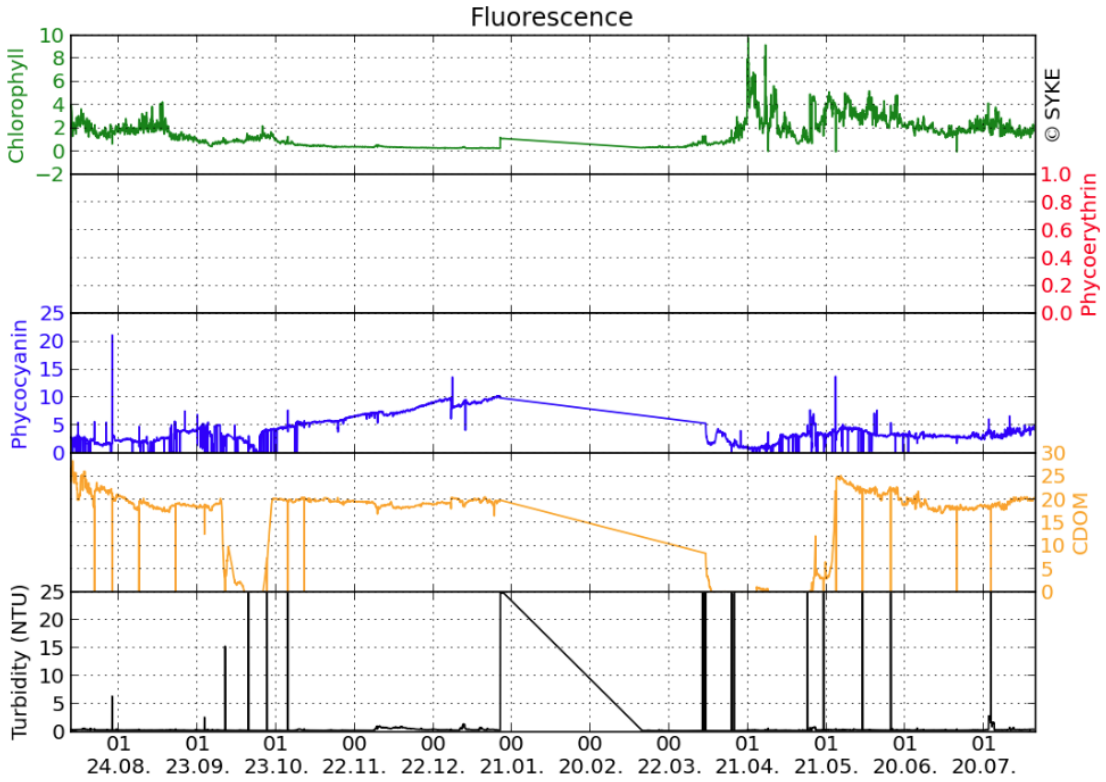


Figure 4. Screen capture from the Utö Atmospheric and Marine Research Station data access interface, showing fluorescence data from the past year.

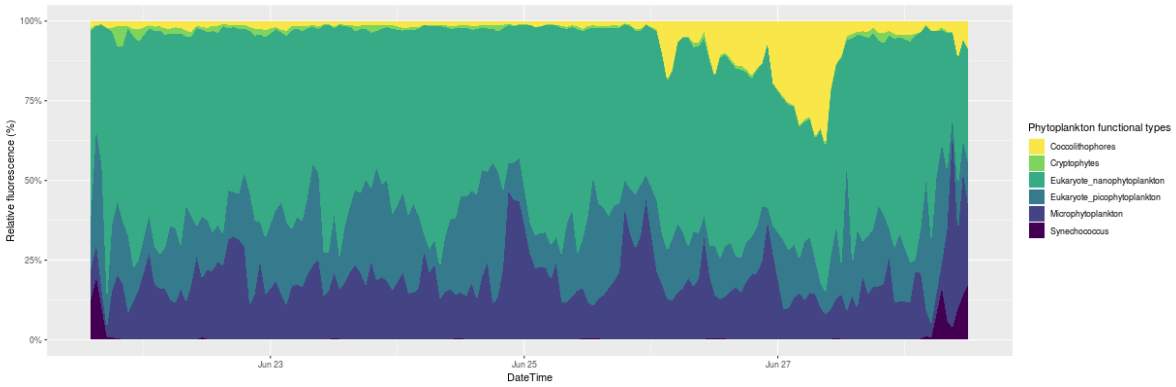


Figure 5. Screen capture of the Cefas Phyto OPS interface to visualise chlorophyll and phytoplankton functional groups data, showing data on relative fluorescence (in %) per functional groups or clusters

4.1.3. Harmonised dataset integrated in EMODnet Biology (direct link to dataset)

Additionally, those datasets that are fully standardized to the EMODnet Biology – Eur(OBIS) format, can be integrated in the EurOBIS data system and become available in an interoperable manner via the EMODnet Biology download toolbox and the OGC compliant data access web services. Those datasets present an additional link in the metadata record, which points to a sub-selection of the EurOBIS system containing records that belong only to that specific dataset. Chapter 4.3 deals specifically with the standardization and integration process. There are currently three JERICO-NEXT datasets that have been fully integrated into EMODnet Biology / EurOBIS, and two that can be harvested from MedOBIS IPT once they are published.

1	IFCB110-SMHI: Imaging flow cytometry from SMHI in Tangesund 2016	An imaging flow cytometer, the Imaging FlowCytobot, was used to investigate phytoplankton abundance and diversity in a fjordal system on the Swedish Skagerrak coast. Samples of 5 mL were collected approximately every 25 minutes. An automated winch was used to move the IFCB to different depths. Chlorophyll fluorescence of individual organisms was used to trigger the camera in the instrument. Several thousand images were collected in each sample. Automated image analyses was used to analyze the images produced to identify and count cells of different plankton taxa.
2	Plankton biodiversity data from a North Sea Cruise with R/V Simon Stevin in May 2017	This joint cruise described in Section 4.1.1 collected data zooplankton diversity using different methods: fluoroprobe, flow cytometer (FCM), phytopam, FRRF, pigment analysis and zooscan. From that cruise, the VLIZ FCM cluster analysis data has been integrated into EMODnet Biology.
3	A*MIDEX CHROME: Western Mediterranean automated flow cytometry surface sample from Ships of O/P crossing Tunis-Marseille and Tunis-Genova between October 2016-January 2017	Phytoplankton functional diversity and spatio-temporal distribution at the meso-scale are studied in the frame of the A*MIDEX CHROME project (Continuous High Resolution Observation of the Mediterranean), thanks to the combined installation of a Ferrybox system and a cytosense flow cytometer onboard the CTN's ferry "Le Carthage". Samples were collected during the cruises between Tunis-Marseille and Tunis-Genova from October 2016 to January 2017. This led to 80 transects and more than 7000 samples collected. Data analyses resolved the abundances of up to six phytoplankton functional groups based on size and pigment content. The data has been analyzed but up to date, only one transect was

		fully processed and uploaded into SeaDataNet, from where it has been made available in EMODnet Biology. Flow cytometry data was submitted to SDN and made available in EMODnet Biology trhough existing data flows between the two systems. The availability of FCM data in EMODnet Biology has been made possible because of the development of new vocabularies in the framework of SDC, and the adoption of the extended Darwin Core OBIS-ENV schema by EMODnet Biology. Further details are provided in Section 4.3.
4	Macrobenthos collected in the Cretan Sea between 2016 and 2017	The dataset contains information on the soft sediment macrobenthic communities from the Bay of Heraklion (Crete, Greece). Data include the species of the taxonomic groups of Mollusca, Annelida, Crustacea and Echinodermata found in the area. Data is uploaded in the Mediterranean node of OBIS (http://ipt.medobis.eu/), but not yet publically visible (under moratorium).
5	Microbial diversity of the Cretan Sea between 2016 and 2017	The dataset contains information on the soft sediment microbial communities from the Bay of Heraklion (Crete, Greece). Data include Bacteria and Archaea OTUs based on sequencing of the 16S rRNA gene. Data is uploaded in the Mediterranean node of OBIS (http://ipt.medobis.eu/), but not yet publically visible (under moratorium).

Table 5. List and short summary of JERICO-NEXT datasets that have been harmonised to DwC standard.

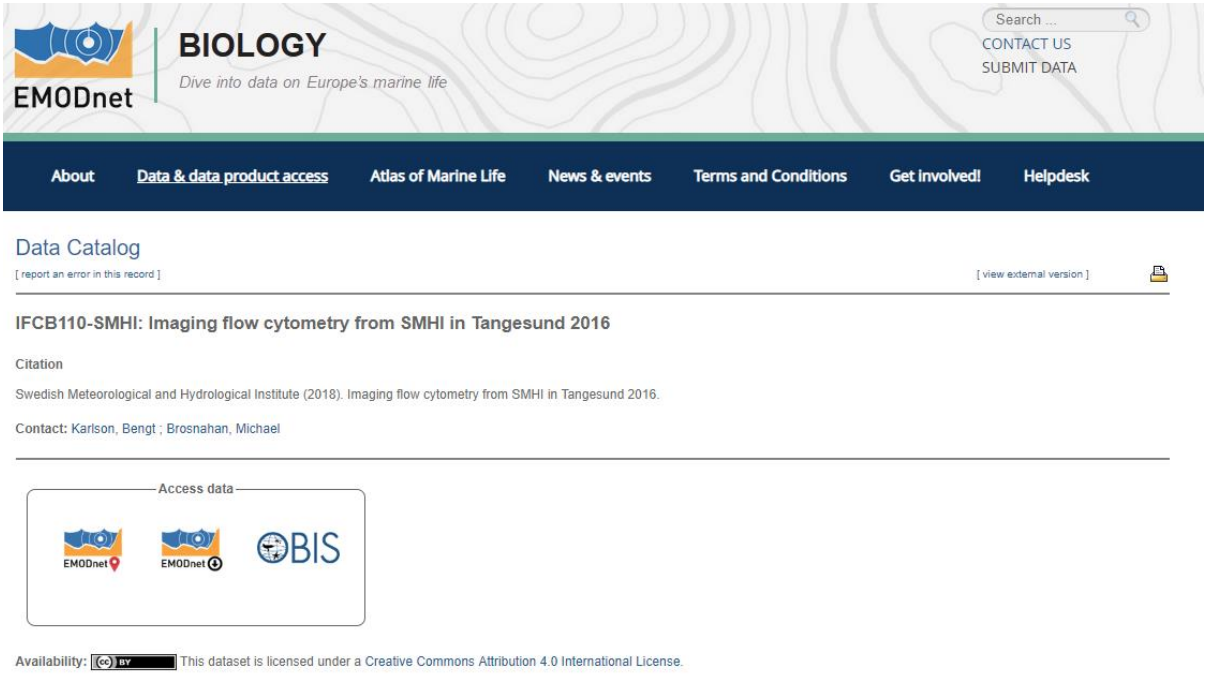


Figure 6. Screen capture of a JERICO-NEXT dataset metadata record showing direct access to the sub-selection in EMODnet Biology download toolbox.

4.2. EMODnet Biology download toolbox for standardized datasets

In addition to accessing data through each individual metadata record, data that are fully standardized and integrated into the EMODnet Biology underlying database (EurOBIS), can be searched through a specific interface, the data download toolbox.

The data download toolbox (http://www.emodnet-biology.eu/toolbox/en/download/occurrence/explore) gives access to the EMODnet Biology data through a step-wise approach. Users can query the full database at the occurrence record level and apply different filters (temporal, geographical, taxonomical, additional biotic/abiotic parameters). Only the data that meets the criteria established by the user will reach the final selection step. Users can either download their data as a csv file or get a web service URL for the data that passes the filters. There are data from three JERICO-NEXT campaigns that can be found via the download toolbox:

IFCB110-SMHI: Imaging flow cytometry from SMHI in Tangesund 2016

Plankton biodiversity data from a North Sea Cruise with R/V Simon Stevin in May 2017

A*MIDEX CHROME: Western Mediterranean automated flow cytometry surface sample from Ships of O/P crossing Tunis-Marseille and Tunis-Genova between October 2016-January 2017

4.2.1. Accessing JERICO-NEXT data in the download toolbox

The advantage of the EMODnet Biology download toolbox is that the user does not require any background information about the data in advance (e.g. where it was collected and by which organizations). With an intuitive gradual selection steps that query the data at the individual occurrence level. However, in the first step of the download toolbox it is also possible to enter metadata-related keywords and organizations to query for specific datasets or data that match those criteria at the metadata level. The following table describes the main steps and sub-steps of the download toolbox, and how JERICO-NEXT data can be found:

1 (Explore)	Data source	The user can find full datasets based on metadata: data creators (organizations or countries), keywords (e.g. “JERICO”).
	Geographical area	Data within a bounding box or that intersects with pre-defined geographical units can be selected.
	Time period	A specific temporal range can be added to the selection criteria.
2 (Select)	Species traits	The user can filter occurrences based on functional groups or characteristics that are relevant for society (e.g. red list species).
	Taxonomic search	Data from specific taxa (multiple selection) can be queried.
	Data precision	The user can exclude data that does not have enough precision (e.g. taxonomic rank at level of genus).
	Parameter filters	Additional biotic or abiotic data can be queried. E.g. Flow cytometry data.
3 (Data download)	Download data	The resulting can be download records as csv file or obtained as a web service URL. The query can be stored as well and be run again in the future.
	Visualize on map	The final selection can be visualized on a map.

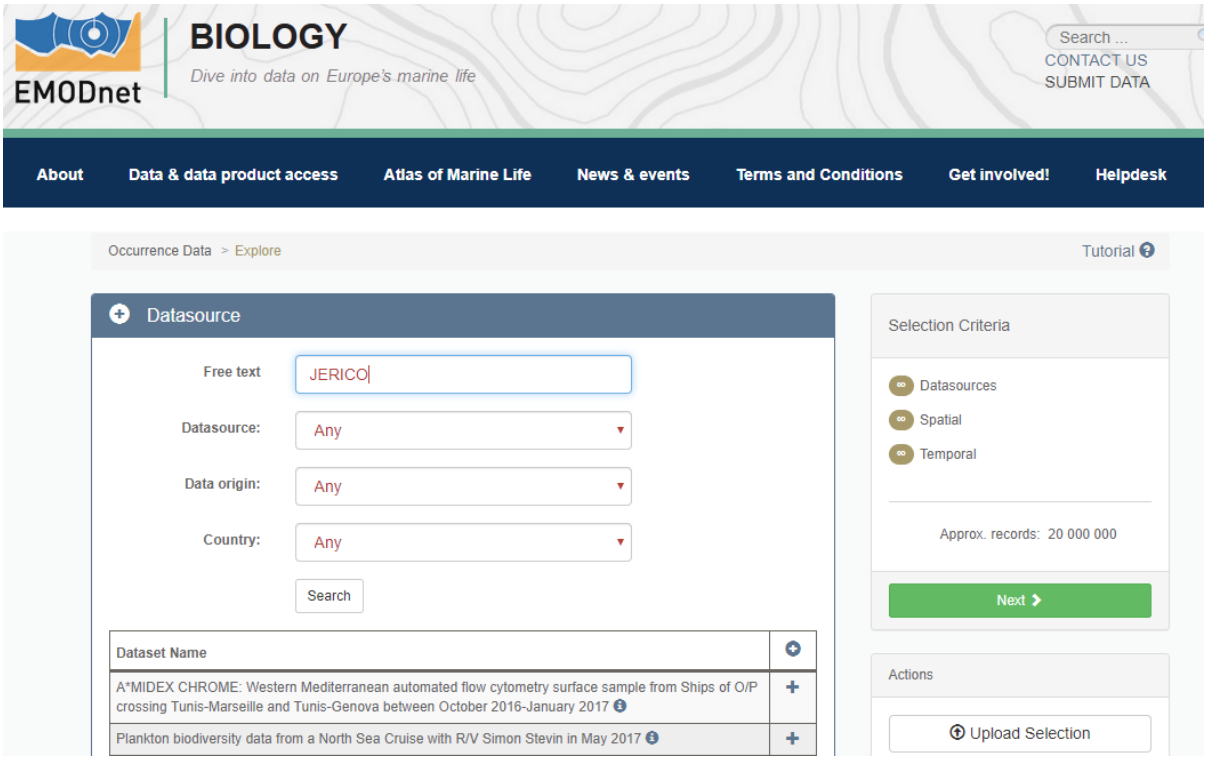


Figure 7. Screen capture of the first step of the EMODnet Biology download toolbox, where any keyword can be entered to look for data.

In the framework of the EMODnet Biology project, a new version of the download toolbox was launched at the begininig of 2019.This version includes new technical implementations that exploit the interoperability advantages gained by the adoption of the Darwin Core (DwC) OBIS-ENV format and the use of the BODC controlled vocabularies. Further details on these standards are provided in Section 4.3. Amongst the new features of the download toolbox, a search option was developed to allow the user to query for data which contains additional parameters beyond taxonomy, such as FCM data, abiotic data or data acquired by specific sampling methods.

4.3. Interoperability, standards and data management practices for marine biodiversity data

The possibility to filter at the record level is enabled by the application of strict syntactic and semantic standards. The used of controlled schemas and vocabularies enables machine-to-machine communications and allows the data in EMODnet Biology to be interoperable with larger biodiversity and marine data initiatives such as OBIS, GBIF or SDN. The common data schema used is the Darwin Core Archive, an internationally recognized biodiversity informatics data standard that simplifies the publication of biodiversity data. Semantic standards include: a) the World Register of Marine Species (WoRMS) for taxonomy, the authoritative and comprehensive global list of names of marine organisms; b) controlled thesaurus from the NERC Vocabulary Server maintained by the British Oceanographic Data Centre (BODC) for additional biotic or abiotic measurements. Data is published through the open source Integrated Publishing Toolkit (IPT), from where it is harvested into other data portals with a global focus (OBIS, GBIF).

4.3.1. Darwin Core standard

Darwin Core (DwC) is a set of syntactic and semantic standards developed and maintained by the Biodiversity Information Standards (TDWG) intended to facilitate the sharing of information about biological diversity. DwC is widely used to share both terrestrial and marine biodiversity data. The DwC archive is a star-based schema with a central component (core data file or table) and irradiating extensions (files / tables) connected to the core.

Based on growing requirements from the biodiversity data community to manage data derived from novel biological sensors and datasets that combine biological and environmental information, EMODnet Biology (and OBIS) transition from a flat DwC to a multiple extension DwC schema[1]. It consists of a DwC Event Core in combination with a DwC Occurrence Extension and a proposed enhancement to the DwC MeasurementOrFact Extension. This new structure enables the linkage of measurements or facts - quantitative and qualitative properties - to both sampling events and species occurrences, and includes additional fields for property standardization by the use of controlled vocabularies.

The adoption of the OBIS-ENV adaptation of the DwC Event core schema has brought the possibility to ingest new data types in EMODnet Biology in an interoperable manner. In the framework of the SeaDataCloud project, a set of new controlled vocabularies for flow cytometry data were developed. A total of new 34 parameter codes were added in the P01 vocabulary collection, after thorough investigation and consultation with the flow cytometry community. The combination of these vocabularies with the use of the Extended MesurementOrFact extension of the new DwC schema enables to look for data in EMODnet Biology that goes beyond the classic taxonomically resolved occurrences, as described in Section 4.2.1.

4.3.2. Taxonomic backbone: World Register of Marine Species

The World Register of Marine Species (WoRMS) is an authoritative and comprehensive list of names of marine organisms, including information on synonymy. The content of WoRMS is controlled by taxonomic and thematic experts globally distributed, who can use the online editing environment of the underlying Aphia data infrastructure system. Aphia uses unique and stable identifiers for each available name in the database, through Life Science Identifiers (LSIDs). This electronic identifiers enable interoperability with other connected systems, such as: EMODnet Biology, OBIS, LifeWatch, GBIF and multiple NODCs.

4.3.3. BODC-NERC controlled vocabularies

The NERC Vocabulary Server gives data managers the means to access lists of controlled terms to describe data that cover a broad spectrum of disciplines of relevance to the oceanographic and wider community. The vocabulary services are technically managed and hosted by the British Oceanographic Data Centre (BODC). The vocabularies are organized in a set of concepts, collections and schemes. Those collections and concepts which are more relevant or frequently used for biodiversity data are documented in EMODnet Biology and OBIS guidelines and online training[2,3].

4.3.4. Quality control of harmonized datasets

Data in JERICO-NEXT that is acquired with new observation techniques is in a development stage in terms of data management practices and standards. Data outputs can be highly variable and for most of the sensors quality control currently consists of expert validation of the data. For data that can undergo full semantic and syntactic standardization, automatisation of quality control procedures is possible. In the framework of EurOBIS and EMODnet Biology, an online QC tool was develop to facilitate the QC of biological data published in DwC format trough the Integrated Publishing Toolkit (IPT). IPT is the platform used by all the OBIS regional nodes, including EurOBIS. The objective of the tool is twofold: a) it checks integrity and compliance with the EMODnet standards (e.g. it reports on records that are not yet semantically standardized to WoRMS or BODC vocabularies); and, b) it reports on potential inconsistent or bad data (e.g. depth values that don't match bathymetry records for that

[1] De Pooter 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://doi.org/10.3897/BDJ.5.e10989
[2] OBIS (2019). Darwin Core guidelines. Available at: https://obis.org/manual/darwincore/
[3] VLIZ (2018). Contributing datasets to EMODnet Biology. Online course available at the Ocean Teacher platform: https://classroom.oceanteacher.org/course/view.php?id=328

location, dubious taxonomy, etc.). The tool also checks whether sampling information is present on the dataset, which is often crucial to understand the given data (e.g. false absences). The online QC tool is available at: http://rshiny.lifewatch.be/BioCheck/. More details about the QC tool are given in Annex 1.

A large part of these sensors (incl. software) have large potential for monitoring purposes, but are currently still in a phase of development. As a consequence the data output is also highly variable and for most of the sensors there is not yet a consensus on protocols for data quality control. Quality control currently consists of expert validation of the data. It is the intention of this WP to document on these procedures and work towards standard output parameters for the different sensors. Once this is achieved and data is stored in a database, the above described delayed mode QC procedures (or parts) could be applied.

5. Summary of data access status and future actions

At the beginning of the project, the different parameters and data outputs to be collected by the JRAP1 and JRAP2 partners were inventoried and categorized according to their potential to be fitted into the existing data standards. Many of the cruises have collected a combination of environmental parameters together with biodiversity related measurements. Although the DwC OBIS-ENV schema allows storing some of these data, it has been decided not to stream data that is not directly linked to biodiversity occurrences. There exist other repositories and/or European data infrastructures, such as SDN, Copernicus or other EMODnet thematic lots, that are better suited for such data and some have indeed been made available using such data flows. The table below summarizes the type of data outputs that was collected for each dataset, with an overview of the status of the data access and what potential actions could be taken in the future. It is worth noting that, regardless the format and/or output, any data file can be made available in a documented manner at the Marine Data Archive, and some of the data is already archived in this system.

Dataset	Parameters / data outputs	Status
A*MIDEX CHROME: Western Mediterranean automated flow cytometry surface sample from Ships of O/P crossing Tunis-Marseille and Tunis-Genova between October 2016-January 2017	Phytoplankton functional composition by FCM.	Partly integrated in EMODnet Biology (one transect)
Baltic sea UVP5 image data set (July 2017)	Plankton communities with Underwater Vision Profiler	Annotated images available online at EcoTaxa. There are open discussions on future connections between EMODnet Biology and EcoTaxa.
Bio-optical measurements from Huvudskär oceanographic buoy 2017	Chlorophyll fluorescence and phycocyanin fluorescence	EMODnet Biology does not focus on optical measurements, but can provide them if they are coupled with occurrences.
Bio-optical measurements in Utö Baltic Sea, spring 2017-spring 2018	Fluorescence for Chlorophyll, phycocyanin, phycoerythrin and CDOM, spectral absorption, spectral fluorescence, FRRF fluorometry from a flow-through system	EMODnet Biology does not focus on optical measurements, but can provide them if they are coupled with occurrences.
Data from the SMILE buoy in the Bay of Seine	Light irradiance surface PAR, fluorescence.	EMODnet Biology does not focus on optical measurements. Data from the SMILE buoy is available through the Seanoe repository.
CAMANOC 2014: Ferry Box CAMANOC 2014 cruise data (English Channel, 2014)	Total Fluorescence, phytoplankton taxonomical and functional composition, photosynthetic parameters	The taxonomic and functional data could be integrated in EMODnet Biology, and the additional optical information could be included linked to the occurrences.

Flow cytometer data during Poseidon survey in the Celtic Sea in 2016	Fluorometry from Ferrybox, phytoplankton functional types with a flow-cytometer (Cytosense) and zooplankton with the Plankton Image Analyser (PIA), Chlorophyll a concentration	The taxonomic and functional data could be integrated in EMODnet Biology, and the additional optical information could be included linked to the occurrences.
Flow cytometer data from RWS during Nephrops survey in the North Sea in 2016	Total fluorescence, phytoplankton functional composition	The taxonomic and functional data could be integrated in EMODnet Biology, and the additional optical information could be included linked to the occurrences.
IFCB110-SMHI: Imaging flow cytometry from SMHI in Tangesund 2016	Phytoplankton abundance and diversity with imaging flow cytometer.	Taxonomically resolved data that is already integrated in EMODnet Biology.
MAREL Carnot: MAREL Carnot data and metadata from Coriolis Data Centre (English Channel, since 2004)	Light irradiance surface PAR, fluorescence.	EMODnet Biology does not focus on optical measurements. Data from the SMILE buoy is available through the Seanoe repository.
Microscopic analysis of water samples from the Aranda cruise in July 2017	Cyanobacteria and zooplankton biodiversity and abundance	Taxonomically resolved data could be integrated in EMODnet Biology.
Phycoerythrin fluorescence and cell counts of phycoeryhtrin containing species, ferry Finnmaid Helsinki - Travemunde, summer 2016	Continuous measurements of phycoerythrin fluorescence, weekly cell counts for phycoerythrin containing species	EMODnet Biology does not focus on optical measurements, but can provide them if they are coupled with occurrences.
Phytoplankton biodiversity data from a North Sea Cruise with R/V Endeavour in June 2016	Biomass and diversity of phytoplankton communities with FCM.	Data is available through a Cefas R-Shiny based tool. It could be made available in EMODnet Biology with use of P01 vocabularies (functional composition and associated measurements).
Phytoplankton biodiversity data from a North Sea Cruise with R/V Simon Stevin in May 2016	Total fluorescence, phytoplankton functional composition (fluoroprobe, flow cytometer, phytopam, FRRF)	Data is available at the MDA. FCM data could be made available in EMODnet Biology with use of P01 vocabularies (functional composition and associated measurements).
Phytoplankton biodiversity data from a North Sea Cruise with R/V Zirfaea in April 2016	Phytoplankton functional groups and biomass dynamics at high frequency ((Flow cytometry, FRRF)	Data could be made available in EMODnet Biology with use of P01 vocabularies.
Phytoplankton biodiversity data from a North Sea Cruise with R/V Zirfaea in June 2016	Phytoplankton functional groups and biomass dynamics at high frequency ((Flow cytometry, FRRF)	Data could be made available in EMODnet Biology with use of P01 vocabularies (functional composition and associated measurements).
Phytoplankton biodiversity data from a Western English Channel Cruise with R/V Endeavour in 2017	Total fluorescence, phytoplankton functional composition, photosynthetic parameters and primary productivity	Data is available through a Cefas R-Shiny based tool. It could be made available in EMODnet Biology with use of P01

		vocabularies (functional composition and associated measurements).
Phytoplankton data from the Etoile cruise in the Bay of Biscay in July - August 2017	Phytoplankton functional composition by flow cytometer and multispectral fluorometry	Data could be made available in EMODnet Biology with use of P01 vocabularies (functional composition and associated measurements).
Phytoplankton data from the PELRAD cruise in the Strait of Dover in July 2017	Total fluorescence, phytoplankton taxonomical and functional composition, photosynthetic parameters, primary productivity	Data could be made available in EMODnet Biology with use of P01 vocabularies (functional composition and associated measurements).
Phytoplankton data from the PHYCO cruise in the Strait of Dover in April 2017	Total fluorescence, phytoplankton taxonomical and functional composition, photosynthetic parameters, primary productivity	Data could be made available in EMODnet Biology with use of P01 vocabularies (functional composition and associated measurements).
Plankton biodiversity data from a North Sea Cruise with R/V Simon Stevin in May 2017	Total fluorescence, phytoplankton functional composition, environmental parameters (CTD)	Data has been made available via the MDA and the functional composition is integrated in EMODnet Biology
Pocket Ferry Box (PFB) and Algae Online Analyser (AOA) data from the IBTS 2015, 2016, 2017 cruises, English Channel and North Sea	Fluorescence, phytoplankton functional composition, on board parameters and additional physico-chemical data from PFB	Data has been made available in the MDA. Data could potentially be made available in EMODnet Biology with use of P01 vocabularies (functional composition and associated measurements).
Silja Serenade ferrybox measurements between Helsinki and Stockholm, spring 2017-spring 2018	Continuous measurements of temperature, salinity, fluorescence for Chlorophyll, phycocyanin, and CDOM, Chlorophyll from a flow-through system.	Data from the flow-through system are made available in CMEMS (FTP server with previous registration). Data from water samples is available in EMODnet Chemistry.
Tavastland ferrybox measurements between Lübeck and Oulu, 2016-2017	Temperature, salinity, chloropyll a, phycocyanin and CDOM fluorescence and turbidity.	EMODnet Biology does not focus on optical measurements. NRT data are already available. Archived data could be made available on the MDA.
Pagure-Next-2016: Benthic biodiversity cruise in the Bay of Brest in October 2016	Taxonomic annotations from underwater images	Species occurrences could be integrated into EMODnet Biology with a temporal and geographical component.
JERICOBENT: Impact study of Gironde input on benthic ecosystems of the West-Gironde mud-patch (2016 -2018)	Taxonomic occurrences and abundances, vertical profiles of sediment column, microganulometry, O2 and nutrients.	Species occurrences and abundances could be integrated into EMODnet Biology with a temporal and geographical component.

Macrobenthos collected in the Cretan Sea between 2016 and 2017	Benthic macrofauna species occurrences and abundance	Species occurrences and abundances are available in the MedOBIS IPT and can be harvested in EMODnet Biology once public.
Microbial diversity of the Cretan Sea between 2016 and 2017	Benthic microbial species occurrences and abundance	Species occurrences and abundances are available in the MedOBIS IPT and can be harvested in EMODnet Biology once public.
Physicochemical characterization of the Cretan Sea between 2016 and 2017	pH or pCO2, Temperature, Salinity, Chlorophyll (Fluorescence), O2	These data can be integrated in EMODnet Biology if it is associated to species occurrences.

6. Conclusions and recommendations

JERICO-NEXT has undertaken substantial efforts to integrate biological data into the marine observation networks in order to address pelagic and benthic biodiversity questions. These data can be obtained with traditional sampling methods but, within JRAP1 and JRAP2, innovative observation techniques and novel sensor data have been explored, revealing a significant potential to deliver operational (near real-time) data usable for monitoring purposes.

In the framework of JERICO-NEXT, datasets from JRAP1 and JRAP2 have been described and metadata records have been made available in the EMODnet Biology catalogue and through a map-viewer interface, that displays the geographical extension of the collected data. The different datasets were screened and considered with respect to their current conformance with the biodiversity data standard (i.e. DwC). Data from 18 datasets has been made accessible using different approaches: a) through a data access link to a local/ad hoc website; b) trough documented archiving in the Nature *Scientific Data* recommended repository: Marine Data Archive; or, c) through EMODnet Biology by means of full standardization and integration into the EurOBIS database. Fully standardized data can be checked using automatic quality control procedures developed in the framework of EMODnet Biology (i.e. online QC tool).

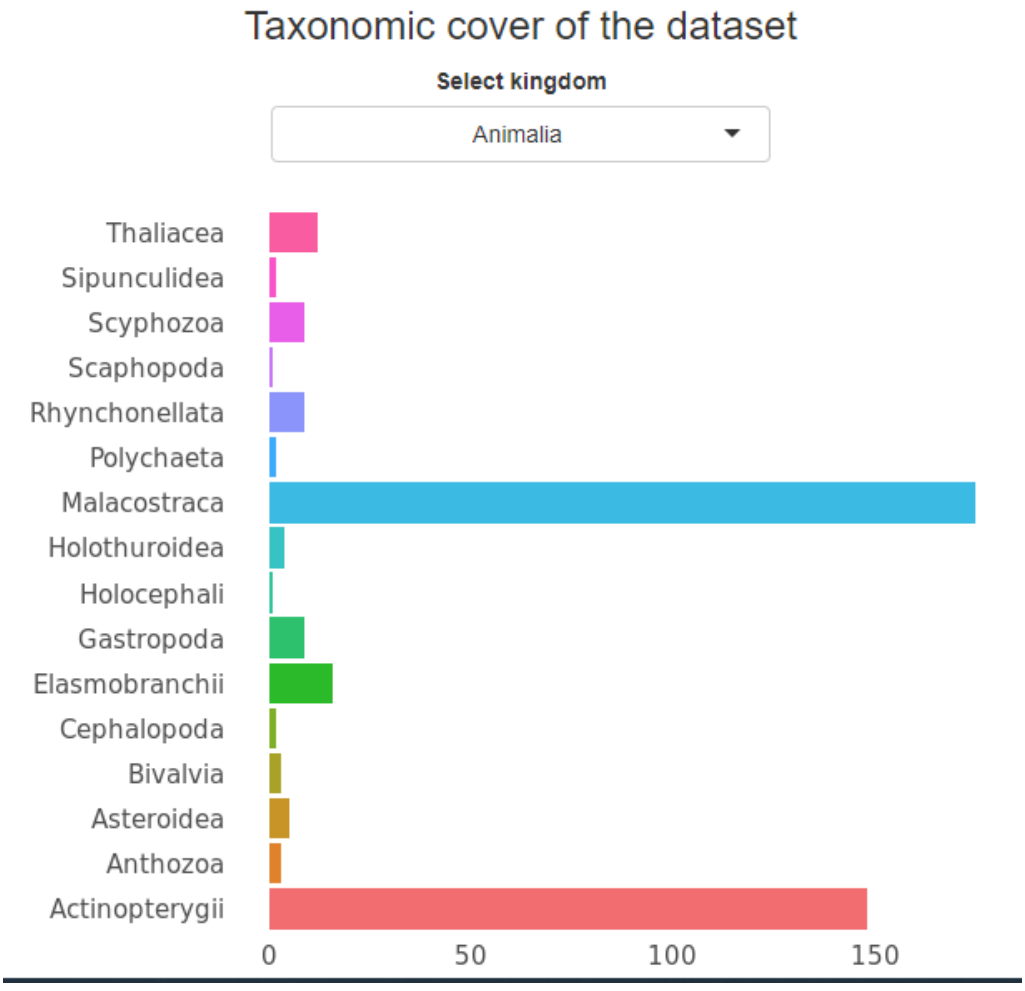
Integration of biological data is relatively straightforward for data that is collected with traditional sampling methods and/or that is taxonomically resolved. However, some of the observation methods deployed in JRAP1 and JRAP2 are in an earlier stage of development and present a great variety of outputs due to a lack of standardization in different phases of the data management cycle: from data collection, (e.g. sensor calibration), to data publication and quality control (e.g. lack of syntactic and/or semantic data standards). Some of these issues have seen a considerable improvement during the JERICO-NEXT project. For example, in relation to WP5 a set of controlled vocabularies has been developed and put into practice for FCM data, and some JRAP1 FCM datasets are now integrated into SDN, EurOBIS and EMODnet Biology. Technical developments were put forward in the EMODnet Biology portal to allow users to search for FCM data. The adoption of new sensor-based biodiversity observations brings further complexities, and additional work on harmonization of data collection and definition of best data management practices is still needed for biological data sensors. This will be a priority in the following phase of the JERICO RI. Through its direct interaction to wider biodiversity data sharing networks such as OBIS or TDWG, EMODnet Biology can provide an interface to translate the needs of the observing community for exploration and documentation of data exchange formats for sensor-based biodiversity data.

The need of developing data management best practices for novel data acquisition methods is widely agreed by the marine observation community. However, it is equally important to keep in mind that it is only by implementing these practices that we can guarantee to collect data of good quality that is FAIR and open. All the different phases of data management, whether at the level of the scientist (i.e. data collector) or at the level of the data manager / aggregator, require economic and/or time resources and need to be appropriately funded. Awareness on the importance of designing data management plans is rapidly increasing and data management is now an eligible cost in many funding mechanisms. However, the lack of appropriate resources continues to be a bottleneck that hampers the publication of data, more specifically interoperable data, after its collection and analysis.

7. Annexes

7.1. Annex 1: Overview of the EMODnet Biology / EurOBIS QC tool and the QC-steps

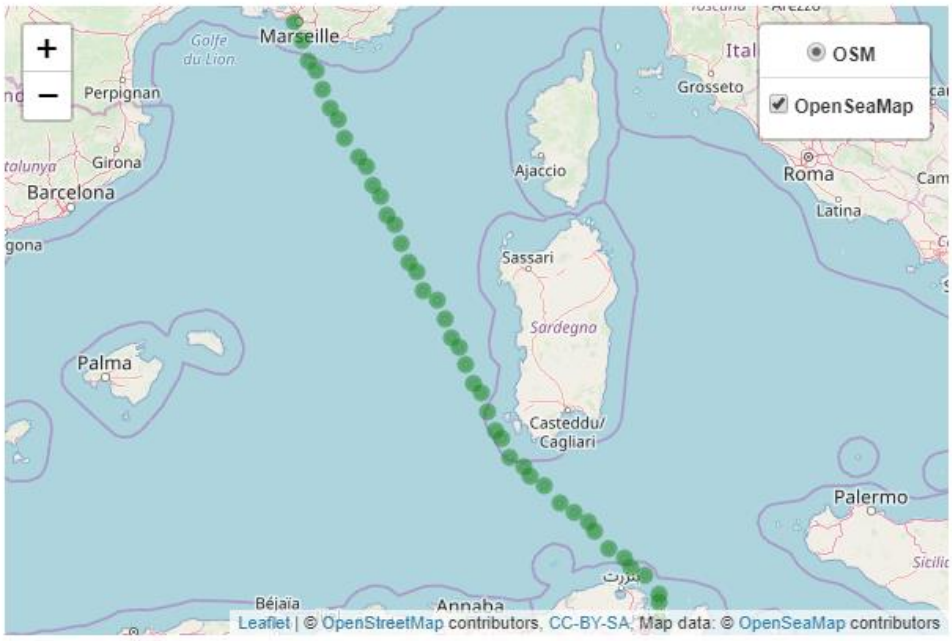
The online QC tool developed in the framework of EMODnet Biology and EurOBIS (LifeWatch) is an R-Shiny based interface that gives a quality report for IPT-based resources. The user can copy-paste the URL of a dataset on IPT and will get a report giving an overview of the data (e.g. total number of records, taxonomic, geographical and temporal coverage, parameters provided in the data, etc.) and issues to be tackled or suggestions for further standardization of the data (e.g. a parameter provided that has not been mapped to a controlled vocabulary). Some examples of the outputs of the online QC tool are given in the figures below. The tool is available at: http://rshiny.lifewatch.be/BioCheck/ and the full list of QC checks is provided in this annex.



Overview of all issues

field	message	count	table
measurementType	Duplicate eMoF record linked to event	26	emof
measurementValue	Biological value of 0 while Occurrencestatus is present	3	emof
coordinates_error	Coordinates are located on land	2	event
eventDate	eventDate does not seem to be a valid date	2	event
maximumDepthInMeters	Depth value is greater than the value found in the bathymetry raster	4	event
minimumDepthInMeters	Recommended field minimumDepthInMeters is missing	1	event
occurrenceStatus	Empty value for required field occurrenceStatus	3	occurrence
scientificName	Empty value for required field scientificName	1	occurrence
scientificNameID	Marine taxon located on land	43	occurrence
scientificNameID	scientificNameID does not resolve	5	occurrence
scientificNameID	Empty value for required field scientificNameID	1	occurrence

Geographical cover of the dataset



List of QC checks:

- Is the dataset integrity ok?
- Do all eventID's in the occurrence extension refer to an Event Record?
- Do all eventID's in the eMoF extension refer to an Event Record?
- Do all occurrenceID's in the eMoF extension refer to an occurrence Record?
- In case of a biometrical parameter, is the eventID in the eMoF link the same one as in the occurrence extension?

- Are there 'duplicate occurrences' meaning is the same taxon listed twice at the same EventID without any difference in any of the biometric parameters?
- Are there 'duplicate measurements' meaning does the same measurement occur twice for the same occurrenceID or the same EventID?
- Are all mandatory fields present in the dataset?
- Are all mandatory field filled out?
- Does eventDate follow the required format?
- Does the scientificNameID follow the required format?
- Are there coordinates located on land? (buffer of 3km is taken under consideration)
- Do the coordinates on land refer to marine taxa?
- Are there depths at the location deeper than the depths stored by EMODnet Bathymetry (Europe) and GEBCO (rest of the world)? (a margin of 150m is taken under consideration)
- Do all measurementtypes have a MeasurementTypeID?
- Does the measurementTypeID / measurementValueID refer to an existing term in the BODC vocabulary?
- Do all measurementValues that refer to facts have a measurementValueID?
- Are there records where measurementValue is NULL?
- Are there records that refer to biological measurement where measurementValue = 0 (and where occurrenceStatus is not absent)?
- Provide overview of the measurementTypes, their units, the min and max values and the prefered labels, definitions and standard units associated to the MeasurementTypeID
- Is there a sampling instrument present?
- Are there other sampling descriptors present?
- Provide an overview of the number of taxa per kingdom and class.
- List the non-matched taxa (including deleted and quarantined matches)
- Plots of the coordinates and the distribution of the temporal cover are provided to allow for quick comparison with the metadata.
- A tree view of the event hierarchy to inspect the structure.