

EMODnet Thematic Lot n° V – Biology

D2.4.3 Report on the availability of data following the workshops

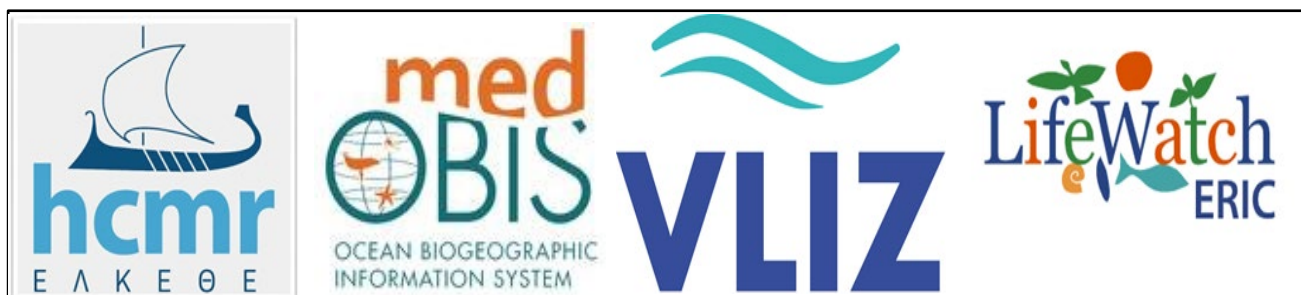
CINEA/EMFAF/2022/3.5.2/SI2.895681

Start date of the project: 10/05/2023 (24 months)

Centralisation Phase

**Report on the availability of data following the workshops
[D2.4.3]**





Disclaimer

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 CINEA or the European Commission. Neither the CINEA, nor the European Commission, guarantees the accuracy of the data included in this study. Neither the CINEA, the European Commission nor any person acting on the CINEA'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 (and reference)	D2.4.3 Report on the availability of data following the workshops
WP title (and reference number)	WP2- Data Management
Task (and reference number)	Task 1- Maintain and improve a common method of access to data held in repositories
Authors [affiliation]	Dimitra MAVRAKI [Hellenic Centre for Marine Research] Vasilis GEROVASILEIOU [Ionian University & Hellenic Centre for Marine Research] Ioannis RALLIS [Hellenic Centre for Marine Research] Melina LOULAKAKI [Hellenic Centre for Marine Research]
Dissemination level	Public
Submission date	2025-03-07
Deliverable due date	2025-03-18

Contents

1. Understanding the need of organising a dedicated Workshop for the Mediterranean	3
2. Autumn Data School	6
2.1 Collecting data from the Mediterranean Region, DATAthon	7
2.1.1. From data collection to standardisation and publication	8
2.1.2. Final Outcome of the DATAthon	9
2.2. Citizen Science platforms facilitate the rescue and standardisation process for historical marine biodiversity data	10
2.2.1. Final Outcome of the Citizen Science Workshop	12
2.3. Autumn Data School Poll and Next Steps	12
3. References	14
4. Annexes	15
AUTUMN DATA SCHOOL DATASETS AVAILABLE ONLINE	15
Agenda of the Autumn Data School	18
Certification	20
Poll Results	21
Workshop Photos	23

Report on the availability of data following the workshops

1. Understanding the need of organising a dedicated Workshop for the Mediterranean

EMODnet Biology is the EU's operational service for in-situ marine biodiversity data, with actions focusing mainly on data publishing, the creation of data products, and the publication of informative material with a focus on, but not limited to, the European Seas and the following groups: (macro)algae, angiosperms, benthos, birds, fish, mammals, phytoplankton, reptiles, and zooplankton. EMODnet Biology provides open and free access to interoperable data and data products on the temporal and spatial distribution of marine species covering the whole marine community.

The European contribution to the global effort of gathering marine biodiversity data is significant with EMODnet playing a pivotal role: more than 43 million records are made available by [EMODnet Biology](#).

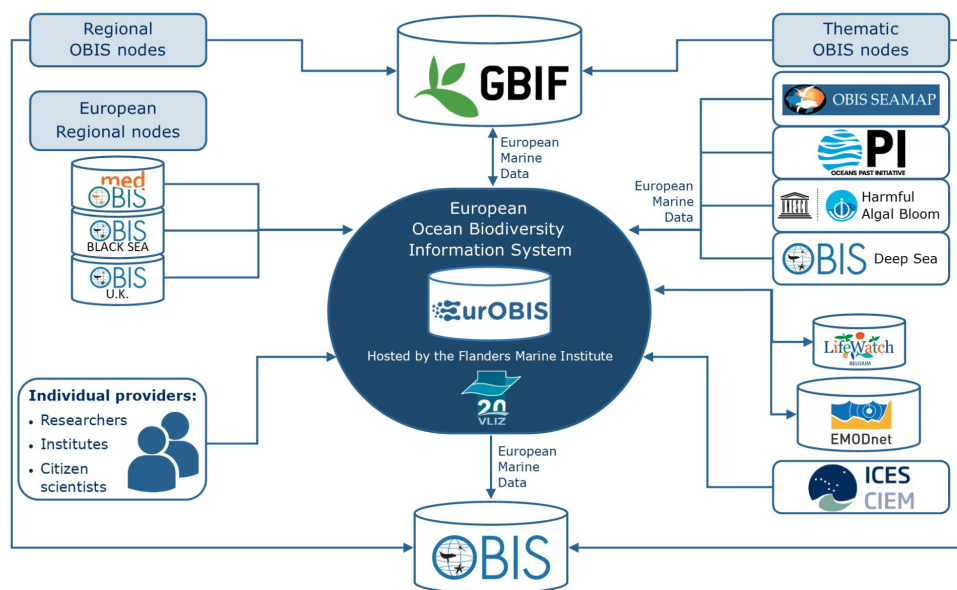


Figure 1: Data flow for marine biodiversity data

Specifically, for the Mediterranean Region, medOBIS, the Regional OBIS node for the Mediterranean Sea hosted by the Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC) of the Hellenic Centre for Marine Research (HCMR), has as its main mission to build and maintain an alliance that collaborates with scientific communities across the Mediterranean Sea, to facilitate free and open access to marine biodiversity data.

Figure 1 illustrates the data flow for marine biodiversity data, HCMR (host of medOBIS) is a part of the EMODnet Biology consortium and the data published by this OBIS node finds its way to the EMODnet Biology

via established data flows. Other data providers to EMODnet Biology comprise of ad-hoc providers, research projects, other OBIS nodes, other initiatives/data repositories. All the data are available in EMODnet Biology and is harvested by OBIS and GBIF in the IPT where it was first published. Ongoing collaboration with these two biodiversity initiatives ensures that duplicates are not found in the system.

OBIS is a global open-access data and information clearinghouse on marine biodiversity for science, conservation, and sustainable development. More than 20 OBIS nodes connect 500 institutions from 56 countries. OBIS is now the world's most comprehensive database on the diversity, distribution, and abundance of life in the ocean in time and space. It is built by a collective effort of thousands of scientists and data managers employed by hundreds of institutions around the world who brought over 50 million observations of 120,000 marine species into the public domain. GBIF is the equivalent initiative that allows for the access to not only marine but also terrestrial and freshwater biodiversity.

At the time of writing this report, the number of datasets published in the EMODnet portal is 1,450 with more than 43 million occurrence records and approximately 107 million extended Measurement or Facts (eMoF) records (<https://www.eurobis.org/>). A detailed breakdown of available occurrence records for each region and EMODnet Biology phase over the last 15 years, is illustrated in Figure 2.

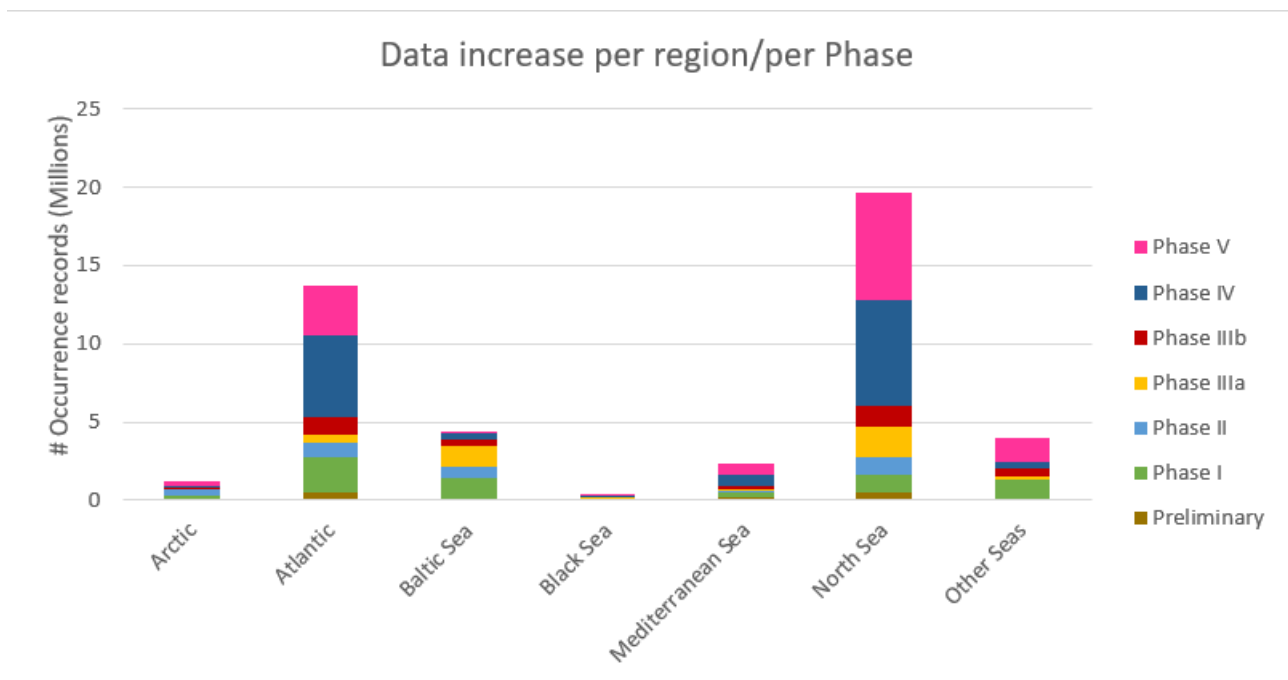


Figure 2: Data increase during each Phase per European Sea region

By the end of Phase V's first year, EMODnet Biology's holdings were composed of 98.7% open data (CC-BY licenses or unrestricted data) and only 1.2% restricted data, this demonstrates the success the initiative has had over the past 1.5 decades in expanding its network of data providers and promoting Open Science and FAIR principles (Wilkinson et al. 2016).

Despite these efforts to collect data, as shown in Figure 2, the total number of biological occurrence records for the Mediterranean Region remains low compared to other marine regions such as the Atlantic Ocean, North and the Baltic Seas. As stated in the United Nations Environment Programme/Mediterranean Action Plan and Plan Bleu (2020), this knowledge gap is the result of many factors, particularly socioeconomic ones, such as uneven demographic development, geopolitical difficulties - mainly in the southern part of the Mediterranean - and high unemployment rates among individuals with advanced education (tertiary level). The aforementioned factors combined with limited awareness of open science best practices (Bezuidenhout

et al. 2018) result in a lack of appropriate data collection and the availability of sufficient long-term time series (Malvarosa et al. 2023) for this specific Region.

Following the acknowledgment of this marine data gap in the Mediterranean, a dedicated workshop was organised to collect biogeographic data from the Region. The medOBIS team, with its expertise in standardisation and quality control procedures for marine biodiversity data, and a good knowledge of the Mediterranean Sea and its biota, as well as its network across the various countries bordering the Mediterranean, took the lead in this effort.

It is important to mention that, since 2013, medOBIS has served as a Marine Research Repository for both historical and modern scientific datasets, to which the principles of FAIR and Open data are applied (Mavraki et al. 2022). MedOBIS accepts data files in any format and publishes the data following specific standardisation and quality control procedures (Figure 5). Currently, medOBIS hosts more than 70 marine datasets covering the period from 1841 to 2024 and contains approximately 100,000 digitally-accessible records, comprising 4,285 taxa, with the top five Classes being Teleostei (15.7%), Scyphozoa (14.2%), Polychaeta (13.1%), Hydrozoa (8.25%) and Gastropoda (7.54%).

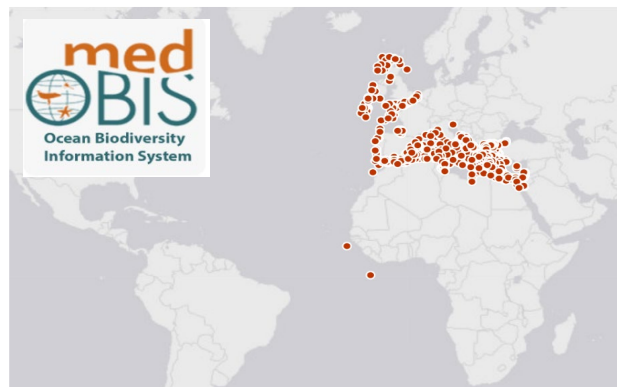


Figure 3: medOBIS Datasets distribution

2. Autumn Data School

For the Phase V workplan, it was discussed that a dedicated Data Workshop for the non-EU countries bordering the Mediterranean Sea should take place. The medOBIS team took on the responsibility of organising, with the support from VLIZ and LifeWatch ERIC (LW), this three-day, in-presence, and free workshop; and gathering researchers mainly from the southern part of the Mediterranean, which is one of the marine areas with the most important data gaps in the Mediterranean basin. The Autumn Data School was advertised through the EMODnet [news page](#) and the LifeWatch ERIC [event page](#).

The Workshop took the form of an “Autumn Data School”. The first two days were organized as a DATAthon (see section 2.1. below), an event where participants gather to solve practical problems through the application of data science tools and techniques, by working together in teams to generate insights and potential solutions. The DATAthon focused on transferring knowledge on data management best practices (from data collection to publication). Participants were required to bring their own dataset(s) to work on and follow all the steps up to its final submission to Marine Research Repositories, such as medOBIS and the LifeWatch ERIC Metadata Catalogue, with a focus on facilitating its reuse. Participants also benefited from applying a DOI to their datasets, as well as by publishing a data paper in the *Biodiversity Data Journal* (<https://bdj.pensoft.net>) with the cost covered by LifeWatch ERIC.

The third day focused on teaching participants how to use Citizen Science platforms, such as Zooniverse and the LifeWatch Citizen Science Platform (see section 2.2. below). The purpose of this training session was to create a Citizen Science project and engage volunteers in the historical data rescue and standardization process. For this session, historical datasets were selected by workshop organisers, based on the work done in the previous phase of EMODnet Biology (report available via the [link](#)).

The **main objectives** of the Autumn Data School were the following:

- a) bringing together Researchers from the non-EU countries bordering the Mediterranean Region and establishing a steady communication and data collaboration channel between them,
- b) sharing best practices on data standardization and data publication on Research Infrastructures and finally,
- c) addressing the data gap that exists for this marine region, by encouraging researchers to bring and share their own marine biodiversity datasets during this three-day workshop.

Several internal HCMR meetings took place and an outline of the organisational targets was set up, the team created a list of potential participants focusing on North African countries with which several research partnerships were established in the past. These countries were Algeria, Egypt, Morocco, Lebanon, Libya and Tunisia. The outcome was a list of potential Organisations and Institutes (Table 1), which were contacted to inform that the data workshop was being planned with a focus on data-poor marine regions.

More than 30 invitations were sent to potential participants, representing 23 different organizations from six countries, and 12 replies expressing their interest in joining the workshop were received. The next step was to make a selection among the applications, based on a set of criteria. For that reason, the following criteria were established, and were communicated to interested individuals through personal communication and e-mails: a. provide a list and description of the dataset(s) covering the Mediterranean Sea, that each participant will bring to the Workshop and b. send a motivation letter regarding its participation and commitment to fulfil the requirements of the Workshop.

The outcome of this communication was the finalisation of a list of in-person participants for the Data Workshop: five (5) early career researcher scientists from Tunisia, one (1) university professor from Algeria, and one (1) early career researcher from Italy, from Sapienza University. Due to visa constraints, which

generally limited the number of participants, one person from Tunisia was only able to follow the workshop online, which was a last-minute arrangement to accommodate the constraints experienced by this colleague.

Conversely, the Citizen Science Workshop invitation was open to Mediterranean Organisations and advertised through various channels, including the MedOBIS list, the EMODnet [news page](#), and the LifeWatch ERIC event page. Registrations for online participation exceeded 35. Finally, 19 people joined the Workshop.

Table 1. List of participating Organisations

Country	Affiliation
Tunisia	TunSea Association
Tunisia	Université de Carthage, Institut National des Sciences et Technologies de la Mer, Laboratoire de Biodiversité Marine
Tunisia	Laboratoire de Biodiversité et Ecosystèmes Aquatiques, Département des Sciences de la vie, Faculté des Sciences de Sfax
Tunisia	Notre Grand Bleu association
Algeria	Universités Benyoucef Benkhadda Université Alger 1

LW ERIC expressed interest in co-organising the Workshop, and the final agreement was that they would host the Autumn Data School, in Lecce, Italy, from 13 to 15 November 2023. LW ERIC also supported the initiative by covering the travel costs (flights and accommodation) for eight (8) participants and providing catering services during the workshop. Their support was valuable in assisting with VISA arrangements for participants and providing two trainers for the Workshop.

This collaboration between EMODnet Biology and LW ERIC proved to be mutually beneficial: the new datasets resulting from the workshop are now available in both medOBIS and [LW ERIC Metadata Catalogue](#), the [LW Citizen Science platform](#) was also tested and eventually, the training material was made available through the [LW ERIC training platform](#).

2.1 Collecting data from the Mediterranean Region, DATAthon

The DATAthon took place in Lecce from 13-14 November 2023, as an in-person event, with seven (7) participants and five (5) trainers. The trainers' team consisted of two trainers from the medOBIS team, two from the VLIZ team, and one from the LW ERIC team. The agenda of the meeting is available in Annex II.

The DATAthon focused on transferring knowledge on data management best practices (from the collection of the dataset until its publication). The participants brought their own datasets to work on and followed all the steps up to their final submission to Marine Research Data Repositories, such as MedOBIS, with the focus on facilitating its reuse by, for example, subsequent republication in EMODnet Biology, OBIS, and GBIF (Global Biodiversity Information Facility).

The learning outcomes were:

- Understanding the FAIR principles, explaining the importance of data sharing, and recognizing the benefits of data standardization;
- Familiarity with the Darwin Core (DwC) standards and know where to locate DwC terms;
- Collecting and documenting metadata;

- Preparing their datasets for publication in medOBIS.

Upon completing the standardisation of their datasets, participants were awarded with a Certificate of Completion (see Annex III) and also were given the opportunity to publish their data in the *Biodiversity Data Journal* as open-access data papers at no cost.

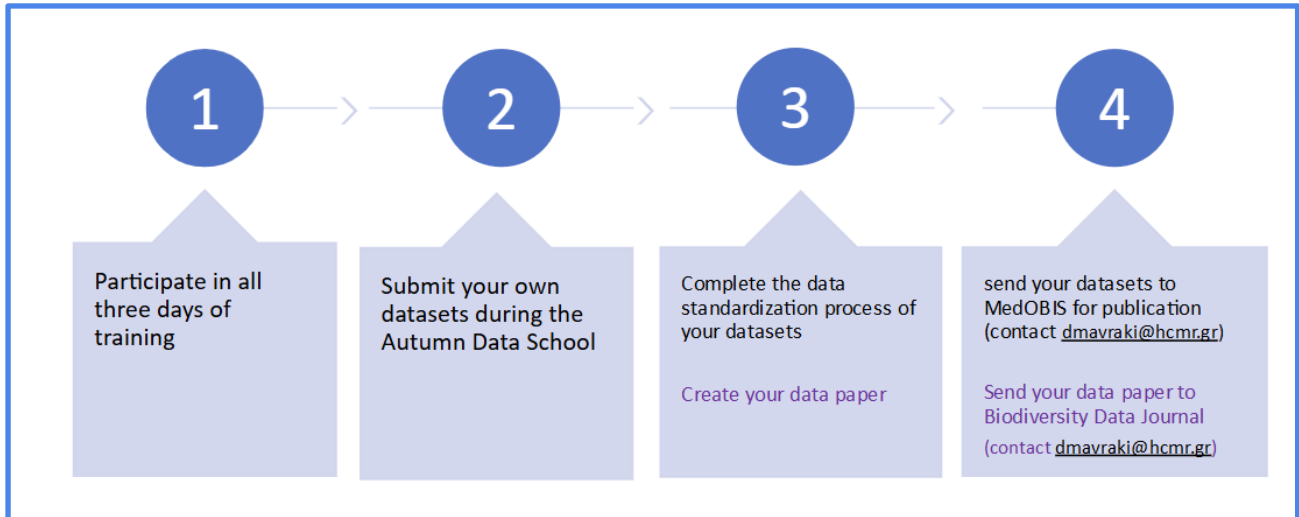


Figure 4: Process of awarding participants with the Workshop Certification

2.1.1. From data collection to standardisation and publication

During the DATAthon the participants became familiar with key concepts such as FAIR and open data. The trainers emphasised the importance of collecting descriptive and accurate metadata records. The first exercise required trainees to describe their dataset by providing essential information, including temporal coverage, taxonomic coverage, citation details, and the people involved in fieldwork. To standardise this process, a metadata template was shared, which participants used to document their dataset information.

In the next step, the participants were asked to focus on their own dataset. After a detailed introduction to DwC standards by the trainers, participants applied them to standardise their datasets ensuring that key attributes such as species names, locations, and dates were properly structured.

Since this process was challenging, trainers provided one-on-one support where necessary. Emphasis was given to data quality control, such as correcting species' scientific names, including accurate locations and/or formatting dates according to standard vocabularies and ontologies. Afterwards, the datasets were structured into a Darwin Core-compliant format, typically as CSV or XLSX files, with each row representing an individual occurrence or record.

Finally, datasets were validated by using the LifeWatch& EMODnet Biology QC tool ([Biocheck Tool](#)), before being submitted to [medOBIS IPT](#). The MedOBIS team conducted the final quality check on both metadata and data provided by each participant before their publication and the data were subsequently republished in EMODnet Biology (Figure 5).

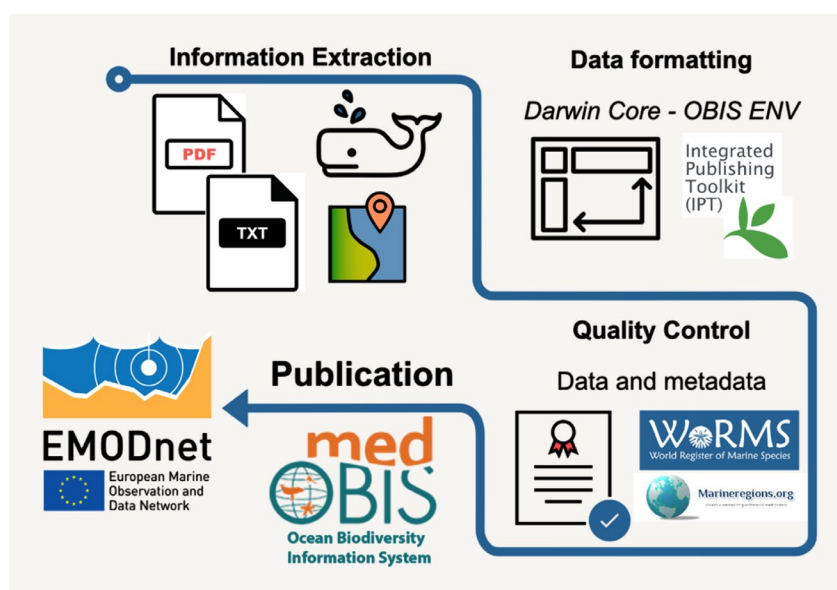


Figure 5: medOBIS data workflow

2.1.2. Final Outcome of the DATAthon

As a result of the above-mentioned process, five (5) new datasets were successfully collected, standardized and published in MedOBIS IPT. A comprehensive representation of these datasets is presented in Annex I, whereas a concise version is available in the table below. It should also be noted that two additional datasets are still in progress, primarily due to the nature of the data (sea turtles stranding data and lessepsian siganid parasite data).

Table 2. List of published datasets after the Autumn Data School

Dataset Title	New Dataset	Historical Dataset
List of marine species observed during diving in Monastir area, Tunisia, 2023	✓	
Tunisian Polychaetes by 2022	✓	
TunSea, 2023. Cryptogenic, Rare and Non-indigenous marine species in Tunisian coastline, recorded by Citizen Scientists	✓	
Seagrass meadow, macrofauna communities, and environmental variables in Southern coast of Algerian Basin (El Melah Lagoon; Algeria) 2019 and 2020	✓	
Yellow legged Gull nesting in Kuriat Island in 2022, Tunisia	✓	
"Pola" Expedition: Mollusks, in the Eastern Mediterranean and the Adriatic Sea, 1890-1894		✓
Expedition of Melita 1892, Mollusks collected on the coasts of Tunisia and Algeria		✓

2.2. Citizen Science platforms facilitate the rescue and standardisation process for historical marine biodiversity data

The medOBIS team has pioneered the workflow for mobilizing historical marine biodiversity data from the [Biodiversity Heritage Library](#) (BHL). Leveraging [Zooniverse](#), one of the world's largest and most popular Citizen Science platforms, the team developed a project where volunteers transcribe and standardize data from historical scientific papers that Optical Character Recognition (OCR) software fails to process accurately. This collaborative approach can significantly reduce the effort and time required for data managers and marine scientists to mobilize these datasets.

The team located papers from historic expeditions containing significant data on the Mediterranean marine region. One focused on the Central and Eastern Mediterranean, while the other covered the Tunisian and Algerian coasts, both were considered data-deficient regions. It is acknowledged that the southern Mediterranean is generally regarded as a data-deficient region. While Tunisia has comparatively more data than other North African countries (e.g., Algeria, Libya, Egypt), it remains data-deficient. Additionally, some unpublished data do exist in the form of reports and other documents, but often in languages other than English, such as Arabic and, in some cases, French.

After converting the datasets into images, the team designed tasks with clear instructions for citizen scientist participants to transcribe and simultaneously format the data according to DwC standards (*Figures 6 and 7*).

On the final day of the workshop, which was a hybrid event, online and in-presence participants attended lectures on the usefulness of Citizen Science and its potential benefits to the scientific community. The training team consisted of two trainers from the MedOBIS team and one trainer from the LW ERIC team. The agenda of the meeting is available in Annex II.

At the beginning of the session, MedOBIS and LW trainers explained the process of building a new project on both Zooniverse and LW Italy Citizen Science platforms. Participants were then guided through an exercise where they followed structured instructions and educational material to create their own Citizen Science projects. Later in the day, they acted as Citizen Scientists, testing both the predesigned and newly developed projects to identify potential errors and, at the same time, contribute to the transcription and standardization process.

Throughout the workshop, trainers requested assistance from online and in-person participants to complete simple, well-described tasks, ultimately rescuing research data from two historical datasets. This approach not only provided Citizen Scientists with valuable knowledge and skills but also streamlined data mobilization efforts for the scientific community, saving significant time and effort. The historical datasets consist of monitoring data covering the period from 1890 to 1894, documenting Mollusca records from dredging events in the Eastern Mediterranean (Expeditions 1890, 1891, 1892, 1893) and in the Adriatic and the Strait of Otranto (Expedition 1894).

The projects and workflows developed in Zooniverse can be accessed at [Digging Up the Ocean's Past](#):

- **Workflow 1:** *Zoologische Ergebnisse VII, Mollusken I. Gesammelt von S.M. Schiff "Pola" 1890-1894, Dr Rudolf Sturany (#25374)*
- **Workflow 2:** *Mollusques Recueillis Sur Les Côtes De La Tunisie Et De L'Algérie (#25410)*

Shortly after the workshop, both projects were completed in terms of transcription and standardization. Once the results were retrieved and cleaned, the MedOBIS team applied quality control to verify the accuracy of the Citizen Scientist contributions. The datasets were then quality controlled using the [Lifewatch](#) and [EMODnet Biology QC Tool](#) and successfully published in the MedOBIS IPT, mobilizing these valuable historical biodiversity records. These data were subsequently made available in EMODnet Biology.

digging up the oceans' past ABOUT CLASSIFY TALK COLLECT RECENTS LAB

Great work! Looks like this project is out of data at the moment!
[See the results](#) or dismiss this message

Nr.	Expedition und Datum	O. L. N. Br.	Tiefe	Opera-tion	Grund	Arten
1	14. Aug. 1890	19°48'20" 39°23' Westlich von Corfu	615 m	Kleine Kurze		<i>Trophon vaginatus</i> Jan. <i>Fusus rostratus</i> Oliv. <i>Nassa limata</i> Chemn. <i>Raphitoma superiua</i> Tib. nov. var. <i>pseudacanthodes</i> . <i>Rissoa cimicoides</i> Forb. <i>subsolida</i> Aradas. <i>Bulla atriculus</i> Brocchi juv. <i>Achaeon pusillus</i> Forb. <i>Dentalium (Antalis) agile</i> Sars. <i>Siphonocentium quinqueangulare</i> Forb. <i>Syndosuya longicollis</i> Seacchi. <i>Noera costellata</i> Desh. <i>rostrata</i> SpGr. <i>Kelliella miliaris</i> Phil. <i>Circe minutus</i> Mont. <i>Arca scabra</i> Poli. <i>(Bathyarca) koreui</i> Dan. <i>Lima (Limatula) sarsii</i> Lovén. <i>subauriculata</i> Mont.

FIELD GUIDE ✕

- WoRMS
- Darwin Core Standard
- Taxonomy
- Datasets
- Find the Correct Coordinates

Figures 6 (top) and 7 (bottom): Snapshots of the Citizen Science project designed on Zooniverse and the "Field Guide" including instructions and supplementary material with tools and guidelines.

2.2.1. Final Outcome of the Citizen Science Workshop

It is important to highlight that the number of historical datasets provided currently by the Biodiversity Heritage Library (BHL) concerning marine species and including at least 100 taxa is almost 2000 (Paragkamian et al. 2022). In recent years medOBIS has made considerable efforts to digitize and publish historical papers. As a result, 36 out of the 78 datasets currently published in medOBIS IPT are historical. However, despite these efforts, this number remains relatively low when compared to the vast quantity of available historical papers.

This low number is due to the fact that the data rescue work is resource-consuming and requires specific expertise that is not easily found. Specifically, several factors turn the curation process of historical documents into a serious challenge (Paragkamian et al. 2022); in terms of georeferencing constraints, location names or sampling points on an old map may be provided instead of the actual coordinates, common taxonomic constraints (e.g. old, currently unaccepted synonyms, lack of authority associated with the taxon names) are combined with the absence of taxonomic literature or voucher specimens (e.g. identifier number for samples of natural history/expedition collections) require the taxonomists' assistance, numerical measurement units often need to be converted to the International System of Units (SI system) (e.g. fathoms to meters), old toponyms and political boundaries that have now changed should also be taken into consideration, as well as coordinates that now fall on land instead of in the sea, due to the changes in the coastline. Lastly, the use of languages other than English is quite common in old scientific publications, so multilingual curators are required.

As a result of the Citizen Science Workshop, two (2) additional historical datasets have been successfully digitized, standardized, and published in MedOBIS IPT. A comprehensive representation of these datasets is presented in Annex I.

2.3. Autumn Data School Poll and Next Steps

According to the poll conducted during the Autumn Data School, researchers from North African countries highlighted the critical importance of sharing data management practices to facilitate data collection and dissemination. In this context, both the EMODnet Biology and medOBIS will continue their efforts by organizing dedicated data workshops in non-EU Countries bordering the Mediterranean and enhancing the availability of reliable data for this area. These initiatives aim to pave the way towards more sustainable and responsible data management policies.

To ensure the successful organization of similar data workshops in the future, experience has shown that the following parameters should be considered: a. To increase the number of participants it would be beneficial for such initiatives to take place in the southern Mediterranean, to avoid visa issues and reduce travel costs, b. Should the workshops be held within the territory of the European Union, providing financial support for participants serves as a key incentive to encourage attendance and engagement, c. Technical constraints do exist and may prevent certain participants from accessing the tools and software typically used in European initiatives and projects.

The poll results are available in Annex IV. It is important to mention that the poll underscores a diverse and multi-sector community actively engaged in marine research, with strong representation from NGOs, academia, and citizen science initiatives. Participants from institutions such as TunSea, University Algiers 1, and the National Institute of Marine Sciences and Technologies (INSTM) emphasized the urgent need for improved data sharing and management across the Mediterranean Region. While citizen science plays a vital role in data collection, respondents identified key challenges, including difficulties in merging and digitizing datasets and coordinating efforts among diverse institutions and stakeholders. Additionally, many

participants highlighted the need to strengthen collaboration through regional training programs and enhanced infrastructure, underscoring the importance of establishing sustainable data management systems.

Furthermore, before the Autumn Data School, familiarity with platforms such as EMODnet Biology and LifeWatch ERIC was limited, reinforcing the need for increased outreach and capacity-building initiatives. Many respondents also stressed the importance of regional collaboration, advocating for targeted training programs, cross-border initiatives, and the creation of a dedicated scientific community to foster innovation and knowledge exchange.

These insights reaffirm the significance of initiatives such as EMODnet Biology and MedOBIS in improving data accessibility and promoting a more integrated, sustainable, and responsible approach to marine research and data management in the Mediterranean region.

The vast wealth of marine data must not remain inaccessible in individual archives or institutional libraries. The value of marine data is exponentially enhanced when it is made openly available and stored in contemporary databases and repositories. MedOBIS is committed to rescuing information that remains hidden in literature while simultaneously providing a support service to assist scientists in publishing their data in an open and FAIR-compliant repository for the Mediterranean Sea.

3. References

- United Nations Environment Programme/Mediterranean Action Plan and Plan Bleu (2020). State of the Environment and Development in the Mediterranean. Nairobi.
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>
- <https://planbleu.org/en/soed-2020-state-of-environment-and-development-in-mediterranean/>
- Malvarosa L., Basilone G., Carbonara P., Carpentieri P., Cozzolino M., Follesa M.C., Gambino M., Gancitano V., Parreno Duque D., Pesci P., Vielmini I., Scarcella G. (2023). Data availability and participatory approach: the right mix for enhancing Mediterranean fisheries' sustainability. *Front. Mar. Sci.* 10:1155762. doi: 10.3389/fmars.2023.1155762
- Bezuidenhout L., Chakauya E. (2018). Hidden concerns of sharing research data by low/middle-income country scientists. *Glob Bioeth.* Feb 26;29(1):39-54. doi: 10.1080/11287462.2018.1441780. PMID: 29503603; PMCID: PMC5827722.
- Mavraki D., Gerovasileiou V., Sarafidou G., Panteri E., Chatzinikolaou E. (2022). MedOBIS: a marine biodiversity data repository for the Mediterranean Sea. In Proceedings of 15th International Congress on the Zoogeography and Ecology of Greece and Adjacent Regions, Hellenic Zoological Society (2022), p. 80.
- Leen Vandepitte, Dimitra Mavraki, Ruben Perez Perez, Georgia Sarafidou, Savvas Paragkamian, Vasilis Gerovasileiou, Laura Marquez (2022). Report on efforts undertaken in rescuing historical data through citizen science. EMODnet Phase IV deliverable report available at https://emodnet.ec.europa.eu/sites/emodnet.ec.europa.eu/files/public/Biology/D2.3_0-Report%20on%20efforts%20undertaken%20in%20rescuing%20historical%20data%20through%20citizen%20science.pdf
- Paragkamian S, Sarafidou G, Mavraki D, Pavludi C, Beja J, Eliezer M, Lipizer M, Boicenco L, Vandepitte L, Perez-Perez R, Zafeiropoulos H, Arvanitidis C, Pafilis E and Gerovasileiou V (2022) Automating the Curation Process of Historical Literature on Marine Biodiversity Using Text Mining: The DECO Workflow. *Front. Mar. Sci.* 9:940844. doi: 10.3389/fmars.2022.940844

4. Annexes

AUTUMN DATA SCHOOL DATASETS AVAILABLE ONLINE

1. **Dataset name:** TunSea, 2023. Cryptogenic, Rare and Non-indigenous marine species in Tunisian coastline, recorded by Citizen Scientists
Dataset citation: Chebaane S (2024). TunSea, 2023. Cryptogenic, Rare and Non-indigenous marine species in Tunisian coastline, recorded by Citizen Scientists. Version 1.5. Hellenic Center for Marine Research. Samplingevent dataset. <http://ipt.medobis.eu/resource?r=sahar&v=1.5>
Data provider: TunSea Association
Data temporal coverage: 2005-05-07 / 2023-08-29
Data geographic coverage: Tunisian Exclusive Economic Zone (EEZ)
 Data can be found via:
 medOBIS IPT: <http://ipt.medobis.eu/resource?r=sahar>
 EMODnet Catalogue:
<https://emodnet.ec.europa.eu/geonetwork/srv/eng/catalog.search#/metadata/6d617269-6e65-696e-666f-000000008753>
 OBIS: <https://obis.org/dataset/0097a1c2-1696-4cc9-9641-4895e0eda497>
 This dataset has appeared in 924 downloads in 2025, with a total of 56,117 records.

2. **Dataset name:** List of marine species observed during diving in Monastir area, Tunisia, 2023
Dataset citation: BOUHLEL R (2023): List of marine species observed during diving in Monastir area, Tunisia, 2023. v1.2. Hellenic Center for Marine Research. Dataset/Samplingevent <http://ipt.medobis.eu/resource?r=tunsea09-2023&v=1.2>
Data provider: TunSea Association
Data temporal coverage: September 2023
Data geographic coverage: Monastir Bay, Tunisia
 Data can be found via:
 medOBIS IPT: <http://ipt.medobis.eu/resource?r=tunsea09-2023>
 EMODnet Catalogue:
<https://emodnet.ec.europa.eu/geonetwork/srv/eng/catalog.search#/metadata/6d617269-6e65-696e-666f-000000008471>
 OBIS: <https://obis.org/dataset/443dc75f-b127-4dfe-807b-0604c46edf14>
 This dataset appeared in 949 downloads in 2025, with a total of 16,663 records.
 GBIF: <https://www.gbif.org/dataset/7b30f242-1317-495e-84e9-417d440ef07a>

3. **Dataset name:** Tunisian Polychaetes by 2022
Dataset citation: AYARI KLITI R (2024). Tunisian Polychaetes by 2022. Version 2.0. Hellenic Center for Marine Research. Occurrence dataset. <https://doi.org/10.25607/iauoit>
Data provider: Institut National des Sciences et Technologies de la mer (INSTM)
Data temporal coverage: 1912-01-01 / 2022-12-31
Data geographic coverage: Tunisian coasts
 Data can be found via:
 medOBIS IPT: <http://ipt.medobis.eu/resource?r=polychaetetestunia>
 EMODnet Catalogue:
<https://emodnet.ec.europa.eu/geonetwork/srv/eng/catalog.search#/metadata/6d617269-6e65-696e-666f-000000008698>
 OBIS: <https://obis.org/dataset/f996a819-8a22-4306-97eb-ca40dee61263>
 This dataset appeared in 816 downloads in 2025, with a total of 174,745 records.

4. **Dataset name:** Seagrass meadow and macrofauna communities in Southern coast of Algerian Basin (El Mellah Lagoon), 2019
Dataset citation: HAMZA H, De Wit R, Mammeria A B (2025). Seagrass meadow and macrofauna communities in Southern coast of Algerian Basin (El Mellah Lagoon), 2019. Version 1.3. Hellenic Center for Marine Research. Samplingevent dataset.
<http://ipt.medobis.eu/resource?r=mammeria&v=1.3>
Data provider: University Algiers
Data temporal coverage: 2019-02-20 / 2019-12-20
Data geographic coverage: El Mellah Lagoon, Algeria
 Data can be found via:
 medOBIS IPT: <http://ipt.medobis.eu/resource?r=mammeria>
 EMODnet Catalogue:
<https://emodnet.ec.europa.eu/geonetwork/srv/eng/catalog.search#/metadata/6d617269-6e65-696e-666f-000000008754>
 OBIS: <https://obis.org/dataset/88b335cb-7d9f-4bb9-9d0a-839fcc753ea6>
 This dataset appeared in 984 downloads in 2025, with a total of 76,554 records.
5. **Dataset name:** Yellow legged Gull nesting in Kuriat Island in 2022, Tunisia
Dataset citation: Notre Grand Bleu and Coastal Protection and Management Agency. Yellow legged Gull nesting in Kuriat Island (2023), Tunisia.
Data provider: Notre Grand Bleu
Data temporal coverage: 2022-04-26 / 2022-04-27
Data geographic coverage: Monastir Bay, Kuriat islands
 Data can be found via:
 medOBIS IPT: <http://ipt.medobis.eu/resource?r=amjed>
 EMODnet Catalogue: This dataset has recently been uploaded and have not yet been harvested by EMODnet Biology
 OBIS: <https://www.obis.org/dataset/447e0dd9-669e-4caf-aa4f-2b91e68d40df>
6. **Dataset name:** "Pola" Expedition: Mollusks, in the Eastern Mediterranean and the Adriatic Sea, 1890-1894
Dataset citation: Dimitra Mavraki, Ioannis Rallis and Melina Loulakaki (2024). Digitization of "Pola" Expedition: Mollusks, in the Mediterranean and the Adriatic Sea, 1890-1894.
Data availability: Biodiversity Heritage Library (biodiversitylibrary.org)
Data temporal coverage: 1890-08-14 / 1894-07-26
Data geographic coverage: Eastern Mediterranean and the Adriatic Sea.
 Data can be found via:
 medOBIS IPT: <http://ipt.medobis.eu/resource?r=pola>
 EMODnet Catalogue:
<https://emodnet.ec.europa.eu/geonetwork/srv/eng/catalog.search#/metadata/6d617269-6e65-696e-666f-000000008696>
 OBIS: <https://obis.org/dataset/e0de7c96-c16f-4815-b4f4-2d2a24e497dc>
 This dataset appeared in 975 downloads in 2025, with a total of 175,722 records.
 GBIF: <https://www.gbif.org/dataset/3be44a65-238e-4412-80ad-3e47c2775b85>
7. **Dataset name:** Expedition of Melita 1892, Molluscs collected on the coasts of Tunisia and Algeria
Dataset citation: Loulakaki M, Rallis I, Mavraki D (2024). Expedition of Melita 1892, Molluscs

collected on the coasts of Tunisia and Algeria. Version 1.1. Hellenic Center for Marine Research. Samplingevent dataset. <http://ipt.medobis.eu/resource?r=melita&v=1.1>

Data availability: Biodiversity Heritage Library (biodiversitylibrary.org)

Data temporal coverage: 1892-08-29 / 1892-10-19

Data geographic coverage: Coasts of Tunisia and Algeria

Data can be found via:

medOBIS IPT: <http://ipt.medobis.eu/resource?r=melita>



EMODnet Catalogue:


<https://emodnet.ec.europa.eu/geonetwork/srv/eng/catalog.search#/metadata/6d617269-6e65-696e-666f-000000008756>

OBIS: <https://obis.org/dataset/b48e27b2-707e-4eef-823f-5782287e42f7>

This dataset appeared in 1,229 downloads in 2025, with a total of 194,173 records.

Agenda of the Autumn Data School



Agenda 13th of November (Day 1)

Day 1: Welcome - Policy, Scientific & Technical Context – Practical Session

09:00 – 09:30	Registration	
09:30 – 09:45	Opening Welcome	Dimitra Mavraki (HCMR)
09:45 – 10:00	Introduction to EMODnet and EMODnet Biology	Leen Vandepitte (VLIZ)
10:00 – 10:20	Data archaeology & rescue. The data management workflow	Vasilis Gerovasileiou (Ionian University & HCMR)
10:20 – 10:40	FAIR data and data standardization process	Dimitra Mavraki (HCMR)
10:40 – 11:00	Data Qualification process	Lynn Delgat (VLIZ)
11:00 – 11:30	Coffee Break	
11:30 – 12:30	Short presentation by each participant (5 min per participant)	
12:30 – 13:30	Lunch Break	
13:30 – 15:00	Practical Session: working on your datasets (focus on metadata)	
15:00 – 15:30	Coffee Break	
15:30 – 16:30	Practical Session: working on your datasets (focus on data)	
16:30 – 17:00	Questions & Wrap up Day 1	






Agenda 14th of November (Day 2)

Day 2: Practical Session (continue from previous day)

09:00 – 09:30	LifeWatch ERIC Metadata Catalogue	Lucia Varia (LifeWatch ERIC)
09:30 – 11:00	Practical Session: working on your datasets (focus on data standardization)	
11:00 – 11:30	Coffee Break	
11:30 – 12:30	Practical Session: working on your datasets (focus on data standardization)	
12:30 – 13:30	Lunch Break	
13:30 – 15:00	Practical Session: working on your datasets (focus on data quality)	
15:00 – 15:30	Coffee Break	
15:30 – 16:45	Participants' feedback (10min per participant)	
16:45 – 17:00	Questions & Wrap up Day 2	

Agenda 15 th of November (Day 3)		
Day 3: Data rescue via citizen science platforms		
9:00 - 9:30	Registration	
09:30 – 09:40	Opening Welcome Overview of the workshop's objectives and structure	Dimitra Mavraki (HCMR)
09:40– 10:00	Presentation of the LifeWatch Citizen Science Platform	Marco Elia (LifeWatch ERIC)
10:00 – 11:00	Introduction to the Zooniverse platform. Case studies of a Zooniverse project focused on marine data transcription.	Ioannis Rallis (HCMR)
11:00 – 11.30	Coffee Break	
11:30 – 12:30	Hands-On Session - Create Your Own Zooniverse Project	
12:30 – 13.30	Lunch Break	
13:30 – 15:00	Hands-On Session – Act as Citizen Scientist	
15:00 – 15:30	Coffee Break	
15:30 – 16:30	Hands-On Session – Act as Citizen Scientist (cont.)	
16:30 – 17:00	Questions & Closure of workshop	

Certification

Autumn Data School

CERTIFICATE OF PARTICIPATION

This is to certify that

"Participant Name"

has actively participated in the three-day Autumn Data School 13-15 November 2023 co-organised by EMODnet Biology and LifeWatch ERIC. The participant has demonstrated their knowledge and understanding of key topics in the field of Research Data Management by successfully completing the final assignment.

Learning objectives

- Identify FAIR data principles;
- Recognise and apply Darwin Core Standards;
- Apply data quality control procedures (for taxonomy and locations);
- Experiment with Citizen Science online tools such as the LifeWatch Citizen Science Platform and Zooniverse.

Dimitra Mavraki
Environmental Data Manager, IMBCC-HCMR



Prof. Alberto Basset
LifeWatch ERIC Service Centre Director



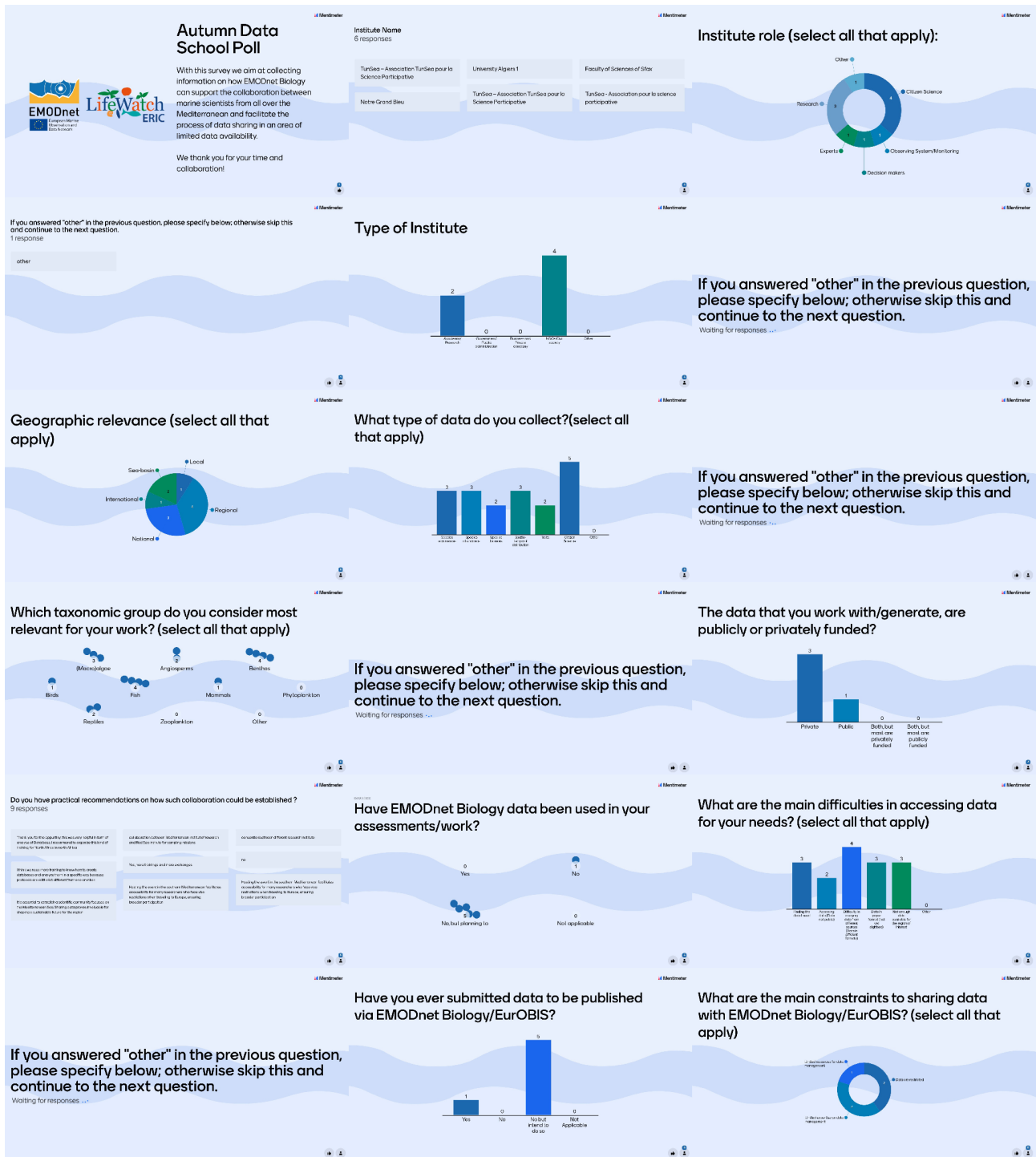

lifewatch.eu



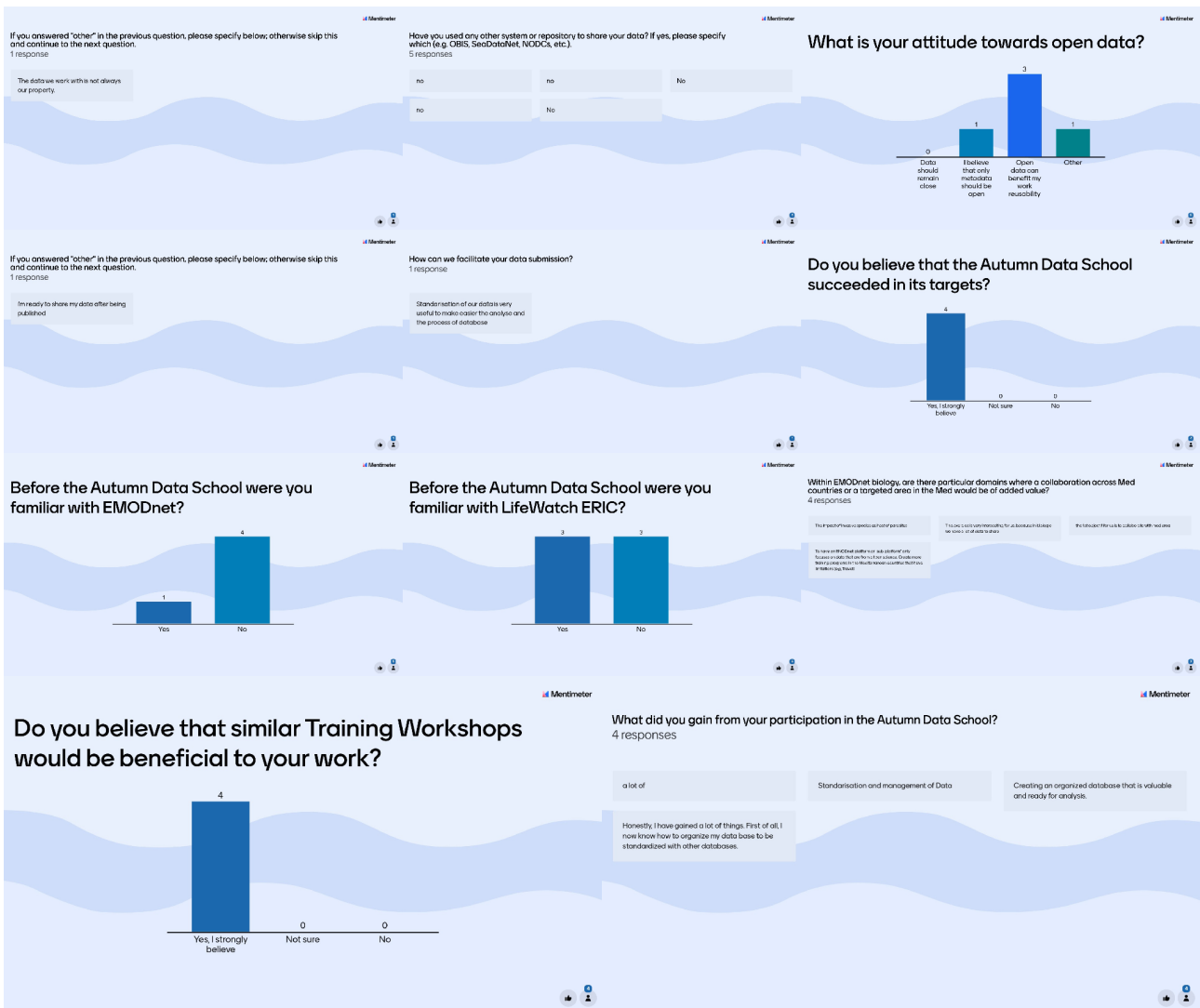
emodnet.ec.europa.eu

Autumn Data School
Lecce, Italy, 13-15 November 2023

Poll Results



Report on the availability of data following the workshops [D2.4.3]



Workshop Photos



