EMODNET DATA INGESTION A KEY RESOURCE FOR THE MARINE SCIENCE COMMUNITY

Leda Pecci, Paolo Diviacco, Mihai Burca

Abstract: Marine and coastal data management is crucial for scientific research, economic activities, and environmental monitoring. Marine data is collected from different sources using various tools and methodologies. There are several methods used for formatting marine data, but to put them in a standardized form takes time, effort and knowledge. To overcome the technical difficulties involved in submitting data, a group of 41 international organizations, led by HCMR (Greece) and MARIS (The Netherlands), is active in the EMODnet Data Ingestion project. Their goal is to simplify the process of data acquisition and transformation, ensuring data becomes interoperable and increasing the availability of marine data. By providing a structured platform for the submission, management, and dissemination of marine datasets, the project fosters an environment where data can be easily accessed, shared, and utilized by a wide range of stakeholders, including researchers, policymakers, and public. This paper presents EMODnet Data Ingestion services within the framework of the broader EMODnet. Examples of data sharing and important aspects of the EMODnet Data portal are delivered.

Keywords: Data submission, Marine open data, interoperability, data sharing

Leda Pecci, National Agency for New Technologies, Italy, leda.pecci@enea.it, 0009-0002-5641-3229
Paolo Diviacco, OGS, Italy, pdiviacco@ogs.it, 0000-0001-9404-9611
Mihai Burca, OGS, Italy, mburca@ogs.it, 0000-0002-2118-4887

Referee List (DOI 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Leda Pecci, Paolo Diviacco, Mihai Burca, *EMODnet Data Ingestion a key resource for the marine science community*, pp. 988-998, © 2024 Author(s), CC BY-NC-SA 4.0, DOI: 10.36253/979-12-215-0556-6.86

Introduction

A holistic approach to environmental sustainability requires addressing global challenges such as climate change and sustainable management of marine resources that rely on data analysis. In the world of marine research, the need for large amounts of reliable and accessible data is critical. In Europe, several initiatives and infrastructures have been developed to improve the collection, storage, interoperability, accessibility and usability of marine data. Even though EU funds and the substantial progress in ocean data management have made possible the implementation of different ocean data infrastructures such as SeaDataNet (https://www.seadatanet.org/) and EMODnet (https://emodnet.ec.europa.eu/en), a large quantity of valuable data remains inaccessible.

Evidence

Ocean data sharing and use require technical and cultural solutions to overcome challenges in uploading, aggregating, and navigating, ultimately to make the data available [1].

EMODnet Data Ingestion plays a crucial role in advancing the principles of open data and data reuse within the marine science domain. As an essential component of the EMODnet initiative, this platform serves as a catalyst for enhancing the accessibility, usability, and sharing of marine datasets across Europe. To understand the significance of EMODnet Data Ingestion, it is essential to be aware of what EMODnet represents.

EMODnet was funded by the European Commission, under the Integrated Maritime Policy, to address fragmented marine data collection, storage, and related issues across Europe [2].

The idea behind EMODnet emerged following the publication of the 2006 Green Paper "Towards a Future Maritime Policy for the Union: A European Vision for the Oceans and Seas" [3] that was a pivotal document aimed at initiating a comprehensive discussion on the future of maritime policy within the European Union [4].

The EMODnet pilot projects began in 2009. In 2010, the European Commission (EC) released the Marine Knowledge 2020 communication, which aimed to enhance data availability and accessibility as well as to highlight the importance of marine data across Europe. This initiative was a step forward in recognizing marine data for sustainable growth and informed decision-making in maritime sectors. According to this document, the EC proposed funding for EMODnet to support its further development [5]. Implementing the objectives of the Marine Knowledge 2020 strategy EMODnet expanded its thematic portals, improving data accessibility by standardizing data formats and integrating various national, regional, and international databases. Collaboration and synergy with EU marine research initiatives, such as SeaDataNet and the Copernicus Marine Environment Monitoring Service (CMEMS), have contributed to the enhancement of EMODnet infrastructure [6].

EMODnet allows access to high quality and harmonised marine data and data products across Europe. The initial pilot projects focused on thematic areas such as biology, bathymetry, chemistry, geology, habitats, and physics, providing a foundation for a more integrated marine data infrastructure [7].

Currently EMODnet functions as the in situ marine data service of the EC, specifically under the Directorate-General for Maritime Affairs and Fisheries (DG MARE) (https://emodnet.ec.europa.eu/en/about_emodnet).

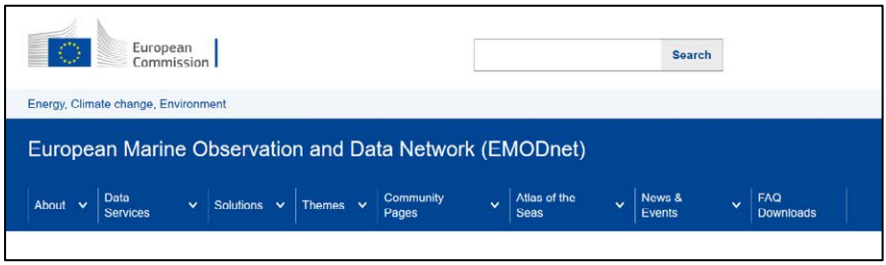

Figure 1 – The EMODnet portal.

Given that the partners involved in EMODnet are not directly funded by the European Commission to provide marine data, their collaboration with other initiatives is essential for the success of the network. EMODnet serves as a data aggregator service, relying on upstream global, pan European and coastal in situ observing networks (GOOS, EuroGOOS, European Marine Research Infrastructures), and associated data assembly and processing infrastructures such as National Oceanographic Data Centres, and SeaDataNet (https://marine.copernicus.eu/news/european-situ-marine-data-service-landscape).

EMODnet Data Ingestion fosters an environment conducive to collaboration, informed decision-making for policy makers, and sustainable management of marine resources. The project website incorporates interactive features that encourages collaboration and engagement (https://www.emodnet-ingestion.eu). The core of the portal is the user-friendly “Data Submission Service”, which simplifies the process for data producers to provide their datasets.

This service breaks down the submission into manageable steps, allowing organizations to upload their data with minimal barriers. This streamlined approach exhortes a wide range of stakeholders, including research institutions, public agencies, and private organizations, to share their marine data by implementing a simplified approach.

An online form, representing a comprehensive tool for users to enter the necessary information, facilitates the data submission service. This digital interface aids the submission process, guiding the data submitter through each step with ease by a series of structured fields, ensuring that all relevant data is accurately captured and submitted in a systematic manner. In the initial phase of the 2-step data submission process, the data contributor fills in a selection of essential fields in the

submission form and attaches a compressed archive (in zip format) that includes the datasets and accompanying documentation.

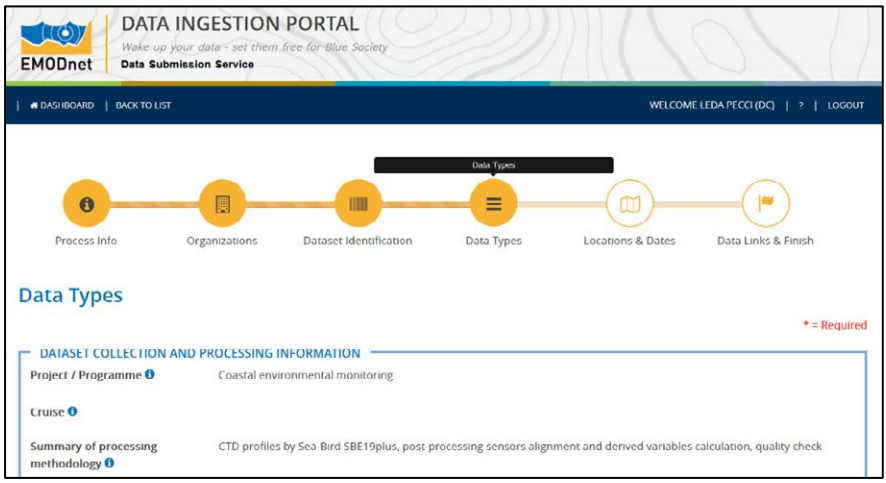

Figure 2 – The first phase of the data submission process.

A data center is chosen to support the completion of the submission process, in the first phase is in charge to try to replace free-text entries, written by data providers, with standardized terms and to make other adjustments in agreements with the data submitter. In the second phase the data center, in close collaboration with data submitter, improve and refine the datasets through a series of systematic processes aimed at improving the quality and usability of the data. Once the datasets has been inspected, their metadata have been finalized and any necessary processing has been performed, they are ready for integration into the national and European infrastructures, such as SeaDatanet, EurOBIS and ICES, which support EMODnet. This integration ensures that the datasets are accessible via the EMODnet portal, broadening their availability to a diverse spectrum of users.

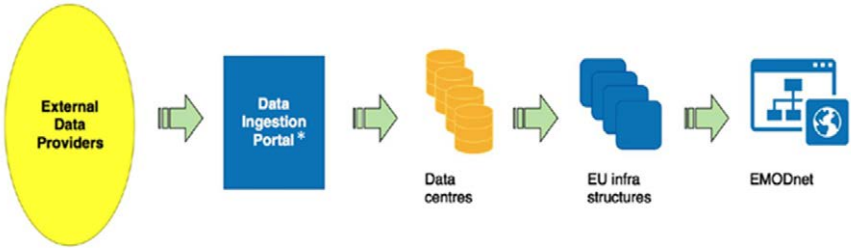

Figure 3 – Workflow from the submission to the publishing. Note: This graph was produced by Dick Schaap (MARIS).

EMODnet Data Ingestion community is creating a culture of open science that offers numerous advantages that enhance the scientific process and its societal impact. One significant benefit is the promotion of transparency and collaboration among researchers, which leads to improved trust in scientific findings. By making research data openly accessible, the academia community can replicate studies using the original data and methods, allowing them to verify the findings. Furthermore, scientists can engage with a wider audience, including policymakers, practitioners, and the public, fostering a more inclusive dialogue around scientific knowledge.

We strongly encourage the submission and use of open marine data expecting that, eventually, these practices will be widely adopted. Infrastructures like SeaDataNet utilizes SEANOE (SEA scieNtific Open data Edition) to enhance data sharing among researchers. By utilizing SEANOE's service, researchers can publish their data and obtain a DOI (Digital Object Identifier), free of charge [8]. Thanks to a strong connection between SeaDataNet and EMODnet Data Ingestion, datasets published in SEANOE are automatically replicated to the EMODnet Data Ingestion portal and published as if they were submitted directly.

One of the possible data types ingestible by Emodnet on which we here will report is anthropogenic underwater noise. This kind of pollution is gaining a lot of importance recently considering the explosion of marine traffic and port related activities.

Marine animals are highly dependent on sound to communicate with their environment. Sounds are crucial for foraging, communication, predator avoidance and general spatial orientation [9]. According to the EC Marine Strategy Framework Directive (MSFD), sounds that have negative effects are defined as noise, and indeed exposure to underwater acoustic noise affects marine life [10].

The effects of underwater noise on marine life encompass a wide range of phenomena, a subset of which are of detrimental ecological significance in terms of long-term consequences for populations. The effects of anthropogenic underwater noise include issues such as behavioral responses masking, hearing loss, physical and/or physiological effects, and even death [11] [12] [13].

To address such topics Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS, in collaboration with other institutions and companies, developed a low cost marine noise monitoring system (CORMA). The structure aims to reconstruct the acoustic pressure in a defined area through the deployment of a real time acquisition and transmission device installed on OGS buoys located within the area of the Gulf of Trieste (North Adriatic Sea). Further details on the acquisition system can be found in [14]. The system recorded two years of continuous data that are made public through the Emodnet Ingestion portal following the FAIR principles and metadata standards there designed.

Files available can be downloaded as standard .wav files that can be playback, processed and analysed using any acoustic data processing software. Data is sampled at 48 kHz 24 bit since the main objective of the project is to correlated underwater noise and marine traffic. Larger spectrum would mean that larger files should be transmitted which would create problems in energy consumption of the system, which being powered by solar panels and batteries need to be carefully tailored for the aim of the research. Data is stored in a central facility from which each recording

can be accessed and downloaded. Each file stores 10 minutes of recording and is catalogued with an easy to be read timestamp that is shown also at filename level.

Analysing the data it is possible to identify the source of the noises. Low frequency that span large time recording can be associated with large vessels, while high frequency short duration signals can be associated with leisure boats.

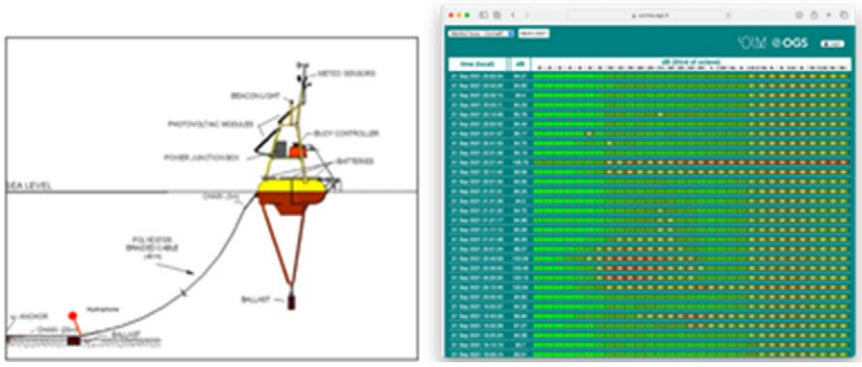

Figure 4 – Corma monitoring system (left) and data storage and access (right).

In our case the datasets are findable on web site of SeaDataNet https://www.seadatanet.org/Metadata/CDI-Common-Data-Index using "QUERY THE COMMON DATA INDEX (CDI) V5 DATA DISCOVERY AND ACCESS SERVICE". The quickest way is to use the tool of "Free search" in a both way: by instrument type (hydrophone) or knowing previously the name of the project (Corma), getting a result as:

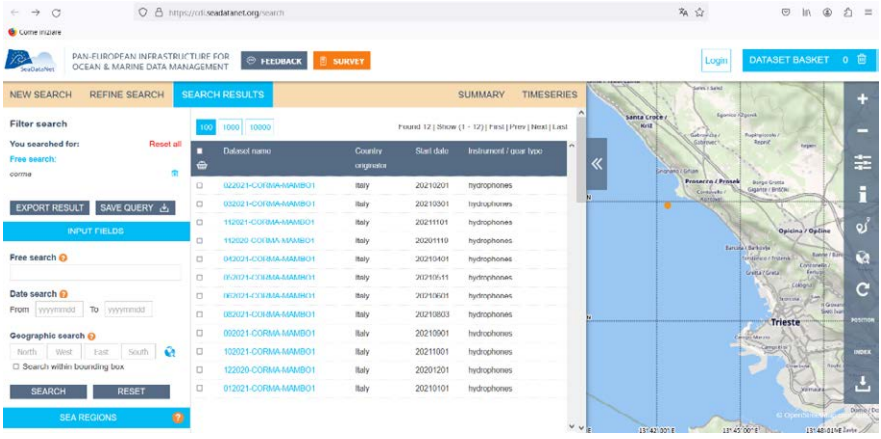

Figure 5 – Visualization of Corma data.

Once you select/click on the dataset(s) you can have access to the data adding it on the basket of SeaDataNet or more easily scrolling down on metadata till the link of the website of the data owner: https://corma.inogs.it/index.php?date=2021-06-01.

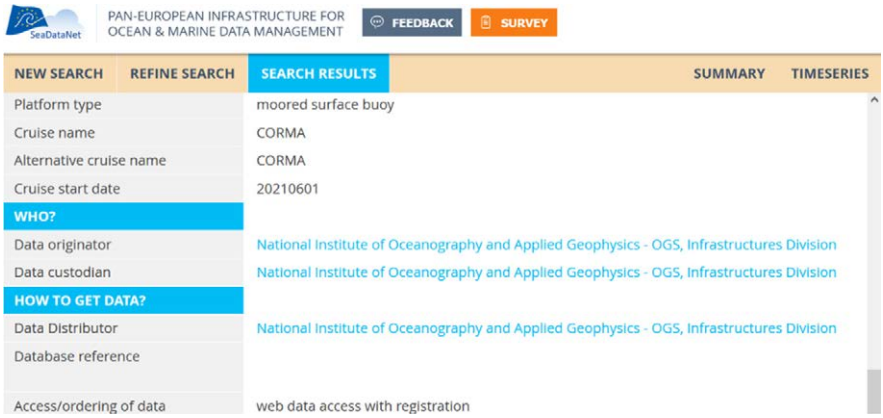

Figure 6 – Visualization of various metadata of Corma's data.

It is enough to change the daily date on the link address to view all data day by day.

Another way to retrieve data is by using the EMODnet service. EMODnet operates as a data service of the European Union, and so, its website is incorporated into the larger EC web framework.

There are different ways to access the data, the most direct is by navigating to the EMODnet portal, starting from a map-based web-page which serves as a primary resource for retrieving various kind of data, mainly organised by disciplines, at web address: https://emodnet.ec.europa.eu/geoviewer/.

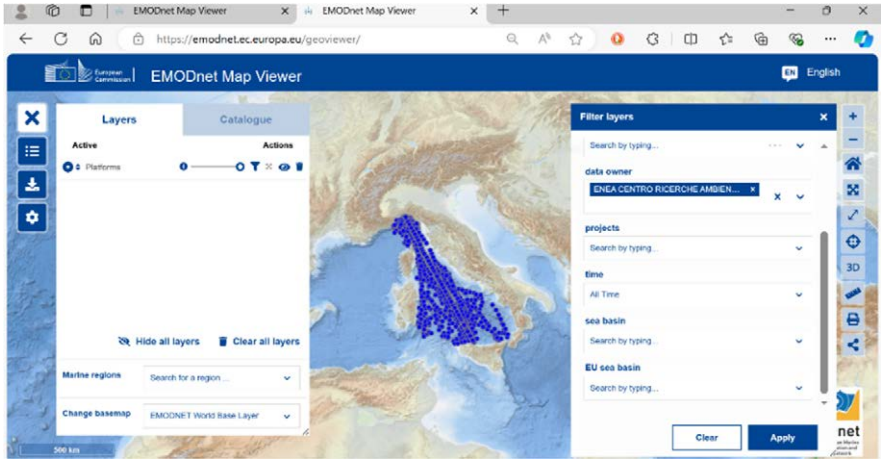

Figure 7 – Visualization of some ENEA's physics data in the Mediterranean Sea.

To activate a layer select it from the “Catalogue” tab, then go to the “Layers” tab. Clicking on the filter icon a list to select appears. It is possible to filter by name of the owner of the data or by other parameters.

Whereas, to search for data, by selecting an area on the map, it is necessary to click the "Download" button in the left toolbar and a button appears on the right side of the map. Click on it and then draw a bounding box on the map. On the right side of the menu appears the format in which you can download the data.

ENEA has shared a diverse array of data throughout the years, showcasing its commitment to collaboration. This data includes various kind of marine data belonging to different disciplines, some of which has been produced in partnership with other Italian institutions and associations.

Here, we want to present an example of citizen science data in the field of the Mediterranean studies: the MedFever project. The public–private partnership between ENEA, the University of Rome, “La Sapienza” and the MedSharks association, makes available valuable data. The MedSharks association is focused on the conservation and monitoring of shark populations in the Mediterranean Sea, but is also involved in various initiatives. This project engaged volunteers, from 12 diving centers, in collecting temperature data.

In 2021, a network of 12 small observatories, by means of marine probes, was established in the Tyrrhenian Sea, with the aim of collecting high-frequency measurements of temperature at standard depths [15]. This initiative was made possible through the collaboration of voluntary diving centers, who played an important role in the deployment and maintenance of the observational infrastructure.

Citizen science offers a unique opportunity for anyone, regardless of scientific background, to contribute to the well-being of our Seas. The collaboration of a working team of highly motivated people in the private sector, professional oceanophers and ocean data managers provides a major impact in the advancements of ocean research. This allows the sharing of a larger quantity of high quality data to a wider audience by using EU infrastructures and standards for the data and metadata.

To retrieve the data submitted by using the EMODnet Data Ingestion Service the simplest way is to go to the “Viewing Submission Service” and select one of the option available, as, for example, the button “Sea Areas” writing “Tyrrhenian Sea” and “ENEA” in the field “Free Search”.

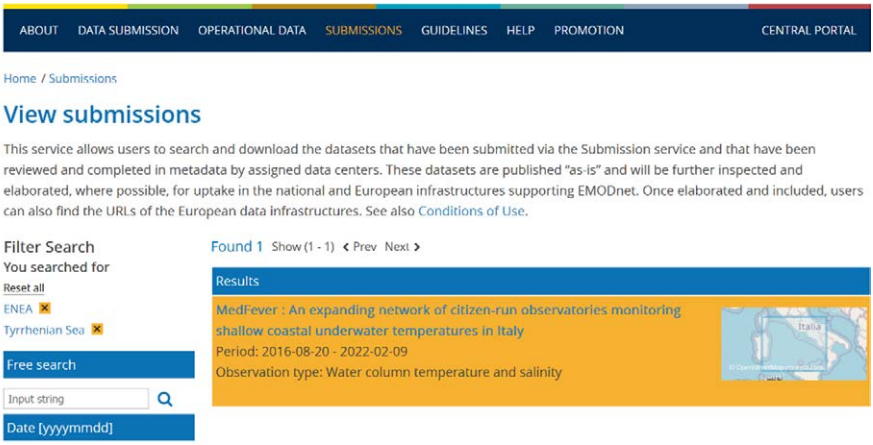

Figure 8 – Example on how to retrieve data in the "View Submissions" service.

Conclusions

The described initiative is designed to support the goals of EMODnet by promoting data reuse and adhering to open data principles, which are essential for effective marine resource management and environmental sustainability. In an era where global environmental challenges such as climate change, pollution, and biodiversity loss are increasingly prevalent, the sharing of ocean data emerges as an important means for dealing with these issues. Data reuse may have other purposes than those for which they were originally acquired. The FAIR (findable, accessible, interoperable, and reusable [16]) principles allow other stakeholders to make use of this data. The project has been playing an important role in advancing the practice of open data sharing, significantly shaping the behaviors and attitudes of researchers within the scientific community. By promoting the principles and the practices of making research data openly accessible, the project not only facilitates the dissemination of knowledge but also fosters a cultural shift among researchers regarding how data is perceived and utilized. This paper highlights the significant benefits that marine data sharing offers as well as the asistance provided by the EMODnet Data Ingestion portal. This portal offers technical support, provides resources and aids with marine data management, submission, and curation processes, and allows for high quality, well-documented data.

Acknowledgements

This paper represents the progress of a collaborative effort by a large and diverse group of marine scientists and data expert contributors. We wish to express our most sincere appreciation to the institutions and teams that supported this project as well as all the projects that have made EMODnet possible. Your

guidance and belief in our vision were fundamental for the outcomes achieved and motivating our large team to work together towards common goals.

The European Marine Observation and Data Network (EMODnet) is financed by the European Union under Regulation (EU) No 508/2014 of the European Parliament and of the Council of 15 May 2014 on the European Maritime and Fisheries Fund.

The EC, by means of the European Climate, Infrastructure and Environment Executive Agency (CINEA) funds the current EMODnet Data Ingestion project.

For what is concerning the data, we want to thank the Transpobank s.r.l. for their assistance in the experiment about underwater noise data, for the citizen science data, the MedShark association, the ENEA team, and the University of Rome "La Sapienza" for their willingness to share data. Regarding the MedFever project, we express our gratitude to the LUSH Italia for the financial support and Mireno Borghini (CNR) for the calibration facility.

References

[1] Pendleton L. H. *et al.*, (2009) - *Disrupting data sharing for a healthier ocean*, ICES J. Mar. Sci., vol. 76, issue 6, pp. 1415–1423, DOI: 10.1093/icesjms/fsz068.

[2] Calewaert J.-B., Weaver P., Gunn V., Gorringe P., Novellino A. (2016) - *The European Marine Data and Observation Network (EMODnet): Your Gateway to European Marine and Coastal Data*», in Zerr B., Jaulin L., Creuze V., Debese N., Quidu I., Clement B., Billon-Coat A. (eds), *Quantitative Monitoring of the Underwater Environment*, Ocean Engineering & Oceanography, vol. 6. , Cham: Springer International Publishing, pp. 31–46. DOI: 10.1007/978-3-319-32107-3_4.

[3] Commission Communication (2023) - *Commission staff working document: Evaluation of the European Marine Observation and Data Network (EMODnet)*, COM(2023)282. Publications Office of the European Union., 26 July 2023. [Online]: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52023SC0281

[4] European Commission (2006) - *Green Paper "Towards a future Maritime Policy for the Union: A European vision for the oceans and seas*7 June 2006. [Online: https://eur-lex.europa.eu/resource.html?uri=cellar:b2e1b06a-6ca9-4e24-ac15-60e1307f32e2.0003.03/DOC_1&format=PDF]

[5] European Commission (2010) - *Marine Knowledge 2020 - Marine Data and Observation for Smart and Sustainable Growth*. Publications Office of the European Union, 8 September 2010. [Online: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM%3A2010%3A0461%3AFIN%3AEN%3APDF]

[6] Martín Míguez B. *et al.* (2019) - *The European Marine Observation and Data Network (EMODnet): Visions and Roles of the Gateway to Marine Data in Europe*, Front. Mar. Sci., vol. 6, p. 313, DOI: 10.3389/fmars.2019.00313.

[7] Novellino A., Gorringe P., Schaap D., Pouliquen S., Rickards L., Manzella G. (2014) - *European marine observation data network — EMODnet Physics»* in *2014 IEEE/OES Baltic International Symposium (BALTIC)*, Tallinn, Estonia: IEEE, pp. 1–3, DOI: 10.1109/BALTIC.2014.6887862.

[8] Pecci L., Fichaut M., Schaap D. (2020) - *SeaDataNet, an enhanced ocean data infrastructure giving services to scientists and society*, IOP Conf. Ser. Earth Environ. Sci., vol. 509, fasc. 1, p. 012042, DOI: 10.1088/1755-1315/509/1/012042.

[9] Codarin A., Picciulin M. (2015) - *Underwater noise assessment in the Gulf of Trieste (Northern Adriatic Sea, Italy) using an MSFD approach*, Mar. Pollut. Bull., vol. 101, issue 2, pp. 694–700, DOI: 10.1016/j.marpolbul.2015.10.028.

[10] Van der Graaf A.J., Ainslie M.A., André M., Brensing K., Dalen J., Dekeling R.P.A., Robinson S., Tasker M.L., Thomsen F., Werner S. (2012) - *European Marine Strategy Framework Directive Good Environmental Status (MSFD-GES) Report of the Technical Subgroup on Underwater Noise and other forms of energy*, EC.

[11] Thomsen F., Campbell J., Fredheim B., Unger S., Ashe S., Middleton B. (2009) - *Overview of the Impacts of Anthropogenic Underwater Sound in the Marine Environment*, OSPAR.

[12] Vasconcelos R. O., Amorim M. C. P., Ladich F. (2007) - *Effects of ship noise on the detectability of communication signals in the Lusitanian toadfish*, J. Exp. Biol., vol. 210, issue 12, pp. 2104–2112, DOI: 10.1242/jeb.004317.

[13] Robinson S. P., Wang L., Cheong S.-H., Lepper P. A., Marubini F., Hartley J. P. (2020) - *Underwater acoustic characterisation of unexploded ordnance disposal using deflagration*, Mar. Pollut. Bull., vol. 160, p. 111646, DOI: 10.1016/j.marpolbul.2020.111646.

[14] Diviacco P. *et al.* (2021) - *Underwater Noise Monitoring with Real-Time and Low-Cost Systems, (The CORMA Experience)*, J. Mar. Sci. Eng., vol. 9, issue 4, issue 4, DOI: 10.3390/jmse9040390.

[15] de Sabata E. *et al.* (2022) - *MedFever: An expanding network of citizen-run observatories monitoring shallow coastal underwater temperatures in Italy*, SEANOE. DOI: 10.17882/86403.

[16] Wilkinson M. D. *et al.* (2016) - *The FAIR Guiding Principles for scientific data management and stewardship*, Sci. Data, vol. 3, issue 1, p. 160018, DOI: 10.1038/sdata.2016.18.