



EMODnet



European Marine
Observation and
Data Network

EMODnet Thematic Lot n° VI – Biology

EMFF/2019/1.3.1.9/Lot 6/SI2.837974 -EMODnet -Biology

Start date of the project: 18/04/2021 (24 months)

Centralisation Phase

Quarterly Progress Report (QR24)

Reporting Period: 01/01/2023 – 31/03/2023

Contents

1. Highlights in this quarter	3
2. Identified issues: status and actions taken.....	11
3. Communication assets	14
4. Monitoring indicators.....	17
5. Annex: Other documentation attached	19

Disclaimer

The information and views set out in this report are those of the author(s) and do not necessarily reflect the official opinion of the CINEA or of the European Commission. Neither the CINEA, nor the European Commission, guarantee the accuracy of the data included in this study. Neither the CINEA, the European Commission nor any person acting on the CINEA's or on the European Commission's behalf may be held responsible for the use which may be made of the information.

1. Highlights in this quarter

Task 1: Maintain and improve a common method of access to data held in repositories

The main aim of the Feasibility Study on Ecological Traits and sampling devices/methodologies identification with text mining is to assess the accuracy of text mining prototypes regarding these entities; advice on next actions; examine whether further development and fine tuning should be sought. More specifically, this study focuses on ecological traits, life stages and body length extraction with text mining and is built upon previous work on traits and text mining of EMODnet Biology. According to the results of the feasibility study, it is concluded that the most suitable models for the data management perspective are SpaCy and BioBERT. To further enhance the accuracy of these tools, an enriched workflow is recommended, consisting of the following four steps: a. Corpus Enrichment and Text Annotation, b. Combining GNfinder with an NLP library, c. Management of the text mining process and d. Visualisation of results.

The full report related to this deliverable, will be published before the end of Phase IV, during the first two weeks in April.

In the framework of this deliverable, a free online training session was organised on March 20th 2023. This training session focused on the interdisciplinary field of text mining and its application on marine ecological traits and was open to marine ecologists, curators, data managers, students who are doing or planning to start research on curation and information extraction of marine ecological literature. During the session, which was advertised through the EMODnet Central Portal; HCMR and EurOBIS social media, as well as the [GitHub of the Open Traits Network](#); participants received an overview of state-of-the-art text mining technologies and bioinformatic tools. The number of registered participants was 32; originating from several Institutes/Organisations, such as: Instituto de Hidráulica Ambiental de la Universidad de Cantabria, University of Salento, University of Crete, Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Portuguese Institute for Sea and Atmosphere (IPMA) etc.

During the reporting period, 10 datasets from partners and from ad-hoc data providers were updated. This harvest brought in an additional 104.149 occurrence records. A list of the published datasets and their details is available in the Annex. In addition, another data harvest is planned to be finished shortly after the end of phase IV which will bring another 83 datasets to EMODnet Biology from both partners and ad-hoc data providers.

Task 2: Construct products from one or more data sources that provide users with information about the distribution and quality of parameters in time and space

Various partners submitted the products and three have been published in this Quarter:

- Species Distribution Models for invasive macroalgae (Ramos et al. 2023, <https://www.vliz.be/en/imis?dasid=8209>)
- Presence/absence data of macrozoobenthos in the European Seas (Herman 2022 <https://www.vliz.be/imis?dasid=8216>)
- R package: Btrait 'Working with biological density, taxonomy, and trait composition data' (<https://www.vliz.be/en/imis?module=dataset&dasid=8217>)

Associated with the product publication, news items were also published to highlight their objective and availability via the Central Portal. The remaining products will be published until the end of Phase IV and a more detailed description will be included in the final report.

The initial plan was to publish 7 products, but due to unforeseen events affecting the partner in question, one product will not be delivered and published before the end of Phase IV. This product, alongside two external products with a focus on birds, mammals and plankton genomics, will be published shortly after the end of the Phase.

WP3 partners have been drafting the position paper (D3.5) over a series of online workshop sessions in February and March 2023. The title is 'Realising the potential of interoperable data products to improve the future outlook for marine biodiversity: lessons from the European Marine Data and Observation Network', with a target journal of Marine Policy. This publication is aimed for after the end of Phase IV.

Task 3: Develop a complete and robust machine to machine (M2M) interface to transfer data and products in bulk, which is easily accessible for other machines and initiatives

A number of technical actions took place during the reporting period, including:

- Implementation HTTP redirection from emodnet-biology.eu to the EMODnet Central Portal, on a per-URL basis as a result of the centralisation and shutdown of the Biology portal;
- Completion of the data publication changes which included the addition of 16 new fields in the data harvest system, all intermediate databases, the production systems. This allows for the publication of datasets with associated media;
- IPT hosted instances were upgraded to version 2.7.2;
- Fixed bugs related to GeoServer log growth and improved the GeoServer restarting policy;
- Revisited fitness for use label on the published data to finalise the last implementation steps in 2023;

Additionally, the first implementation of the semi-automated procedures to query, request, download and reformat data from SeaDataNet's API has been completed and is now operational. This means that we will be able to more quickly publish biological data available through this aggregator, while spending less time with the reformatting. Constraints are still in place when it comes to vocabulary mapping which cannot rely on a fully automated process, but require manual intervention by a data manager.

Task 4: Contribute data, data products and content to a central portal that allows users to find, view and download data and data products

In late January, centralisation reached a stage where all thematic lots shutdown their portals. This meant that all Biology data/products/content became available via the EMODnet Central Portal.

The publication of data is an invisible process to users, the layer EurOBIS database observations in the map viewer is updated every time data are published. More details about the publication of data can be found in Task 1.

Product publication is also done either via the addition of new layers in the viewer, or by the update of content, if referring to new tools or services. More details about the product publication can be found in Task 2.

Content updates have been requested to the EMODnet Secretariat and done within the agreed deadlines. More details about the content can be found in Task 5.

Task 5: Contributing content to dedicated spaces in Central Portal

Various news items and one use case were published in the Central Portal during the reporting period. The news items covered mostly the publication of products and a status updated of the project was included in the Jan-Mar newsletters. The published news can be found via the links below:

- [New Biology outputs: products, use case, data](#) (Published on 2023-03-22)
- [New EMODnet biology data product: Species distribution model of invasive macroalgae](#) (Published on 2023-03-15)
- [New EMODnet biology data product: Presence/absence of macrobenthos in European Seas](#) (Published on 2023-03-15)
- [Training session in the framework of the EMODnet Biology Phase IV on the feasibility study for recognition of specific ecological traits and sampling devices/methodologies in text](#) (Published on 2023-03-01)
- [EMODnet Biology contributing to the TRASMARES itinerary \(Massive Open Online Courses\)](#) (Published on 2023-02-17)
- [EMODnet Biology reaches 30 million data point occurrences, and other highlights from 2022](#) (Published on 2023-01-26)

A [use case](#) referring to the publication of a dataset with standardised associate media information was published on 2023-03-14. It highlights the technical developments needed to publish this type of information and the impact it will have for EMODnet Biology. More datasets containing this information will be published in the coming months.

Task 6: Ensure the involvement of regional sea convention

EMODnet Biology has identified the need for a flexible, bespoke level of engagement with each of the Regional Sea Conventions, due to differing levels of technical support, data product level requirements and mechanisms for interaction. As such we have previously identified EMODnet Biology 'ambassadors' within organisation already engaged with the 4 RSC's through D4.2 - the 'connectivity map'. In this reporting period we convened an internal meeting of these 'ambassadors' to review progress and undertake a lessons-learned exercise to inform future EMODnet Biology engagement with, and support for, the RSC's.

Task 7: Contribute to the implementation of EU legislation and broader initiatives for open data

EMODnet Biology and its partners form a key part of the Horizons Europe funded MARCO-BOLO (Marine Coastal Biodiversity Long-term Observations) project. The MARCO-BOLO project aims to improve the flow and utility of marine biodiversity observations, including those coming from autonomous sources and eDNA derived occurrence data. EMODnet Biology provides a key role in the promotion of data standards and ensuring the long-term accessibility of the mobilised

data. The kick-off meeting for MARCO-BOLO took place 14-16th March, with strong EMODnet Biology representation across the whole consortium.

Task 8: Monitor quality/performance and deal with user feedback

Four user requests/feedback were received either directly to the biology email or via the EMODnet JIRA. Details can be found in the Table on Section 3.

EMODnet Biology's OGC webservice are continuously monitored via the OGC Health check implemented by VLIZ for the Central Portal. An analysis of Figure 3 indicates that the lot's services were always within INSPIRE guidance, with a 100% operationality and 99.6% reliability. It is noted that due to not being able to select a specific time range, the information present covers the period from 2022-04-04 to 2023-04-06.

Due to centralisation and the shutdown of the Biology portal, all analytics are now been collected via the Europa Analytics instance. The information collected is included in the Indicators spreadsheet

Task 9: Maintain the existing thematic web portal for a maximum of six months from the start of the projects

The portal was kept until late January, after which it was shutdown following the completion of the EMODnet centralisation.

Status of the Milestones and Deliverables listed in the workplan					
Milestone/Deliverable	WP	Date due	Status (Delivered/Delayed)	Date delivered	If Delayed: reason for delay and expected delivery date
D1.1: Help desk operational and contact details published online (M3)	WP1	QR17 (30 Jun 2021)	Delivered	30/06/2021	
D1.2 Quarterly progress reports (scheduled defined in the tender) (M3-M24)	WP1	QR17 (15 Jul 2021)	Delivered	15/07/2021	Due to EMODnet centralisation
		QR18 (15 Oct 2021)	Delivered	15/10/2021	
		QR19 (15 Jan 2022)	Delivered	15/01/2022	
		QR20 (15 Apr 2022)	Delivered	15/04/2022	
		QR21 (15 Jul 2022)	Delivered	15/07/2022	
		QR22 (15 Oct 2022)	Delivered	14/10/2022	
		QR23 (15 Jan 2023)	Delayed	15/02/2023	
QR24 (15 Apr 2023)	Delivered	15/04/2023			
D1.3 First interim report after phase (M12)	WP1	19/04/2022	Delivered	17/06/2022	
D1.4 Final report (M24)	WP1	19/04/2023			

D1.5 Attend Steering Committee and Technical Working Group meetings (M0-M24)	WP1	Sep 2021 Apr 2022 Sep 2022 Nov 2022 Apr/May 2023	Delivered Delivered Delivered Delivered Pending	08/09/2021 26- 28/04/2022 21- 22/09/2022 07- 08/11/2022	
D2.1 Inventory of possible historical data resources within the consortium (M6)	WP2	29/10/2021	Delivered	18/11/2021	
D2.2 Technical implementation of data flows for the new project partner (M6)	WP2	29/10/2021	Delivered	18/11/2021	
D2.3 Report on efforts undertaken in rescuing historical data through citizen science (M18)	WP2	15/10/2022	Delivered	01/12/2022	
D2.4 At least 3 linkages with initiatives outside of the original Consortium, resulting in extra data/information available via the Portal (M20)	WP2	31/12/2022	Delivered	31/12/2022	
D2.5 Feasibility study for recognition of specific ecological traits and/or sampling devices/methodologies in text (M22)	WP2	25/02/2023	Delayed		Will be delivered before end of Phase
D2.6 Report on the standardisation and integration of the proposed new and updated datasets (M24)	WP2	19/04/2023	In progress		

D3.1 Quarterly WP3 community calls; call leader or other nominated team member to produce summary report of each call for publication on EMODnet website (M3-M24)	WP3	Call QR17 (30 Jun 2021)	Delivered	18-19/05/2021	
		Call QR18 (30 Sep 2021)	Delivered	22/09/2021	
		Call QR19 (31 Dec 2021)	Delivered	10/12/2021	
		Call QR20 (31 Mar 2022)	Delivered	Week of 28/02/2022	
		Call QR21 (30 Jun 2022)	Delivered	Week of 02/05/2022 +	
		Call QR22 (30 Sep 2022)	Delayed	18/07/2022 20/10/2022	
		Call QR23 (31 Dec 2022)	Delayed	07/02/2023	
		Call QR24 (31 Mar 2023)	Delivered	02/03/2023	
D3.2 Annual intensive workshops, in person with online participation options. Workshops involve collaborative product development on one or more targeted themes derived from WP4 user needs questionnaire (M12-M22)	WP3	Workshop 1 (19 Apr 2022)	Delivered	Week of 28/02/2022	
		Workshop 2 (28 Feb 2023)	Delivered (workshop transformed to multiple WP calls to draft D3.5)	Various dates	
D3.3 Publish R package to link EMODnet biology data with data from other EMODnet sources (M12)	WP3	19/04/2022	Delivered		
D3.4 Develop method to use Phase III presence-absence maps to display time series of distribution change (M12)	WP3	19/04/2022	Delivered		
D3.5 Produce position paper outlining questions that can be addressed using EMODnet data, together with remaining gaps, and	WP3	19/04/2023	In Progress		

strategies for filling these (M24)					
D3.6 Add/update data product metadata in the EMODnet Biology catalogue (M24)	WP3	19/04/2023	In progress		
D4.1: Questionnaire to inform cross-lot product development (M3)	WP4	30/07/2021	Delivered	06/09/2021	
D4.2: EMODnet Biology connectivity 'map' of projects, institutes, initiatives and networks to inform targeted engagement (M12)	WP4	19/04/2022	Delivered		
D4.3 EMODnet Biology participation in each of the RSCs to inform and advise of available data products and mechanisms to access and influence the development of data, products, tools and services (M12)	WP4	19/04/2022	Delivered		
D4.4 "Launch" of the European MBON node (M24)	WP4	19/04/2023	To be delivered		See QR23 Task 7
D4.5 Creation of engaging and informative use-cases for EMODnet Biology to illustrate uptake and utility of data products across a range of stakeholders across the quadruple helix of engagement (M6, M12, M18, M24)	WP4	Use case 1 (29 Oct 2021) Use case 2 (19 Apr 2022) Use case 3 (15 Oct 2022) Use case 4 (19 Apr 2023)	Delivered Delivered Delivered	31/01/2022 01/04/2022 14/03/2023	
D5.1 User portal operation and maintenance (M0-M6)	WP5	19/10/2021	Delivered		

D5.2 Webservice operation and maintenance (M0-M24)	WP5	19/04/2023	In Progress		
D5.3 Technology stack upgraded (M12)	WP5	19/04/2022	Delivered		
D5.4 Evaluation and implementation of bulk data transfer technologies (M24)	WP5	19/04/2023	In Progress		

2. Identified issues: status and actions taken

A. Priority issue(s) identified and communicated by CINEA/ DG MARE/ SECRETARIAT				
Priority issue	Status (Pending/ Resolved)	Action(s) taken/ remaining actions planned	Date due	Date resolved
Including in metadata records online resources (This will be provided as part of the comprehensive metadata review, planned for end 2022. This will however require a CSW service that provided access to the Portal's metadata records)	Pending	No changes since the last QR	QR24	
Including data and metadata URL in web services of data products	Ongoing	No changes since the last QR	QR18	
Progress solving identified issues				
EM-83 - Biology - Web Services MetadataUrl and DataUrl	Pending	Resolution is planned to take place in the second half of 2021 or early 2022		
EM-78 - Biology to report on status or plans to support INSPIRE Compliant CSW	Ongoing	Customised GeoNetwork (DCAT and INSPIRE compliant) upgrade in progress. Fixing and debugging is almost complete, however the Log4j bug that was identified worldwide at the end of 2021 also needs to be addressed for this piece of software		

A. Issues / challenges identified by the thematic assembly group itself				
Priority issue / challenge	Status (Pending/ Resolved)	Action(s) taken / remaining actions planned	Date due	Date resolved
Users do not have access to absence occurrences	Ongoing	Species' absence in surveys can provide invaluable information for data product creation and overall ecosystem knowledge. Within the thematic lot, there are several datasets where data providers have reported true absences (species were	QR22	

		looked for but not found) or derived absences (i.e a list of species was defined for the project and systematically used throughout. When a species is absent from the data submitted, we know that this is because such individuals were absent from the samples). Work is underway to allow for these data to be available to users		
Users do not have access to image data associated with occurrence records	Completed	As part of Phase IV proposed work, changes to the Database and data publishing procedures are necessary to allow users to obtain images associated to the taxa occurrence records.	QR24	QR24
Users do not have access to genomics data	Ongoing	In the Phase IV workplan it was stated that the thematic lot would endeavour efforts to implement procedures to handle and publish genomics data. This work is not meant to be done in isolation with the international community therefore most actions have been limited to accompanying the situation and internally assessing what changes need to occur in the DB so that these data can be properly managed.	Phase V	
Delays due to situation in the Ukraine	Ongoing	The partner in questions, UkrSCES is based in Odessa and contributes to WP2. Situation is being followed up within the project. The ongoing conflict might mean that the partner is unable to supply the data within the project's deadlines, as initially agreed		
Thematic lot metadata records not in the EMODnet catalogue	Ongoing	Due to centralisation, the lot will lose its catalogue, where users can search and download data. Work is ongoing to allow for a sync	QR24	

		between the lots' catalogue (hosted by VLIZ) and the EMODnet GeoNetwork instance to ensure all metadata remain FAIR		
Users do not have access to extended measurements and facts data	Pending	Due to centralisation users querying the EurOBIS observations layer only have access to occurrence data and no associated parameters. Technical developments are needed to ensure that the Central Portal viewer replicates the same functionalities and provides access to the same data as via the Biology portal		

3. Communication assets

A. Communication products developed				
Date	Communication material	Short description (of the material, title, ...) of the asset	Main results	Name of event at which material was disseminated (if applicable)
QR24	Tweets	25 tweets on published datasets, data management tips and activities by the EMODnet Biology team	Increase exposure to activities and outputs, Outreach	NA
QR24	News	Articles published in the EMODnet Central Portal, see Task 5	Increase exposure to activities and outputs, Outreach	NA
QR24	Use case	See Task 5	Increase exposure to activities and outputs, Outreach	NA

Your Tweets earned **9.1K impressions** over this **90 day period**

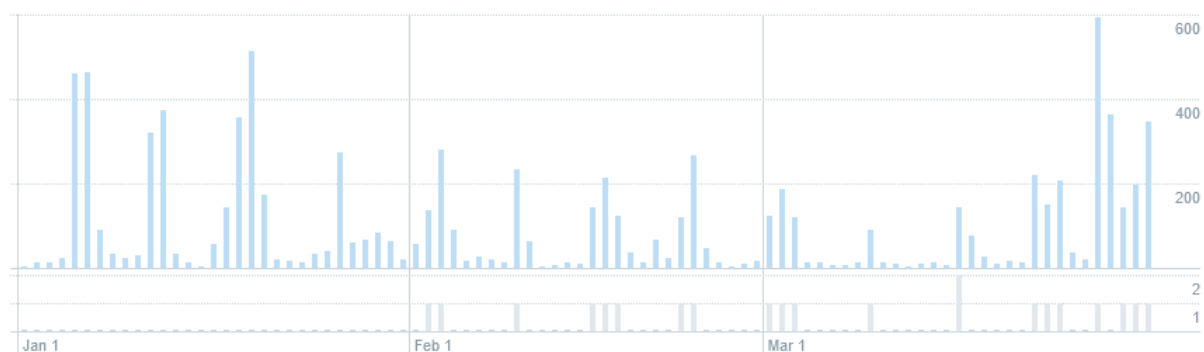


Figure 1. Twitter Impressions for @EurOBIS_VLIZ Twitter account during the reporting period

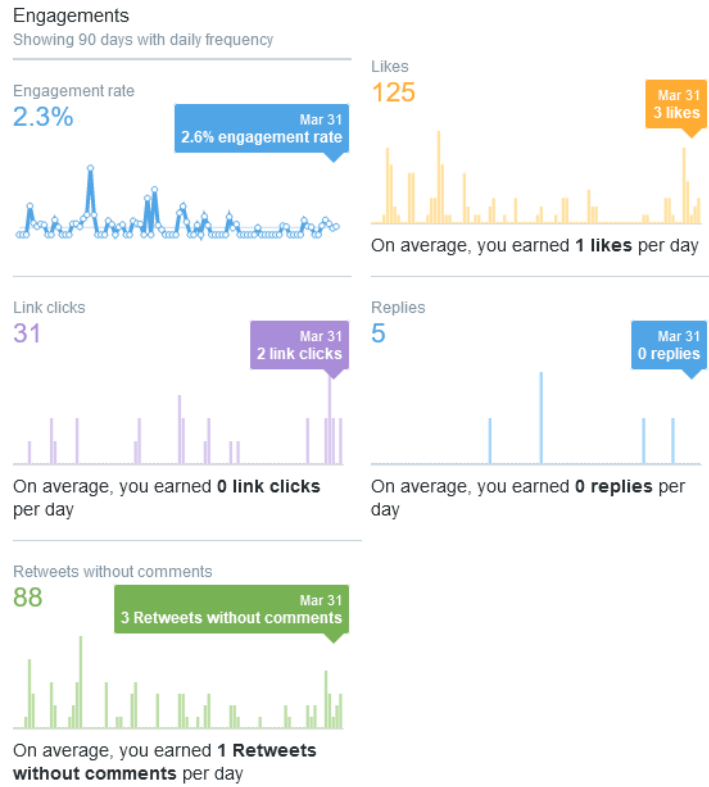


Figure 2. Twitter Engagements for @EurOBIS_VLIZ Twitter account during the reporting period

B. Planned communication products			
Date	Communication material	Short description (of the material, title, ...) and/or link to the asset	Main results expected
There is no contract signed for Phase V, so there are no outputs planned for QR25			

A. (Co-)Authoried peer-reviewed publications in the quarter

Date of publication	Type of publication	Full reference	ISBN	DOI	Is it open access? Yes/No
NA					

B. Other/non-peer reviewed types of publications (co-)authoried in the quarter

Date of publication	Type of publication	Full reference	ISBN	DOI	Is it open access? Yes/No
NA					

For a compressive overview of publications referring to/making use of EMODnet data and/or data products, please consult Google Scholar.

4. Monitoring indicators

Comments on the progress indicators in the indicators spreadsheet		
Progress indicator	Means of collecting figures	Comment
1. Current status and coverage of total available thematic data A) Volume and coverage of available data	In-house systems	Due to the updated datasets that were published, the volume of data increased slightly across most functional groups, except for birds. The published datasets covered only two regions, the Atlantic and Baltic.
What is your opinion on the data coverage within EMODnet for your thematic? B) Usage of data in this quarter		No change since last QR Due to centralisation, there is no information about the data usage/downloads for this quarter. All data are being served via GeoServer and there are no systems in place to capture relevant information for this indicator
2. Current status and coverage of total number of data products A) Volume and coverage of available data products	In-house systems	Six new products were finalised, but only three were published during the reporting period. One of these was an R Package and the other two a (macro)algae and a benthos focused product, both of which covered all European Seas.
B) Usage of data products in this quarter		Due to centralisation, there is no information about the data usage/downloads for this quarter. All data products are being served via GeoServer or ERDDAP and there are no systems in place to capture relevant information for this indicator
3. Internal and external organisations supplying/approached to supply data and data products within this quarter		One dataset was from an organisation outside the consortium, the remainder were from partners.
4. Online 'Web' interfaces to access or view data		Due to centralisation several changes occurred. Data and products are now available via the EMODnet viewer. OGC services for data are hosted by the VLIZ geoserver instance and the main data catalogue is available via the EurOBIS website. OGC webservice for data products are available from both the VLIZ and EMODnet Geoserver instances. Products are also available via the EMODnet ERDDAP instance. The Data Products metadata is available via the VLIZ catalogue. Products code, input and outputs are also available via the EMODnet GitHub

5. Statistics on information volunteered through download forms		NA
6. Published use cases		NA
8.1. Technical monitoring		NA
9. Visibility & Analytics for web pages		NA
10. Visibility & Analytics for web sections		NA
11. Average visit duration for web pages		NA

The monitoring numbers reported as part of the progress monitoring of EMODnet performance are collected through Matomo and/or Europa Analytics, unless reported otherwise.

5. Annex: Other documentation attached

Task 1: Maintain and improve a common method of access to data held in repositories

Dataset title	New/Update	Phase	Organisations	License
SHARK - National marine bacterioplankton monitoring in Sweden since 1989	Update	IV	Partner (SMHI)	CC 0 (No Rights Reserved)
SHARK - National phytoplankton monitoring in Sweden since 1983	Update	IV	Partner (SMHI)	CC 0 (No Rights Reserved)
SHARK - National zoobenthos monitoring in Sweden since 1971	Update	IV	Partner (SMHI)	CC 0 (No Rights Reserved)
SHARK - National marine environmental monitoring of Ringed seal in Sweden since 1995	Update	IV	Partner (SMHI)	Attribution (CC BY)
SHARK - Regional monitoring, recipient control and monitoring projects of zoobenthos in Sweden since 1972	Update	IV	Partner (SMHI)	CC 0 (No Rights Reserved)
Explore Your Shore	Update	NA	Ad-hoc (NBDC)	Attribution (CC BY)
MAREANO - Base-line mapping of epifauna obtained with Beamtrawl	Update	IV	Partner (IMR)	Attribution (CC BY)
MAREANO - Base-line mapping of fauna obtained with grab	Update	IV	Partner (IMR)	Attribution (CC BY)
MAREANO - Base-line mapping of hyperbenthic crustacea fauna obtained with RP-sledge	Update	IV	Partner (IMR)	Attribution (CC BY)
Type locality distributions from the World Register of Marine Species	Update	IV	Partner (VLIZ)	Attribution (CC BY)

Task 8: Monitor quality/performance and deal with user feedback

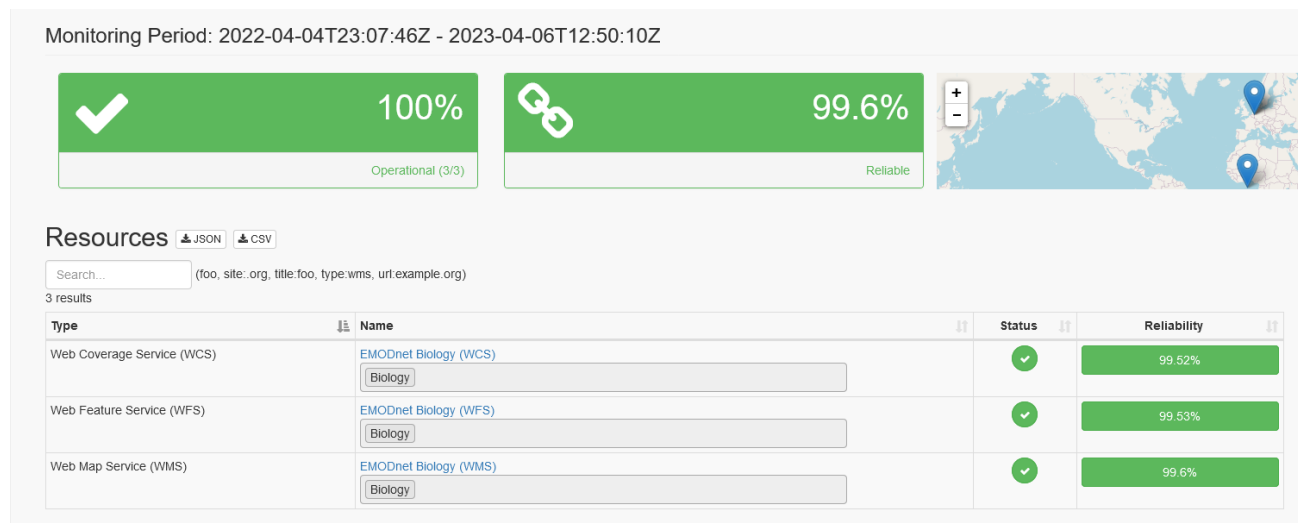


Figure 3. GeoHealthCheck monitor for EMODnet Biology OGC services for the period between 2022-04-04 to 2023-04-06