EMODnet ingestion and safe-keeping of marine data

CINEA/EMFAF/2021/3.4.10/02/SI2.868290

Start date of the project: 30/03/2022 (48 months)

Centralisation Phase

Quarterly Progress Report (10)

Reporting Period: 01/07/2024 – 30/09/2024

The European Marine Observation and Data Network (EMODnet) is financed by the European Union under Regulation (EU) 2021/1139 of the European Parliament and of the Council of 7 July 2021 establishing the European Maritime, Fisheries and Aquaculture Fund.

Contents

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, further develop and migrate a web-portal

Migration to Central Portal, Submission service:

The migration implementation plan, which includes the division between the backend and frontend components, has progressed further. The backend, encompassing the database, cloud storage, and API, will remain on the servers of HCMR. In contrast, the frontend will be hosted on the CP servers at VLIZ, for both testing and production environments.

Actions:

- The integration of ECAS as the authentication mechanism has progressed. The communication with the ECAS authentication endpoints has been established, by addressing issues with the available EU software libraries, including the compatibility with latest stable PHP versions. Development efforts are now focused on adjusting the MarineID centric backend model, as well as user/role management, to ECAS user schema.
- A request has been submitted to grant access to ECAS endpoints for the additional development environment (https://ingestion2.getmap.gr/ingestion/) which has been established to facilitate broader testing of the new API-based approach by members of the Ingestion consortium.
- Testing is ongoing to ensure the new application functions as intended. This includes unit tests, integration tests, and user acceptance testing.
- The operational testing environment, available at https://submission-test.emodnet-ingestion.eu/, remains unchanged until the migration procedure is fully accepted and operational.

Further progress was made with drafting the narrative of EMODnet Ingestion, that will be published at the EMODnet Central Portal. Following the earlier defined structure, texts are drafted for each section, trying to comprise and cover major content of the stand-alone website. Once ready, the full narrative will be discussed at a CP migration meeting for following publishing at the EMODnet development server.

Progress was made with a new map for as-is stations, which are the result of the joint activities of EMODnet Physics and EMODnet Ingestion, aiming at identifying, convincing, and connecting new NRT data providers and their NRT data streams from operational stations. The new map layer has been added to the Central Map Viewer and is driven by a new EMODnet Physics web service.

Summarizing, the migration is now depending on completing the deployment of the new submission service, the integration with ECAS AAI service, the migration of existing users from Marine-ID to ECAS accounts, and finalizing the narrative.

Task 2: Implement pathways for delivering data to final repositories

The total number of received submissions increased from 1681 to 1745, while the number of processed and published data submissions increased from 1549 to 1583, and of which, the number of fully elaborated data submissions went from710 to 727 data. The KPI excel sheet provides more details.

Task 3: Facilitate machine-to-machine transfers

During the reporting period, the team continued working on ingesting new data sources, supporting the CP to include the Ingestion activities and layers. The inclusion of the source "as is" (phase one) involves the adoption of machine-to-machine technologies, ranging from FTP file access to web APIs for real-time exchange. It also involves the adoption of services like the Data Access Broker (DAB) or ERDDAP.

The “ingestion-erddap.emodnet-physics.eu” platform provides users with an easy way to interact with the latest ingested real-time (RT) collections, which reached 2342 datasets during the reporting period. However, not all of these datasets are immediately visible in the Central Portal under the new "Ingestion – as is" section. New data streams need to be integrated into the metadata-controlled collections, which may take some time. Once a new dataset is integrated into these controlled collections, it is offered to the Central Portal under a specific WMS layer.

Error! Reference source not found. shows the newly connected stations that are coming from four main sources covering BioArgo (ARGO GDAC), diving loggers (SSI), and Atmospheric station in New Zealand. Some SSI diving profiles are more inland than the primary focus of EMODnet; therefore, these data will not be considered for integration into the data collections.

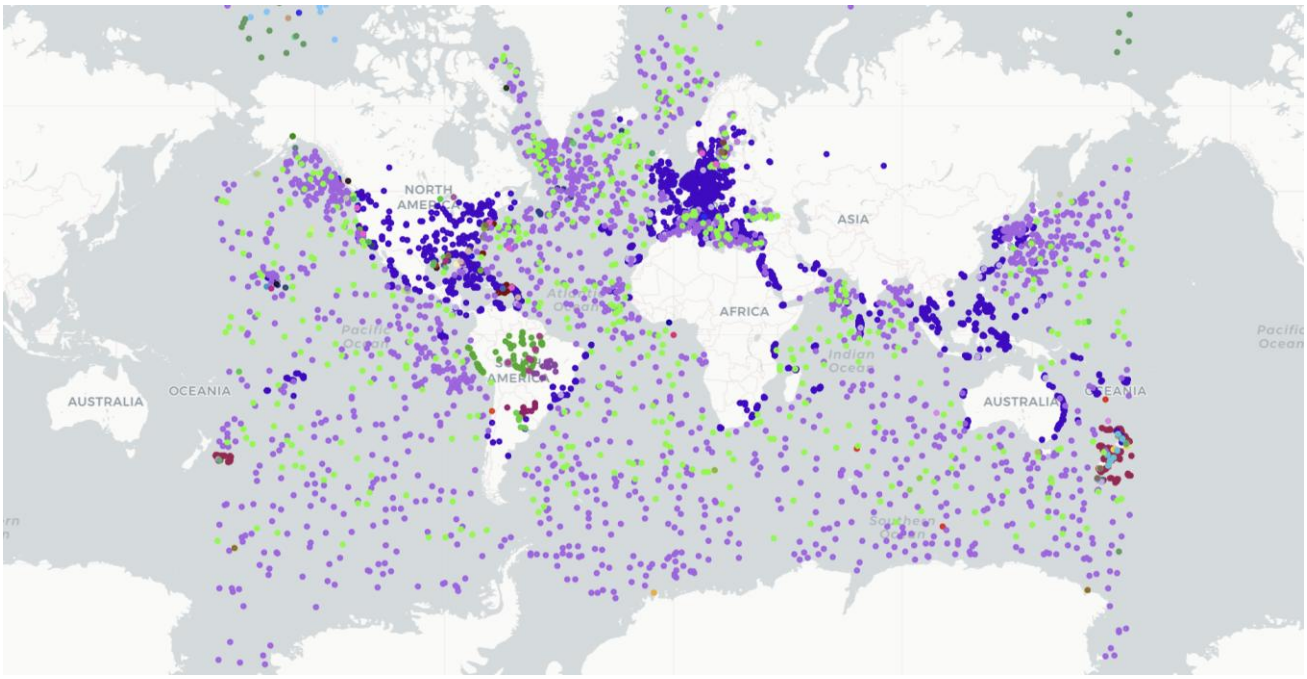

Figure 1. newly integrated platforms[1]

Some data were directly integrated into the CMEMS INSTAC validated collections, mostly coming from Coriolis (GDAC). Notably, when CMEMS directly includes data into operational collections, it may happen that metadata are not fully updated. As described in the previous report, some data that are integrated in Physics via the CMEMS – CORIOLIS/GDAC may miss some metadata that Physics workflow need for proper management. These new data are now mapped and counted “as is”.

[1] https://data-erddap.emodnet-physics.eu/erddap/tabledap/EP_PLATFORMS_METADATA.csv?PLATFORMCODE%2Ccall_name%2Clatitude%2Clongitude%2CdataFeatureType%2CfirstDateObservation%2ClastDateObservation%2Cdata_owner_longname%2Cdata_owner_country_code%2Cdata_assembly_center_longname%2Cplatform_type_longname%2Cintegrator_id%2CIntegrationDate&IntegrationDate%3E=2024-04-01&integrator_id!=%22argo%22&integrator_id!=%22hfradarrv%22&integrator_id!=%22instac%22&integrator_id!=%22tad%22&call_name!=%22UNKNOWN%22

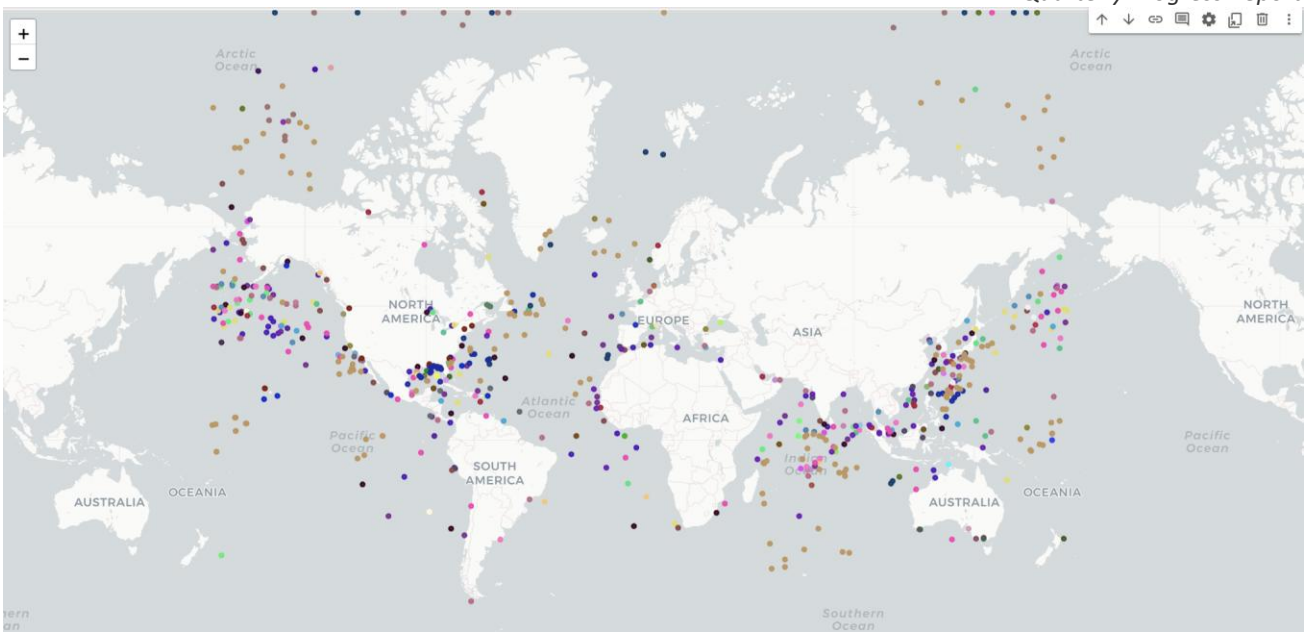

Figure 2. newly data in CMEMS INSTAC[2]

Task 4: Operate a help-service for users to provide their data in the most appropriate format

Users can either email their questions or ask for a call back. All queries are saved and tracked in the Open-source Ticket Request System (OTRS) at Ifremer. In the reporting period four questions were received and answered. As part of the migration the proprietary system at Ifremer will be replaced by using the CP help forms, which are assigned to Thematics / Ingestion by the EMODnet Secretariat. Responses will then be given through JIRA tickets.

Task 5: Allow providers of data to track the progress of their data from submission through to their storage in a repository

Data providers can follow the processing of their data submissions in the Submission Service, which is done in several steps each indicated by a status field. Data providers are contacted by assigned data centres, in case there are additional questions about the ingested data sets.

Task 6: Participate in discussions with EMODnet partners in order to improve the efficiency of the whole collection, assembly and dissemination process

All coordinators of EMODnet Thematic projects are partners in EMODnet Ingestion which guarantees a mutual tuning with EMODnet Ingestion. Moreover, EMODnet Ingestion coordinators are involved in the communication of the EMODnet Steering Committee and Technical Working Group.

Task 7: Maintain a summary record of data delivered

[2] https://ercompwebapps.emodnet-physics.eu/erddap/tabledap/EP_PLATFORMS_METADATA.csv?PLATFORMCODE%2Ccall_name%2Clatitude%2Clongitude%2CdataFeatureType%2CfirstDateObservation%2ClastDateObservation%2Cdata_owner_longname%2Cdata_owner_country_code%2Cdata_assembly_center_longname%2Cplatform_type_longname%2Cintegrator_id%2CIntegrationDate&lastDateObservation%3E=2024-04-01&IntegrationDate%3E=2024-04-01&integrator_id!=%22argo%22&integrator_id!=%22hfradarrv%22&integrator_id!=%22tad%22&call_name!=%22UNKNOWN%22&integrator_id=%22instac%22

This function is offered by the View Submissions service. Each completed submission is migrated to that service for publishing as part of a discovery and access service. Distinction is made in phase I and II which is one of the search facets. Editing activities take place aimed at replacing so-called orphan data for organizations from free text into controlled EDMO terms, orphan data for projects into controlled EDMERP terms, and orphan terms for Cruises into controlled Cruise Summary Reports (CSR) terms in order to improve the integrity and richness of the metadata.

Task 8: Engage in outreach activities towards significant holders of marine data whose data are not yet available.

A central document compiling all outreach activities and events of partners promoting the project is being continuously updated with input by all the partners. A full overview will be given in the interim and final reports. The current document is available here[3].

Early September a meeting took place at RBINS to discuss the update of EMODnet communication material. It was agreed to create a new design for a sticker to advertise the EMODnet data ingestion once it has been integrated in the Central Portal. The figure 3 shows the former sticker.

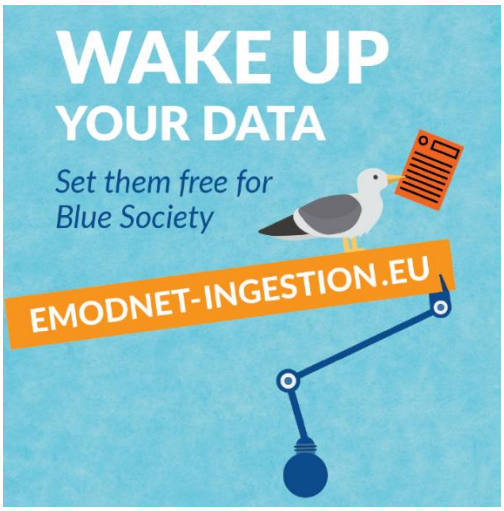

Figure 3. current sticker which needs updating

Advertisement was made for the EMODnet Marine Data for Ports, Marinas and Boating Online Workshop that is planned for 22nd of October. The advert was posted on RBINS networks, and the EMODnet secretariate was contacted to offer support for dissemination of the event.

In communication with the EMODnet secretariate it was arranged that EMODnet and EMODnet Ingestion were presented on 25th September 2024 at an event of the Sustainable Blue Economy Partnership (SBEP)[4] in Madrid – Spain. EMODnet Ingestion member IEO-CSIC, the Spanish NODC, presented the EMODnet EC data service as the focal point for possible submission of SBEP project new data via EMODnet Data Ingestion.

[3] https://docs.google.com/document/d/1cwuwM1oZYon1vAPwqLqSguLdSkSlt4ioR9MKh7cN9Yw/edit

[4] https://bluepartnership.eu/events/embarking-together-kick-meeting-funded-projects-2023-joint-call

Task 9: Improve and document the availability of data provided for coastal and offshore licensing.

Introduction

Deltares is currently in the process of selecting pilot cases to be used in the development of a roadmap for harmonizing data management in offshore licensing procedures. In parallel, a literature review has been started to provide background information for the roadmap. Additionally, a draft interview script has been prepared for use in the pilot case interviews.

Selection of Pilot Cases

Several partner countries have been contacted in order to select relevant pilot cases. It is aimed to select a total of four pilot cases that can be used for developing the roadmap for the harmonization for data management within offshore licensing procedure. It is foreseen that two cases are selected for offshore aquaculture and two cases for offshore renewable energy. An update on the final selection of cases will be part of the next quarterly report, as in line with the work plan.

First Outcomes of Literature Study

In the previous EMODnet Ingestion phase III, an inventory of current license data practices has been made. At the time of inventory, in the offshore renewable energy sector, it was found that the development and implementation of licensing procedures vary among countries. Out of the 27 countries analysed, 17 had established licensing procedures, while 10 did not. Regarding offshore aquaculture, practices also differ among countries. Among the 27 countries analysed, 23 had established procedures for offshore aquaculture licensing, while 4 did not. Licensing procedures vary among countries, which is due to different national regulations.

In several countries, different governmental bodies are needed for licensing procedures, for both offshore renewables as well as offshore aquaculture. A difference can also be found in data sharing, some countries provide licenses where sharing data via public access is included, while other countries also provide licenses where commercial confidentiality is included. This means not all data is shared.

Data collection needs during licensing and planning phase are similar between countries. Four phases were identified during EMODnet Ingestion Phase III:

- Site selection
- Licensing and Planning
- Construction and Installation
- Operation and Maintenance

Many countries indicate that they require environmental impact assessment data for the licensing and planning phase of offshore aquaculture. Some countries indicate to also use environmental impact assessment data for the offshore renewables. From the study, it is unclear if data is shared to the governmental bodies in standardised data formats. After the license is granted, many countries require monitoring in operational and maintenance phase. It is also unclear whether the monitoring data is delivered in a standardised data format.

The first steps to a roadmap for harmonisation of data management for offshore licensing procedures would include:

- Understand better differences between countries in public access licenses and commercial confidentiality licenses and data sharing procedures within these licenses.
- Understand in which phase of licensing procedures the need for better data harmonisation is the most urgent
- Understand what data standard data is commonly used to share data with governmental bodies nationally and internationally, and if this differs in different phases of the licensing procedure

Draft Script for Pilot Case Interviews

At least one interview is planned for each pilot case, with the possibility of conducting an additional interview if deemed relevant to the project's objectives. The interviews may be conducted individually or structured as

a group interview. Interviews are expected to last between 60 to 90 minutes. The main goal of the interview is to gain a deeper understanding of the current data management procedures in the country, as well as to identify the needs and challenges related to harmonization. The interview topics and questions will be tailored to focus on either offshore renewable energy or offshore aquaculture, depending on the case. The draft script is as follows:

Introduction

- Introduction of the interviewer
- Purpose of the interview
- Background of the project

Overview of Data Management Practices

1. Current Data Management Process
    a. Can you describe the data management procedures currently in place for offshore licensing in your country?
    b. What types of data are typically collected? (e.g. environmental, geophysical, or maritime traffic data)
2. Data Collection Standards
    a. Are there specific standards or protocols that provide guidance for your data collection process?
    b. Does your organization use specific tools or platforms for collecting and managing offshore licensing data?
    c. Are the selected data formats/standards national or international?
3. Data Sharing
    a. Is offshore licensing data shared with stakeholders or accessible to them?

Needs and Opportunities for Harmonization

1. Benefits of Harmonization
    a. What are the potential benefits of harmonizing data management procedures across different regions and stakeholders?
    b. Which opportunities for improvement for harmonizing data management procedures do you see?
2. Areas for Improvement
    a. Which aspects of data management do you believe need the most improvement in order to achieve harmonization?
    b. What do you think are the main obstacles for achieving harmonized data management across offshore licensing process?
3. Needed Support
    a. What kind of support, resources, or guidance would be most helpful for you in further harmonizing data management procedures?
    a. Are you familiar with any specific tools, technologies, or standards that could help harmonize data management?
        a. Note: this might also be an opportunity for informing the interviewee on the work of EMODnet

Conclusion

- Concluding remarks
- Next steps

Task 10: Service continuity during operation and for transition

Coordination of the consortium is undertaken by MARIS and HCMR to ensure the continuity of the EMODnet Ingestion portal and its array of services. The Q9 report has been prepared and submitted which was accepted by the EU.

EMODnet Data Ingestion joined three EMODnet Vision 2035 Drafting Group (VDG) on-line meetings for contributing and sharing perspectives to shape the future evolution of EMODnet (8 July 2024; 5 September 2024; 30 September 2024) and one EMODnet Governance meeting (30 August 2024).

During this reporting period, also two meetings took place with the EMODnet Central Portal (CP) technical team (19 July 2024 and 6 September 2024) to overview and discuss the further progress of the migration of the Submission service to the CP.

.

Status of the Milestones and Deliverables listed in the workplan					
Milestone/Deliverable in numerical order	WP	Date due	Status (To do/ Delivered/ Delayed)	Date delivered	If Delayed: reason for delay and expected delivery date
D0.1: Quarterly concise progress reports	0.1	M4, M7, M10, M13, M16, M19, M22, M25	D0.1-a,b Delivered	M4, M7	
D0.2: Interim report	0.1	M12	To do		
D0.3: Final report	0.1	M24	To do		
D0.4: Transition and hand over protocol	0.1	M24	To do		
D0.5i: Consortium Agreement and subcontracts	0.1	M1	Subcontracts done		
D0.6i: Short minutes - action lists of internal coordination meetings	0.1	Regularly			
D1.1: Migrated Ingestion site maintained at Central Portal	1.1	M6 – M24	Underway		
D1.2: Users migrated from Marine-ID to ECAS	1.2	M6	Underway		
D1.3: Migrated Data Submission Service operational	1.3	M6 – M24	Underway		
D1.4: Data tracking service operational	1.3	M1 – M24	Delivered	Operational since M0	

D1.5: Migrated View Submissions service operational	1.4	M1 – M24	Delivered	Operational since M0	
D2.1: Pathways operational	2.1	M1 – M24	Delivered	Operational since M0	
D2.2: Many submissions processed and published at phase 1 and phase 2	2.1	M12, M24	Underway		
D2.3: Help service operational	2.2	M1 – M24	Delivered	Operational since M0	
D3.1: Updated documentation, standards, tools, and procedure for NRT and RT data published	3.1	M12, M24	Will be integrated in Interim Report		
D3.2: Connections with new NRT and RT monitoring stations operational	3.1	M12, M24	Underway		
D3.3: Upgraded Viewing service for NRT and RT stations operational	3.2	M6 – M24	Done	M4	
D4.1: Inventory of data providers and use cases engaged in EMODnet Ingestion	4.1	M12, M24	Will be integrated in Interim Report		
D4.2: Updated promotion material	4.3	M12, M24	Will be integrated in Interim Report		
D4.3: Pilot cases selected for interviews about licensing data	4.2	M9	Underway		
D4.4: Interviews done with selected pilot case stakeholders	4.2	M16	To do		
D4.5: Pilot case Workshop hosted	4.2	M22	To do		
D4.6: Final Webinar hosted	4.2	M23	To do		

D4.7: Roadmap for harmonisation of Data Management within Offshore Licensing Procedures	4.2	M24	To do		

2. Identified issues: status and actions taken

*[Provide an **overview of issues** identified by CINEA/ DG MARE/ Secretariat (Table A) in the past quarter - new as well as pending ones, the status of those issues, and actions taken to address them and/or roadmap with remaining actions planned to resolve the issues. In Table B, provide information about any issues and challenges identified by yourself.]*

A.				
EM-768: An epic to collect together JIRA for the DIP centralisation	Pending	Folder to follow all DIP migration activities	End 2024	
EM-305/322 Content Inventory Data Ingestion	Pending	CP team has made a narrative compilation and composed a first Ingestion narrative at CP test site for check and completion by DIP; Ingestion is underway with a revised structure and drafting the narrative	End Nov 2024	

B. 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**
EM-976 Not all details of Ingestion records are displayed in the HTML pages of the Central Products Catalogue	Pending	CP Team to check and update the presentation pages of the Central Products Catalogue	End 2024; will be solved by CP as part of deploying and customising new version of GeoNetWork	
EM-893 Setup EMODnet styled page for the submission service	Pending	CP Team to create an example page with the EMODnet/Europa styling header/footer/page styling	15 Jan 2024	

B. 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**
		so EMODnet Ingestion can test the php script and CSS for the submission form.		
EM-912 New Ingestion layer – validated platforms	Resolved	New map layers for Physics as part of Ingestion	Mid July 2024	
EM-911 New Ingestion layer – as-is platforms	Resolved	New map layers for Physics as part of Ingestion	Mid July 2024	

3. Communication assets

[In Table A, list peer reviewed publications directly (co-)authored by consortium and project partners in the reporting period. In Table B, list all non-peer reviewed publications (co-)authored. In all cases, indicate the type of publication, provide the full reference incl. title, volume and issue etc., and whether the publication is open or closed access.]

A. (Co-)Authored peer-reviewed publications in the quarter					
Date of publication	**Type of publication**	**Full reference**	**ISBN**	**DOI**	**Is it open access? Yes/No**

B. Other/non-peer reviewed types of publications (co-)authored in the quarter					
Date of publication	**Type of publication**	**Full reference**	**ISBN**	**DOI**	**Is it open access? Yes/No**

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

4. Monitoring indicators

[Refer to the standardised monitoring tool, i.e. Europa Analytics, to complete the indicators excel template, and provide a short explanation in the table below on the numbers and trends for each indicator when possible/applicable. Indicate clearly if monitoring was carried out using tools other than Europa Analytics.]

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	Submission Viewing service	The total number of new phase 1 + phase 2 submissions in the current quarter is 34 and of this 17 were elaborated to phase 2. The overall number of published submissions went from 1549 to 1583.
B) Usage of data in this quarter	Cloud storage of Submission Viewing service	The total number of download transactions is with 300 much smaller than in the previous quarter which was very high with 1.000 downloads.
3. Internal and external organisations supplying/approached to supply data and data products within this quarter	Submission Viewing service	34 new data submissions were received and published from 12 organisations, mostly academic and research.
9) Visibility & analytics for web pages	Grafana	The visits to the Homepage, Submission service and Viewing service are slightly decreasing over the quarter. The homepage is most visited.
10) Visibility & analytics for web sections	Grafana	The Viewing Service which publishes the completed submissions generates most traffic, and is stable compared to the last quarter. Overall, there is no high traffic on the site, but is also not to be expected considering the function of EMODnet Ingestion in the EMODnet framework.
11) Average visit duration for web pages	Grafana	The average visit duration for the Homepage is about 35 seconds.

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