

EMODnet Data Ingestion: ‘Wake up your data’

Dick Schaap, MARIS (Netherlands), dick@maris.nl
Sissy Iona, HCMR (Greece), sissy@hnodc.hcmr.gr
Francis Strobbe, RBINS (Belgium), fstrobbe@naturalsciences.be

The ‘**EMODnet Ingestion and safe-keeping of marine data**’ project, started mid-2016, seeks to identify and reach out to organisations from research, public, and private sectors who are holding marine datasets and who are not yet connected and contributing to the existing marine data management infrastructures which are driving EMODnet. Those potential data providers should be motivated and supported to release their datasets for safekeeping and subsequent freely distribution and publication through EMODnet. The EMODnet Data Ingestion portal facilitates submission of their sleeping marine datasets for further processing, Open Data publishing and contributing to applications for society.

The activities are undertaken by a large European network that is geographically anchored in the countries bordering all European marine basins, and covers all EMODnet data themes. The EMODnet Data Ingestion members are national and regional marine and oceanographic data repositories and data management experts. The coordinators of the EMODnet thematic portals are also part of this new initiative. Moreover the data centres work together on pan-European and international scales in organisations such as IODE, ICES, EuroGeoSurveys, EuroGOOS, and IHO, and for pan-European marine data management infrastructures such as SeaDataCloud, EurOBIS and EGD. The latter are feeding into several EMODnet thematic portals.

The emphasis of activities in the first year has been put towards developing the EMODnet Data Ingestion portal and its services for ingesting and publishing data sets, developing the pathways for processing and elaborating of data submissions, laying a basis for promotion and marketing activities, and making an initial inventory of potential data sources and their providers.

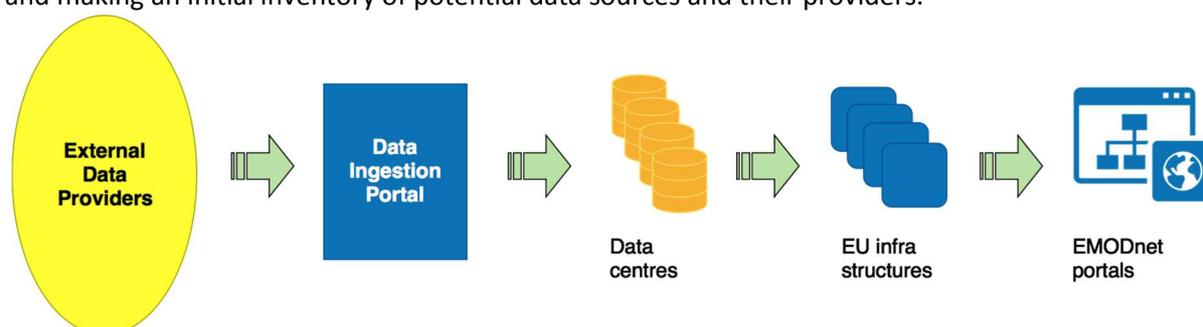


Figure 1: Flow of data sets from external data providers to EMODnet portals

The EMODnet Data Ingestion portal been launched early February 2017. It encourages data providers to share marine data, gives marine data management guidance information, and provides a range of services such as:

- submission service for easy ingestion of marine data packages
- view submissions service to oversee submitted data sets ‘as is’
- data wanted service to post requests for specific data types

Submission forms with data packages are assigned to qualified data centres depending on the country of the data provider and the type of EMODnet theme. This group includes not only the EMODnet Ingestion consortium but also the groups of data centres who are involved in each of the EMODnet Thematic portals.

A distinction is made between 2 phases in the life cycle of a data submission:

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

This split allows to publish already in an early stage the original data package with high quality metadata.

For operational oceanography a close cooperation takes place with EMODnet Physics. This aims at identifying and arranging inclusion of additional stations for Near Real Time (NRT) data exchange. The Data Ingestion portal explains how the NRT exchange is organised with EuroGOOS – Copernicus and guidance how to connect in practice. Furthermore a Sensor Web Enablement (SWE) pilot is set-up for Real Time data exchange. A client service to locate stations and to retrieve data streams in a time series viewer is hosted at the EMODnet Physics portal and ‘advertised’ at the EMODnet Data Ingestion portal.

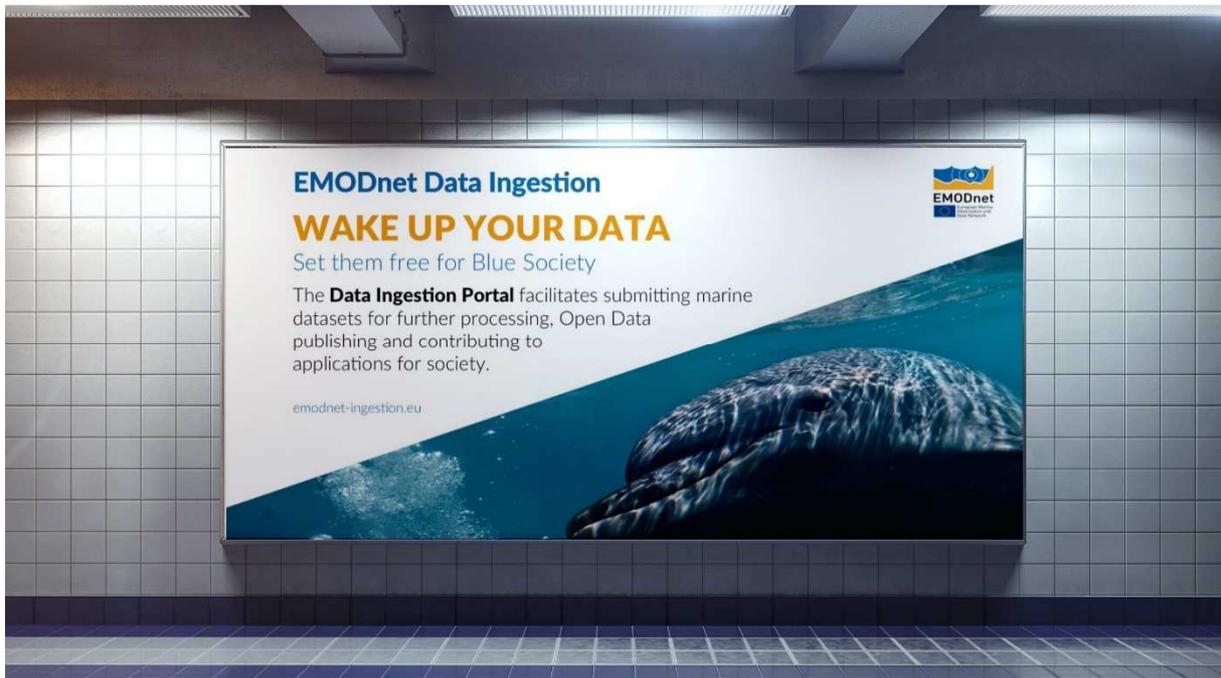


Figure 2: Poster in the Brussels metro

Promotion and outreach activities are equally important as technical developments. In the first year it has focused on establishing cooperation and synergy within the EMODnet community. A portfolio of promotional items has been developed, such as leaflets, posters, presentations, stickers, and a wonderful animation. These are part of the promotion and marketing strategy that was designed to reach out to potential data providers. In the second year this plan has been put into motion on full scale for a wider outreach and marketing to potential data providers in government, science and industry. This has so far resulted in many submissions and also in development of special use cases, such as for monitoring data from offshore renewable energy projects or minting DOIs for research data to support data citing for data submitters.