



EMODnet



European Marine
Observation and
Data Network

EMODnet Thematic Lot n°VI – Biology

EMFF/2019/1.3.1.9/Lot 6/SI2.837974 -EMODnet Biology

Start date of the project: 18/04/2021 - (24 months)

EMODnet Phase IV

Deliverable 2.1: Inventory of possible historical data resources within the consortium





Disclaimer¹

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 EASME or of the European Commission. Neither the EASME, nor the European Commission, guarantee the accuracy of the data included in this study. Neither the EASME, the European Commission nor any person acting on the EASME's or on the European Commission's behalf may be held responsible for the use which may be made of the information.

Document info

Title	D2.1: Inventory of possible historical data resources within the consortium
WP title	WP2: Access to marine biological data
Task	Task 1: Maintain and improve a common method of access to data held in repositories.
Authors	Dimitra Mavraki (HCMR) < dmavraki@hcmr.gr > Leen Vandepitte (VLIZ) < leen.vandepitte@vliz.be >
Dissemination level	Public
Submission date	26/10/2021
Deliverable due date	29/10/2021

¹ The disclaimer is needed when the document is published



Contents

1	Inventory of possible historical data resources.....	4
1.1	List of proposed datasets	5
1.2	Prioritisation Process for the Datasets	8
1.3	Next Steps.....	8



D2.1: Inventory of possible historical data resources within the consortium

1 Inventory of possible historical data resources

In EMODnet Phase IV, the Work Packages on 'Data access to marine biological data' and 'Data archaeology and rescue' are merged. As both Work Packages deal with data management, standardisation, quality control and interoperability of marine data, it is no longer felt relevant to split up recent and historical data in separate Work Packages. Although processing of historical data and data rescue in most cases starts with digitisation, all other necessary steps to make the data available and interoperable are common.

The main objective for WP2 is covered in Task 1: Maintain and improve a common method of access to data held in repositories. The data covered by this inventory will primarily include the following groups: algae, angiosperms, benthos, birds, fish, mammals, phytoplankton and zooplankton in European seas, more specifically defined in six regions: Arctic, Atlantic, Baltic Sea, Black Sea, Mediterranean Sea and North Sea, including their coastal and estuarine zones. Data from other regions are also covered within the proposal, even though it is not the main focus of the ongoing work.

As it was underlined in the previous EMODnet phase III (https://www.emodnet-biology.eu/sites/emodnet-biology.eu/files/public/documents/EMODnet_Biology_III/Deliverables/D3.7.pdf), in scientific terms, historical biodiversity data are crucial for assessments with the same credibility as new data (Griffin E. (2019)). They are valuable for studies on biodiversity loss, for the design of future samplings and for predictions of future trends. Furthermore, historical data facilitate the comprehensive and integrative conservation plans since past patterns and processes are compared with current ones (McClenachan et al. (2012)). The recently launched European Green Deal provides a bright example of the necessity of estimating the effects of climate change and environmental degradation on the economy and welfare of our societies. Historical Data Resources therefore, can offer the kind of evidence needed for conservation policy and marine resource management.

Historical data are usually in the form of specimens stored in Biodiversity Collections and in historical documents, (also known as legacy, ancient or simply old documents), which comprise past-period's literature stored in an analogue and/or obsolete format. These old documents cannot be found online; but are available only in institutional libraries, publications, books, expedition logbooks, project reports, newspapers (Mavraki D. et al. (2016)) or other types of legacy formats (stored on floppy disks, microfilms, or cd). That is the ultimate target of this inventory: to mobilise the participating partners to find in-house historical datasets.

The inventory of possible historical data resources within the consortium was initiated over the Summer (July - September 2021), requesting all partners to provide an overview of in-house historical datasets. Within EMODnet Biology and this Work Package, historical data are defined as data collected at least 70



years ago (before 1950). The inventory was however also used to identify possible rescue datasets, defined as 'data published between 1950-2000s, which currently only exist on paper or in text files.

The outcome of this inventory is presented in Table 1. The list of the proposed resources contains 18 datasets in total and includes among others, information on the current file format, the type of data, the geographical, taxonomic, and temporal scope. During Phase IV, partners will still be able to identify historically relevant data, which can be added to this inventory at a later time.

1.1 List of proposed datasets

The table presents the 18 proposed datasets. For every dataset, the temporary title is given along with the acronym of the EMODnet partner (Organisation or Institution) which proposed the relevant dataset. The third and fourth columns provide information regarding the format of the file and if the dataset is already digitised (not in paper format) and in which format the digitised document is available (file format). Additional columns present the language in which the dataset is written, the marine region(s) covered, the temporal and the taxonomic coverage.

It is worth mentioning that apart from the above-mentioned information, participating partners provided also more detailed information regarding: the source of the dataset (e.g. library), the country/sub-region that is covered by each one of the datasets, a short description of the proposed datasets, as well as available measurements, such as presence/absence and abundance.

Table 1: List of proposed historical datasets

EMODnet Partner	Temporary title	Digitised File dataset format		Language	Marine Region	Temporal cover	Taxonomic coverage
		yes	format				
HCMR	R.H. Whitehouse (1933). Report on fish eggs and larvae taken during 1931	yes	PDF	EN	Mediterranean Sea	1931	Fish
HCMR	Report on the Danish Oceanographical expeditions 1908-1910 to the Mediterranean and adjacent seas. Vol II Biology. J.1 Mediterranean Ceratia By E. Jorgensen (1920)	yes	PDF	EN	Mediterranean Sea	1908-1910	Phytoplankton
HCMR	Topsent, E. 1894. Éponges du Golfe de Gabès	yes	PDF	FR	Mediterranean Sea	1892	Benthos

HCMR	Molluscan diversity and distribution of Heraklion bay	yes	PDF	EN/GR	Mediterranean Sea	1990-1991/2014-2015	Benthos
HCMR	CINCS - Pelagic-benthic Coupling IN the oligotrophic Cretan Sea	yes	Excel	EN/GR	Mediterranean Sea	1994-1995	Benthos, Fish
HCMR	Report on the Danish Oceanographical expeditions 1908-1910 to the Mediterranean and adjacent seas. Vol II Biology. J.3 Mediterranean Tintinnidae By E. Jorgensen (1924)	yes	PDF	EN	Mediterranean Sea	1908-1910	Zooplankton
HCMR	Report on the Danish Oceanographical expeditions 1908-1910 to the Mediterranean and adjacent seas. Vol II Biology. J.2 Mediterranean Dinophysiaceae By E. Jorgensen (1923)	yes	PDF	EN	Mediterranean Sea	1908-1910	Phytoplankton
HCMR	Walter Klie (1935) V. Ostracoda. Die fischereigrunde vor Alexandrien. Notes and Memoirs of the Fisheries Research Directorate of Egypt, 12	yes	PDF	DE	Mediterranean Sea	1933	Zooplankton
HCMR	Report on the Danish Oceanographical expeditions 1908-1910 to the Mediterranean and adjacent seas. Vol II Biology. A.8 Lepadogaster By Frederic Guitel (1919)	yes	PDF	FR	Atlantic Ocean	1904-1910	Fish
HCMR	MATER a study of the benthic environment at selected stations of the South and North Aegean Sea	yes	Excel	EN/GR	Mediterranean Sea	1997-1998	Benthos

EMODnet Partner	Temporary title	Digitised dataset	File format	Language	Marine Region	Temporal cover	Taxonomic coverage
HCMR	Forbes, E. (1843). On the Radiata of Eastern Mediterranean. I. Ophiuridae. Transactions of the Linnean Society of London 19: 143-153	yes	PDF	EN	Mediterranean Sea	1843	Benthos
HCMR	Forbes, E. (1843). On the species of Ophiura inhabiting the Aegean Sea. Proceedings of the Linnean Society of London 1: 174-177	yes	PDF	EN	Mediterranean Sea	1843	Benthos
OGS	Algae in Mediterranean Sea islands (1876-1877)	yes	PDF	IT	Mediterranean Sea	1876-1877	Algae
OGS	Mollusca in Mediterranean Sea, off the coasts of Tunisia and Western Libya (spring 1913)	yes	PDF	IT	Western Libya	1913	Benthos
OGS	Phytoplankton North Adriatic PRISMA2 Project 1996-1997	yes	Excel	IT	Mediterranean Sea	1996-1997	Phytoplankton
SYKE	Algaline semiquantitative phytoplankton data from the Baltic Sea	yes	Excel	FI/EN	Baltic Sea	1993-2011	Phytoplankton
VLIZ	Scientific papers on the Belgian Oceanographic Expedition (1948-1949) along the African Coast (South Atlantic Ocean)	yes	PDF	FR	South Atlantic Ocean	1948-1949	Benthos, Fish
VLIZ	Scientific results of the Mercator training ship expeditions	yes	PDF	FR/DE	South Atlantic Ocean	1937-1951	Benthos, Fish



1.2 Prioritisation Process for the Datasets

The temporal coverage (date coverage) is an essential factor to prioritise the list of the 18 datasets, taking into account the “pre-1950” definition of historical data within this project. To be more specific, five out of 18 datasets cannot be considered as historical because the period they cover is more recent than 70 years ago:

1. Molluscan diversity and distribution of Heraklion Bay (1990-1991/2014-2015),
2. CINCS - Pelagic-benthic Coupling IN the oligotrophic Cretan Sea (1994-1995),
3. MATER (MAss Transfer and Ecosystem Response) a study of the benthic environment at selected stations of the South and North Aegean Sea (1997-1998),
4. Phytoplankton North Adriatic PRISMA2 Project 1996-1997 (1996-1997)
5. Algaline semiquantitative phytoplankton data from the Baltic Sea (1993-2011).

In conclusion, for the purposes of this inventory, the following 13 datasets are identified as Historical Data Resources:

6. R.H. Whitehouse (1933). Report on fish eggs and larvae taken during 1931
7. Report on the Danish Oceanographical expeditions 1908-1910 to the Mediterranean and adjacent seas. Vol II Biology. J.1 Mediterranean Ceratia By E. Jorgensen (1920)
8. Report on the Danish Oceanographical expeditions 1908-1910 to the Mediterranean and adjacent seas. Vol II Biology. J.2 Mediterranean Dinophysiaceae By E. Jorgensen (1923)
9. Report on the Danish Oceanographical expeditions 1908-1910 to the Mediterranean and adjacent seas. Vol II Biology. J.3 Mediterranean Tintinnidae By E. Jorgensen (1924)
10. Topsent, E. 1894. Éponges du Golfe de Gabès
11. Walter Klie (1935) V. Ostracoda. Die fischereigrunde vor Alexandrien. Notes and Memoirs of the Fisheries Research Directorate of Egypt, 12
12. Report on the Danish Oceanographical expeditions 1908-1910 to the Mediterranean and adjacent seas. Vol II Biology. A.8 Lepadogaster By Frederic Guitel (1919)
13. Forbes, E. (1843). On the Radiata of Eastern Mediterranean. I. Ophiuridae. Transactions of the Linnean Society of London 19: 143-153
14. Forbes, E. (1843). On the species of Ophiura inhabiting the Aegean Sea. Proceedings of the Linnean Society of London 1: 174-177
15. Algae in Mediterranean Sea islands (1876-1877)
16. Mollusca in Mediterranean Sea, off the coasts of Tunisia and Western Libya (spring 1913)
17. Scientific papers on the Belgian Oceanographic Expedition (1948-1949) along the African Coast (South Atlantic Ocean)
18. Scientific results of the Mercator training ship expeditions

1.3 Next Steps

The work performed in the previous EMODnet phases has allowed the consortium to establish a solid foundation of European marine data and a network spread across countries. In phase IV, WP2 aims to continue expanding and consolidating the network in order to make more data available. WP2 will



continue to use the standards, vocabularies and data formats implemented in previous phases, thereby providing consistency and ensuring interoperability for providers and users:

- The OBIS-ENV Darwin Core (DwC) format, not only allowing the inclusion of presence/absence information of marine biodiversity data, but also the storage of additional measurements or facts sampled alongside with the biological sampling
- The World Register of Marine Species (WoRMS), the authoritative and comprehensive list of names of marine organisms worldwide
- The Marine Regions Gazetteer, a standard list of marine georeferenced place names and areas
- The BODC controlled vocabularies, lists of standardised terms that cover a broad spectrum of disciplines of relevance to the oceanographic and wider community

Throughout the next year, a selection of the 13 datasets identified as Historical Data Resources, will be used to test existing online citizen science platforms, through which these datasets can potentially be standardised (according to the above mentioned standards and vocabularies) to flow to the EurOBIS database and subsequently, EMODnet Biology. The test will include - amongst others - the user friendliness of the platforms both from the data manager and the citizen volunteers' perspective, and the general speed (time required) to digitise a dataset through these platforms. If one or more platforms are positively evaluated, these can serve as a digitisation resource during the remaining time of Phase IV.

The next steps - to be carried out by October 2022 - can be summarized as follows:

- Identify possible suitable citizen science platform(s)
- Pre-usage test and evaluation of the identified platform(s)
- Upload the digitally available format (pdf or Excel) of one or more datasets per platform
- Mobilise a group of citizen scientists within each platform to carry out the formatting and quality control steps
- Establish a data flow from the citizen science platform to EurOBIS and EMODnet Biology
- General post usage test and evaluation of the used platform(s) in the framework of the EMODnet Biology needs
- Reporting on the efforts undertaken in rescuing historical data through citizen science by the end of October 2022 (=M18).

1.4. References

1. Griffin E. (2019), Getting necessary historical data out of deep freeze. *Polar Science* 21: 238-239. <https://doi.org/10.1016/j.polar.2019.05.008>
2. McClenachan et al. (2012), From archives to conservation: why historical data are needed to set baselines for marine animals and ecosystems. *Conservation Letters* 5 (5): 349-359. <https://doi.org/10.1111/j.1755-263X.2012.00253.x>
3. Mavraki D. et al. (2016), Rescuing biogeographic legacy data: The "Thor" Expedition, a historical oceanographic expedition to the Mediterranean Sea. *Biodiversity Data Journal* 4 <https://doi.org/10.3897/BDJ.4.e11054>