



# EMODnet Ingestion and safe-keeping of marine data

2nd Quarterly Report

Reporting Period: 1/09/2016 – 30/11/2016

Date: 06/12/2016

## Contents

1. Highlights in this reporting period.....	3
2. Work package progress .....	3
WP0 – Project Management: .....	3
WP1 – Construct and operate central Data Ingestion portal with services:.....	3
WP2 – Implement pathways to forward submitted data to the appropriate repository:.....	6
WP3 – Facilitate machine-to-machine transfers:.....	6
WP4 – Marketing and outreach activities:.....	6
3. Unexpected difficulties encountered .....	6

## 1. Highlights in this reporting period

---

- The 1<sup>st</sup> meeting of the Technical Working Group (TWG) took place on 7<sup>th</sup> – 8<sup>th</sup> November 2016 in Athens – Greece with participation of the TWG partners to discuss progress and further planning of the portal, associated services, pathways and promotion.
- Agreement in the consortium on the submission form and the overall procedure for completing the form and related pathways.

## 2. Work package progress

---

### ***WP0 – Project Management:***

The Coordinator has undertaken further activities for the Consortium Agreement (CA) and the Subcontracts. This has resulted in the status that all partners have signed the CA and have received a digital copy of the CA, which is also added to the extranet. Furthermore, bilateral subcontracts have been established between MARIS and each of the subcontractors; with exception of the subcontract for the Turkish subcontractor. However also this last subcontract will be signed soon.

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

Further developments have taken place for the central portal. Its domain has been registered as <http://www.emodnet-ingestion.eu>. A sitemap has been prepared by the Coordinator that has been agreed with the consortium. Actions have been formulated for several partners to contribute texts for the portal. This is now well underway. MARIS has drafted a design layout, which has been agreed with the consortium. MARIS has recently concluded the configuration of the Content Management System and a test version of the public portal. Work is ongoing for editing and entering texts and images to complete the public portal for publication.

A meeting of the Technical Working Group (TWG) has taken place on 7<sup>th</sup> – 8<sup>th</sup> November 2016 in Athens – Greece to discuss and take decisions on the specifications and further developments of the planned services for:

- Data Submission
- User Management
- Data Tracking
- Summary Records

- Help Desk
- Data Wanted
- Guidelines

Key topics for the Data Submission service are the submission form as well as the workflow how the submission form will be completed and followed up. HCMR together with MARIS had finalised the draft for the submission form based upon the ISO19115-2 standard and also considering compliance with INSPIRE and practices at other ingestion systems, such as such as NOAAs Send2NCEI service. As much as possible use is made of SeaDataNet vocabularies, SeaDataNet Directories (EDMO, EDMERP, CSR) and other standard lists (e.g. from INSPIRE) to structure and standardise the completing of the metadata fields of the form. This form was presented at the TWG meeting and received acceptance as a good basis. Furthermore the TWG agreed with a proposal from MARIS for the workflow how to go from an initiation of a data submission by a data submitter up till publishing of the submitted datasets and for the associated services. In practice, it is planned to distinct between 2 phases in the life cycle of a data submission:

- **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.

For the implementation of **Phase I** it is very relevant that potential data providers are organisations that are not yet routinely submitting data sets to national data repositories. They will need encouragement and support for the actual submission, and they are not (yet) used to practices and standards as used by the international marine data management community. For that purpose it has been decided to make the treshold for submission relatively low which is translated into a procedure that the submission form (ISO19115 – INSPIRE compliant model) is split in 2 parts:

- **Part 1 submission form:** a number of key fields to be completed by the Data Submitter, including uploading of a zip file with the datasets and related documentation;
- **Part 2 submission form:** review of the received datasets package and part I metadata, and consecutive completion of the additional metadata fields of the submission form by the appointed Data Centre.

Initiating and completing a data submission form will be implemented as part of the **Data Submission service**. As part of the workflow part 1 data submissions will be initially received by the master of the Data Submission service (in casu MARIS and HCMR) and assigned by the master to the most appropriate Data Centre. This will be based upon data discipline, country and agreements to be made with the networks of Data Centres, engaged in each of the EMODnet Thematic portals. The appointed Data Centre then has the responsibility to finetune and complete the submission form, possibly in contact with the Data Submitter to the level that it can be published. The publishing will be implemented by the **Summary Records service** which will be a public Discovery and Access service, based upon the completed data submission metadata, and giving open access to the received Datasets

Package *as is*. This approach guarantees that the time from submission to publishing can be relatively short which is essential for creating momentum and satisfaction at the Data Submitter. All users of the Data Submission service, in casu Data Submitters, Data Centres and Master, will be facilitated to oversee the progress of the workflow per data submission and associated processing log information. Data Submitters can only oversee their own submissions, and Data Centres can oversee the submissions in their account, while the Master can oversee all data submissions. This functionality will be implemented by the **Data Tracking service** which will be integrated in the Data Submission service. In practice, a **dashboard** is included in the Data Submission service to facilitate access to various services, depending on the role of the user (Data Submitter / Data Centre / Master). The Data Submitters, Data Centres and Master are maintained with the **User Management service** which is another service in the dashboard. The User Management service works together with the SeaDataNet Marine-ID Authentication and Authorisation service to give each user a uniq logon. Note: the **Summary Records service** is not integrated in the dashboard of the Data Submission service, but is another main service which can be accessed from the main EMODnet Data Ingestion portal.

For the implementation of **Phase II** it is planned that the appointed Data Centres undertake extra efforts to work up (subsets of) the original datasets package up to the level that the datasets and associated metadata will be integrated in the data management infrastructures of the Data Centres themselves for discovery and access by their national portals. Furthermore the worked up datasets will be populated into the relevant European infrastructures such as SeaDataNet, EurOBIS and others, and from there included in the EMODnet portals. This Phase II is to be developed and implemented as part of WP2 (see below). An important feature will be that information about all further processing is logged in the **Data Tracking service** and this way available to its stakeholders. And that exchange links will be inserted between the national, European and EMODnet Thematic portals and the Summary Records service to serve users.

HCMR had already started with the development of the core Data Submission service well before the TWG and in communication with MARIS. Following the TWG developments have been increased and expanded also to include related services (User Management service, Data Tracking service, Part 1 – part 2 completion) and the agreed workflow. The services are now almost completed and ready for beta testing. The Summary Records service development will start soon by MARIS and is required to be ready in 2 – 3 months, once an initial number of data submissions has been booked. The service makes full use of the data submission database and awaits full completion of that database model before undertaking follow-on activity.

The Help Desk is developed and implemented by IFREMER and during the TWG already a working service has been presented. Following the TWG further action has been dedicated to establishing an email identity, which has been set to: [support@emodnet-ingestion.eu](mailto:support@emodnet-ingestion.eu). The Help Desk will be integrated in the Central Portal as one of the main services.

During the TWG also specifications have been presented by IFREMER for the **Data Wanted service**. These will be finetuned in the coming 2 months, while the development by IFREMER will aim for delivery by Month 12 (end May 2017).

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

As indicated under WP1, the pathways will be further specified and implemented as part of **Phase II** of the workflow. Following the kick-off meeting coordinators of EMODnet Thematic lots have had the time to discuss with their networks how they should be engaged in the further elaboration and integration of (subsets of) the received datasets submissions. In principle, each EMODnet Thematic portal and associated network will get responsibility for processing submitted datasets for their discipline and existing Data Centres in these networks accept this perspective way of working. However all groups are awaiting the results of the latest EMODnet Phase 3 tender evaluations and can not make firm agreements in this stage (except for the EMODnet Bathymetry group which already received a positive result a few weeks ago). The networks of Data Centers also are given a major role in Phase 1 as described under WP1 and therefore it is hoped that the situation will become clear soon.

### ***WP3 – Facilitate machine-to-machine transfers:***

At the TWG the WP3 leader and individual taskleaders have given presentations of their plans for the coming months and all seems to be under control.

### ***WP4 – Marketing and outreach activities:***

At the TWG the WP4 leader has given a presentation of a further developed plan for promotion and marketing for the EMODnet Ingestion initiative which will be put into action on short term. An important role will be played by the partners in the EMODnet Ingestion consortium that will identify potential data providers in their countries and that will act as ambassadors to open doors and to support data providers to submit their datasets, centrally supported by promotion material, social media and portal. A number of use cases are already foreseen, such as datasets collected for wind farms in the North Sea. These use cases will serve as extra promotion at the portal for gaining momentum.

## **3. Unexpected difficulties encountered**

---

A delay in the planning, mostly because of having the holiday period (July – August) just after the start of the project and the somewhat higher complexity than earlier thought. However the developments are under control for launch of the portal and first core services by January 2017. Moreover the planned workflow with 2 phases is considered as an excellent way to please data submitters and users.