EMODnet Ingestion and safe-keeping of marine data

3rd Quarterly Report

Reporting Period: 1/12/2016 – 31/03/2017

Date: 16/04/2017

Contents

1. Highlights in this reporting period

- The Ingestion portal including several services has been officially launched, 7th February 2017, at www.emodnet-ingestion.eu

2. Work package progress

WP0 – Project Management:

All bilateral subcontracts have been established between MARIS and each of the subcontractors. The coordinator has undertaken preparations for the 2nd Plenary Group meeting with all consortium members which will take place from 10 to 12 April 2017. This meeting will be held to monitor the progress of the project since its start 19 May 2016, to discuss the coming activities for reaching out to potential data submitters, and to provide input for preparing the 1st annual progress report to the EU. In the quarterly reporting period the coordinator has participated in the following meetings:

- Progress meeting with EASME, DG-MARE and EMODnet Secretariat, 8th December 2016, Brussels – Belgium
- EMODnet Steering Committee meeting, 15th-16th February 2017, Brussels – Belgium

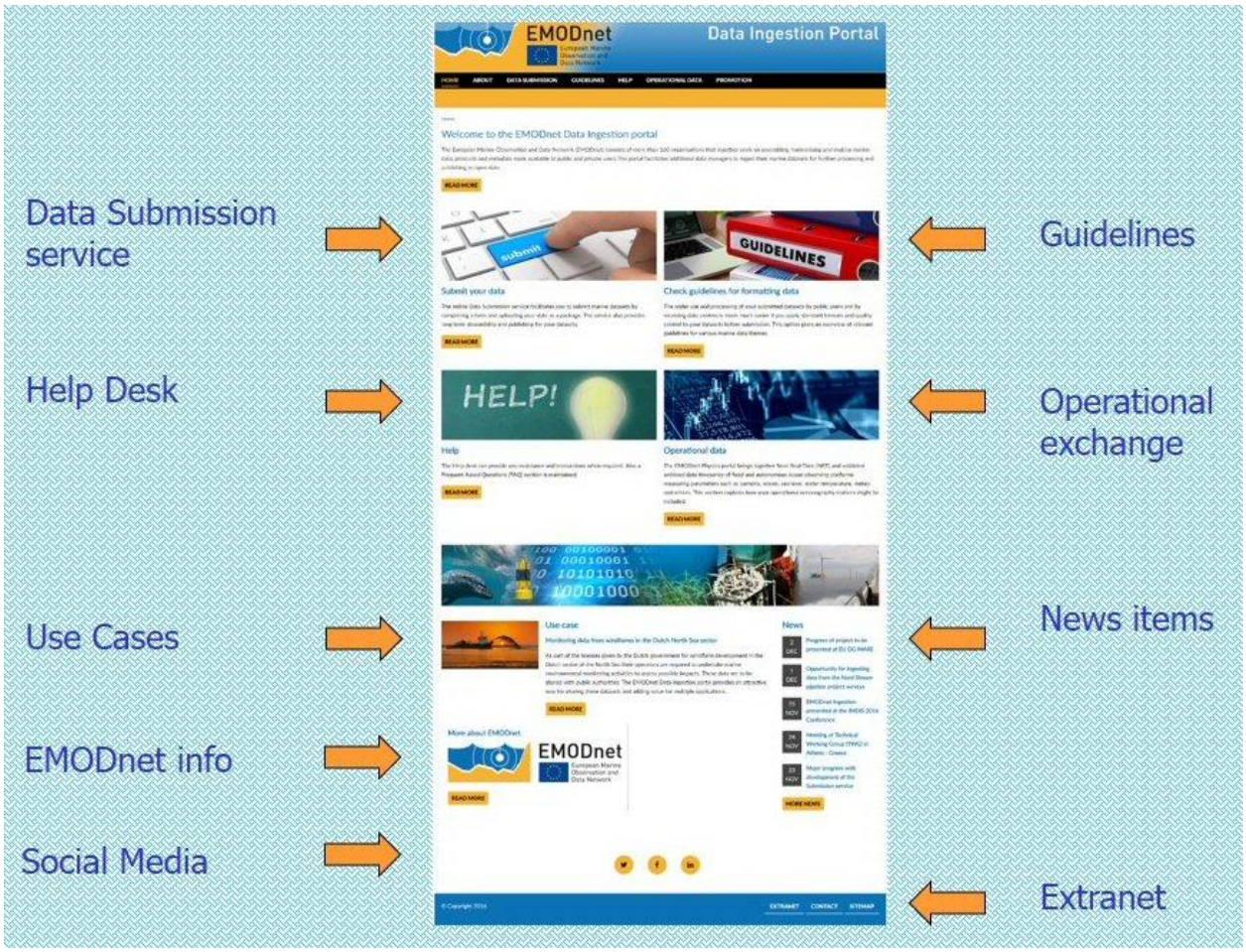
The coordinator has also prepared the 2nd quarterly progress report which was accepted by the EU (EASME and DG MARE).

WP1 – Construct and operate central Data Ingestion portal with services:

The portal has been launched 7th February 2017 at https://www.emodnet-ingestion.eu. The website has been implemented with responsive design so that it can be used at multiple platforms (PC, tabs and smart phones). The website already includes several services:

- **Data Submission service** – for Phase I – includes **User Management service** in connection with **MarineID service**
- **Guidance documents –** collection discussed at TWG - included in portal
- **Data Tracking service** for submitters, data centres and master – key fields are maintained in Data Submission service – operational including **key indicators**
- **Help desk service –** operated by IFREMER
- **Operational exchange –** invitation and instructions for opertors of operational oceanography stations to connect to the NRT data exchange of EMODnet Physics
- **Use cases –** explanations of the ongoing use cases for ingesting data from renewable energy developments in the waters around the UK and the Netherlands

The following image gives an illustration of the website and overview of the included services.



For promotion and cooperation it has been agreed with the EMODnet Secretariat and the coordinators of the EMODnet Thematic portals that each will include a promotion text and a link to the ingestion portal in their portals. This is already making good progress. MARIS maintains the website and aims to include regularly news items and project progress such as about the use cases.

The workflow for submission is divided over 2 phases:

- Phase I: from data submission to publishing of the submitted datasets package '*as is*'
- Phase II: further elaboration of the datasets package and integration (of subsets) in national, European and EMODnet thematic portals.

The Data Submission service is a core component and the ISO 19115-2 standard has been adopted for the submission form. It is supported by SeaDataNet vocabularies and directories, while also free text are allowed when required terms can not be found in the existing vocabularies and directories. The form is divided in part 1, to be completed by the data submitter, and part 2, to be completed by the assigned Data Centre. The service works with roles for Data Submitter (DS), Data Centre (DC) and Master (M).

There is a Dashboard where the functions for each of the roles are displayed. All users have to register in Marine-ID (the User Management service) and this way can be tracked. The Master can register specific data centres in the internal Data Centres Matrix from which Data Centres can be selected and assigned by the Master for handling specific data submissions. This way Data Centres are coupled to specific Data Submitters, depending on country and data theme. The Data Centres can oversee in their Dashboard all data submissions in their account and undertake action for reviewing part 1 of submissions and complete part 2. Most likely this will be supported by direct communication between the Data Centre and the Data Submitter in order to get a better understanding of the submission and to allow for completing the submission form part 2. The Data Centre also will keep a log of major actions which will be available as Process Info for both the Data Submitter and the Data Centre as well as the Master. This will be administered as part of the Tracking Service. Moreover the Dashboard of DC and M contains a Key Indicators function as an additional way to give overviews of systems performance. This allows to set filters on the received collection of submission forms and to extract and oversee in graphics and tables the range of the received submission forms.

Next to the operational submission service which is included in the Ingestion Portal recently also a TEST Submission service has been developed as a complete copy in order to facilitate training by the Data Centres in the operation of the system and to get used to the roles of Data Submitter and Data Centre. The test system is available at: **http://submission-test.emodnet-ingestion.eu**

The following developments are underway:

- **Submission Summary Records** service = **Public Discovery and Access service** – specified and implementation planned by end May 2017
- **Data Wanted service** –specifications ready – implementation planned by end June 2017
- **Data Submission service** – **extension for Phase II** – specified and implementation planned by end May 2017

WP2 – Implement pathways to forward submitted data to the appropriate repository:

The pathways concern the step going from the ingestion towards the Data Centres for further elaboration of ingested data sets. The reasoning is that ingested data should be reviewed and, where feasible, worked up by those Data Centres and to be included in their repositories for long term safe-keeping and for sharing thereafter through the European infrastructures and EMODnet thematic portals. The routing of data to repositories will depend on types of data and nationality of the data submitter. The existing pathways as in place for EMODnet thematic portals will provide a basis, while improvements might be considered for streamlining.

A further analysis on the organisation of the pathways and this has resulted in the idea to engage not only the Data Centres as present in the EMODnet consortium, but also the group of Data Centres which are involved in each of the EMODnet Thematic portals. Each of the thematic groups has recently signed their new contracts with the EU and are now gearing for their kick-off meetings. MARIS is asking all thematic coordinators to engage their networks in the pathways for EMODnet Ingestion and to bring this forward to their consortium members at their kick-off meetings. So far MARIS has done this for EMODnet HRSM and the Hydrographic Service organisations have agreed to perform this role for bathymetry data sets submitted for their countries. The fact that all thematic coordinators are members of the Ingestion consortium is very useful for this purpose. Only Seabed Habitats is not a member of the Ingestion consortium but they have indicated to be interested in collaborating too. This way the number of Data Centres involved in the processing and handling of submitted data sets will increase considerably, giving a better coverage of countries and data themes. All involved Data Centres are included with their responsibilities (themes and countries) and their contacts (via Marine-ID user names) in the Data Centre Matrix that MARIS is maintaining and which is included in the Submission Service. Submitted data packages are assigned by HCMR and MARIS to capable Data Centres from this DC Matrix. The present DC Matrix contains all Data Centres from the EMODnet Ingestion consortium and has been compiled by MARIS by asking all Consortium members for their contacts (with Marine-IDs) and preferred responsibilities.

In addition it is considered to analyse and compile a number of conversion tools that can be made available as online services, **first to Data Centres** and later possibly **also to Data Submitters**. The online services might be provided as cloud services. This approach is in line and can benefit from ongoing activities in related projects such as SeaDataCloud for developing a Virtual Research Environment (VRE) with workflows and dedicated services for processing different data types. This will be further worked out.

WP3 – Facilitate machine-to-machine transfers:

It focuses on operational oceanography and concerns 2 separate tasks:

- Establishing new connections to the existing EuroGOOS – Copernicus INSTAC Near Real Time (NRT) exchange progress
- Pilot for Real Time (RT) exchange using Sensor Web Enablement (SWE).

The NRT machine-to-machine ingestion focuses on additonal NRT data for EMODnet Physics. These will originate from additional operators that are willing to get connected and share their NRT data freely in EMODnet Physics. At the launch of the Ingestion portal already standards and procedure for NRT data and a list of receiving repositories for NRT data have been published at the portal. The portal explains how the NRT exchange is organised between operators of observing stations and EuroGOOS – Copernicus and it gives a link to the EMODnet Physics portal to show how NRT stations can be displayed

and used. Moreover a set of stepwise instructions is included in the Ingestion portal explaining how to connect in practice and also a list of EuroGOOS – Copernicus contact persons who to contact depending on region. This implementation covers the activities WP3.11 and WP3.12. Open is now WP3.13: Connecting new monitoring stations. This concerns putting into practice connecting new monitoring stations.

The Sensor Web Enablement (SWE) pilot concerns real time monitoring systems, allowing direct standardised access to selected data types from selected monitoring instruments. There are several subtasks. Good progress has been made with formulating SWE profiles for platforms and sensors which are expected end June 2017. Setting up a pilot with a few partners has been completed and in practice concerns BODC, IFREMER, ETT and OGS which already had SOS implementations in place. Developments are ongoing now for including a pilot SOS service with a number of stations and instruments as demonstrator in the Data Ingestion portal and EMODnet Physics portal. Moreover detailed instructions and tools should be made available for adoption by other operators of comparable monitoring stations who wish to join the pilot service. It is expected to be ready by end June 2017.

WP4 – Marketing and outreach activities:

So far most action has focused on the technical development of the portal and its ingestion services and involving only a few core partners. However the marketing and outreach activities as planned in WP4 are equally important and require involvement of all consortium members in the coming 2 years of the project. An important aim of WP4 is to identify, approach and convince potential data submitters to make use of the EMODnet Ingestion service for ingesting their data sets. In addition WP4 aims to undertake promotion.

To prepare for the coming 2nd Plenary Group meeting all consortium members have been requested to prepare a national inventory of potential data providers and data sources. For that purpose MARIS has prepared and distributed to all a guidance document. This guideline gives an overview of the data types that have priority for the EMODnet thematic portals and general guidance and hints how the EMODnet Ingestion members might compile their national overviews. Finally it gives a template for reporting identified potential data providers and possibe sources. Moreover an example of a national inventory was compiled and circulated by HCMR.

The plan is to go through the total inventory at the 2nd Plenary Group meeting by EMODnet theme to analyse the results in order to identify the best candidates for which follow-up activities will be undertaken by the national members. The thematic sessions will be coordinated by each relevant thematic coordinator who can ask for further details and consider whether provided data options are priority or not. This way a number of thematic sessions took place to filter and mark-up the total inventory. The refining action will result in the ‘shopping list’ that will be given back to the consortium members for giving a follow-up. This includes undertaking action towards the identified data providers.

The ultimate goal is to achieve willingness and cooperation with the identified potential data providers for submitting selected data sets to the EMODnet Ingestion portal. At the coming Plenry Group meeting brainstorming will take place how this follow-up can be implemented in an effective way.

RBINS as WP4 leader is developing a set of promotion material which can be used for central promotion activities and also by each of the consortium members as part of the WP4.2 activities. RBINS is developing the following promotional items:

- Bookmarks (ready for distribution)
- Poster (ready)
- Generic presentation (short and long version) (ready)
- Leaflet (under development)
- Movie (graphical – under development)
- Possible tutorial (youtube) how the submission service works (to be developed)
- EMODnet Data Ingestion USB-card with presentations (under development)
- EMODnet Notebooks (personalised with 'Data Ingestion') (under development)
- EMODnet gadgets…. (under development)

3. Unexpected difficulties encountered

Nothing to report.