



Topic	H2020 – INFRAIA-2018-2020
Short Title	Eurofleets+
Title	An alliance of European marine research infrastructures to meet the evolving requirements of the research and industrial communities
Project Number	824077
Delivery Date	11/07/2023
Deliverable No	4.12
Lead Beneficiary	RBINS
Dissemination Level	Public

D4.12: Second data management follow-up report for funded cruises



Document information	
Document Name	D4.12: Second data management follow-up report for funded cruises
Document ID	D4.12 Second data management follow-up report for funded cruises RBINS v1.4.docx
Revision	1.4
Revision Date	2023-07-11
Author	Iona Sissy, Le Hong Minh, Partescano Elena
Security	Public

Approvals			
	Name	Organisation	Date
Coordinator	Aodhán Fitzgerald	Marine Institute	
Activity Coordinator	Anneli Strobel	AWI	
WP Leader	Nicole Biebow	AWI	

History			
Revision	Date	Modification	Author
v1.0	2023-05-22	First creation of chapters structure and bullet points	Hong Minh
V1.1	2023-06-19	Narrative content, progress overview tables and case-by-case summary for RBINS and OGS cruises	Hong Minh Le, Elena Partescano
V1.2	2023-06-21	Addition of cruises managed by HCMR in case-by-case summary and Table 1	Hong Minh Le, Elena Partescano, Sissy Iona
V1.3	2023-07-05	Review of HCMR and OGS	Hong Minh Le, Elena Partescano, Sissy Iona
V1.4	2023-07-11	Merged coordinator's comments into the document	Hong Minh Le, Elena Partescano, Sissy Iona

This document contains information, which is proprietary to the EUROFLEETS+ consortium. Neither this document nor the information contained herein shall be used, duplicated or communicated by any means to any third party, in whole or in parts, except with prior written consent of the EUROFLEETS+ Coordinator. The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user there for uses the information at its sole risk and liability

TABLE OF CONTENTS

Contents

1. Cruise data management.....	1
2. Data Management Plan	6
3. Cruise Summary Report	7
4. Cruise data handling	8
5. Case-by-case overview.....	10
6. Ingestion progress.....	35
7. Difficulties and lessons learned	36
8. Coming data management strategy.....	38

1. Cruise data management

The EUROLLEETS+ project currently counts a total of 28 cruise projects distributed over 23 cruises scheduled between September 2020 and June 2023. It includes four Co-PI projects paired to four of the existing cruises. One RTA project (IsoMed) has been granted in May 2022 and has been paired to TalPRO 2022 cruise. POSEIDON cruise has been granted and added to the scheduled cruises on March 2023. Four projects scheduled in 2023 have been cancelled (ICON, IDEA, ROSEBUD and SLOGARO II) and are already deduced from the total count of 28. Some cruises have been postponed of several months from their initial schedule (GRACE, CABLE leg 1) where OASIS has been postponed from July 2022 to March 2023. Also, ERODOTO formerly planned in June 2023 has been postponed for 2 weeks in July because of late granting of permits for the cruise.

Data collected during the cruises need to be managed according to the FAIR guiding principles (Findable, Accessible, Interoperable, and Reusable). In order to achieve FAIRness, all Principal Investigators (PIs) have submitted a full Data Management Plan (DMP) before their cruise which ensures the collection, preservation and dissemination of the data to be properly planned. Two weeks after the cruise, the Cruise Summary Report (CSR) needs to be submitted to the SeaDataNet CSR front-end. To achieve FAIRness, EUROLLEETS+ project prescribes that all data collected during the cruise are to be made openly available and published on open data repositories with SeaDataNet (SDN) or EURODIS infrastructures as reference end points by the end of the project. However, an embargo of up to 2 years can be requested by the data providers in order to publish their results in scientific journals.

Three Responsible Data Centres (RDCs) i.e. HCMR, RBINS, OGS are in charge of the cruise data management. The granted project cruises have been assigned evenly between the RDCs according to the surveyed areas and the data types. Due to its experience in handling bathymetric data, HCMR has been assigned the bathymetric datasets. Some cruises presenting bathymetric datasets and other dataset types are shared between two RDCs. Their task of data curator consists in making sure the data reach eventually the FAIR principles. They ensure the DMPs are submitted in time, they check whether they are well and enough described and suggest updates to become more complete. They check whether requested embargoes do not exceed 2 years. They ensure the CSRs are submitted on time and give support for their submission. But the major work of the RDCs is the cruise data curation. Each cruise conducts its specific scientific objective for which various amount of datasets are collected with a large diversity of data types. To achieve FAIRness of the data, the RDCs ensure the cruise data is complete, findable, accessible, interoperable and reusable which requires for being interoperable and reusable a harmonization process of the data.

The cruises are assigned as presented in Table 1 between the 3 RDCs. HCMR is leading the DMPs and CSRs follow-up for 8 cruises, OGS for 10 cruises and RBINS for 9 cruises. Regarding the data curation, HCMR is responsible for the datasets of 13 cruises, OGS and RBINS both for 11 cruises.

At the present stage, all the cruises, except ERODOTO, already have taken place. Data publication and transfer have started and a close follow-up of the data submissions is performed by the RDCs who are in continuous contact with the PIs to lift their data to reach the FAIRness objectives.

ID	Year	RV	Cruise Name	Study Area	Location	Timing	DMP DC	Data DC	DMP	CSR	Embargo	Transferred	DIP1	CDIs	SEA NOE
3	2020	UGOT Hugin AUV	Focus-AUV	Geology, Marine Biology, Sedimentology	Kaikōura Canyon, New Zealand	Sept 30th - Oct 29th	HCMR	HCMR							
8	2021	Celtic Explorer	PORO-CLIM	Climate dynamics, Geophysics, Sedimentology, Training	NE Atlantic (S Rockall Plateau; Eriador Seamount; Porcupine Basin & Ridge; East Thulean Rise)	05-30th May	HCMR	HCMR							
1	2021	Pelagia	iMAR	Biological Oceanography, Geology, New technologies	Mid-Atlantic Ridge inside the Portuguese EEZ of the Azores	May 18th - June 3rd	OGS	HCMR OGS							
5	2021	DANA	GSHARK	Marine Biology	Bredefjord, Greenland.	July 30 - 12 August	RBINS	RBINS							
2	2021	GO Sars_ROV Aegir	BENCHMARK	Marine Biology Physical Oceanography	Denmark Strait	01-10 August	RBINS	HCMR RBINS							
7	2021	Tubitak Marmara	PHYCOB	Biological Oceanography	Western Black Sea	9th - 15th Sept	RBINS	RBINS							
4	2021	SOCIB	GRASSMAP	New technologies, Marine Biology	Mallorca and Cabrera islands	14th - 20th September	OGS	OGS							
6	2021	Aegeo	MYRTOON	Climate dynamics	Eastern Mediterranean, SW Aegean Sea, Myrtoon Basin	29 Sept - 08 Oct	HCMR	HCMR							
9	2022	Pelagia	CALYPSO	Physical Oceanography	Alboran or Balearic Sea	16 Feb - Mar 10	OGS	OGS							
10	2022	Aranda	CABLE	Biological Oceanography	Gulf of Finland, Baltic Proper	22-29 April & 10-13th October	OGS	OGS							

ID	Year	RV	Cruise Name	Study Area	Location	Timing	DMP DC	Data DC	DMP	CSR	Embargo	Transferred	DIP1	CDIs	SEA NOE
10	2022	Aranda	DOMUSE (Co-PI on CABLE)	Biogeochemistry, biological oceanography	Gotland Deep Baltic Sea	22-29 October 14th October	OGS	OGS							
11	2022	Belgica II_AUV Barabas	GRACE	Geology Geophysics Physical Oceanography Sedimentology	Ceuta Canyon and adjacent areas (West Moroccan Mediterranean margin).	28 April - 11 May	HCMR	HCMR							
11	2022	Belgica II_AUV Barabas	SEAQUAKE (Co-PI on GRACE cruise)	Geology, Geophysics, New Technologies, Sedimentology	Ceuta Canyon and adjacent areas (West Moroccan Mediterranean margin).	28 April - 11 May	OGS	HCMR OGS							
14	2022	Belgica II	TAIPro2022	Physical Oceanography	Algero-Provencal Basin, Sicily Channel, Tyrrhenian Sea, Ligurian Sea	17-26th May	RBINS	RBINS							
14	2022	Belgica II	IsoMed (RTA during TalPro22)	Biological Oceanography	Tyrrhenian Sea	17-26th May	N/A	RBINS							
12	2022	Arni Friedrikson	SYNERGY	Marine Biology	North Western Iceland	18 - 25 June	OGS	HCMR OGS							
15	2022	Sanna	IOPD	Biogeochemistry Climate dynamics New technologies Marine Biology Polar Biology Training	Godhabfjord, Ameralik fjord and the shelf area connecting these fjords in Nuuk, Greenland	28th June - 10th July	RBINS	RBINS							

ID	Year	RV	Cruise Name	Study Area	Location	Timing	DMP DC	Data DC	DMP	CSR	Embargo	Transferred	DIP1	CDIs	SEA NOE
16	2022	Atlantic Explorer	FIGURE	Biological Oceanography Biogeochemistry Microbiology Physical Oceanography	Gulf Stream (NW Atlantic)	21-30 July	RBINS	RBINS							
16	2022	Atlantic Explorer	CARING (Co-PI on FIGURE cruise)	Biological Oceanography Biogeochemistry Microbiology Physical Oceanography	Gulf Stream (NW Atlantic)	21-30 July	RBINS	RBINS							
13	2022	Ramon Margalef	CARBO-ACID	Biogeochemistry, Marine Chemistry, Physical Oceanography	Iberian margin	3-12th August	HCMR	HCMR RBINS							
17	2022	Sanna	GLICE	Biological Oceanography Biogeochemistry Physical Oceanography	Disco Bay, West Greenland coastline	10-24 Aug	RBINS	RBINS							
18	2022	Sarmiento de Gamboa	SINES	Biogeochemistry	Northeast Atlantic, Western Iberian Margin.	11-20th September	OGS	OGS							
19	2023	Tangaroa	VISIT	Geophysics	East coast of North Island, New Zealand	31 March - 11 April	HCMR	HCMR OGS							
20	2023	Tangaroa	Hydee-Obs	Geophysics	Hikurangi Margin, North Island, New Zealand	21 - 31 March	HCMR	HCMR							

ID	Year	RV	Cruise Name	Study Area	Location	Timing	DMP DC	Data DC	DMP	CSR	Embargo	Transferred	DIP1	CDIs	SEA NOE
25	2023	RV Laura Bassi	POSEIDON	Geophysics	Ionian Islands	12-22 June	HCMR	HCMR							
23	2023	Thalassa_ROV	OASIS	Geology New technologies Marine Biology Physical Oceanography	SE Alboran Sea (W Mediterranean)	26 March - 7 April	OGS	OGS							
24	2023	Thalassa_ROV	UNSEEN (Co-PI on OASIS cruise)	Biological Oceanography, Pollutants or aerosols, Sedimentology	bassi	26 March - 7 April	RBINS	RBINS							
22	2023	RV Aegaeo_ROV Max Rover_AUV Barabas	ERODOTO	Deep Sea Research, Geology, Geophysics	Squillace Canyon, Italy	19-30 June 2023 1-12 July	OGS	HCMR OGS							

Table 2: Overview RDC assignment and data management progress follow-up

Legend:

	completed
	datasets curation started and in progress
	embargo
	data partially under embargo
	not applicable
	on hold until end of embargo
	missing

2. Data Management Plan

DMP review

To apply to the EF+ programmes, the PIs submitted a preliminary DMP on the EF+ DMP submission website (DMP EF+ Roadmap - <http://dmp.ef-ears.eu>). When granted for ship hours, PIs have been requested by the RDCs to complete their DMP in full two months at the latest before the campaign. When the full DMP is ready, assigned RDCs review its completeness and ask for updates when required. RTA cruise (IsoMed) is not required to provide a DMP but needs to provide an overview of the collected datasets, request of embargo, licence and availability time. RDCs send reminders when full DMP is not completed on time.

For 2021 and 2022 cruises, all DMPs have been completed. For 2023 cruises, four DMPs (VISIT, Hydee-Obs, OASIS, UNSEEN) are already completed and two (POSEIDON and ERODOTO) still need to be. A cruise-by-cruise description of the DMP follow-up is reported in Chapter 5. Almost all DMPs comply with FAIR data management principles and when there are not complete enough, updates have been requested. Ensuring that the DMP is updated after RDCs review can sometimes not be pursued by the cruise PI, despite many reminders.

Most of the DMPs were of very high quality with reoccurring remarks of:

- missing information on harmonisation process and missing references to standard vocabularies (the PIs were encouraged to include these mappings in the data files and update the DMP accordingly).
- limited lineage information

Embargo

It is also in the DMP that PI define whether the data should be kept under embargo and for how long (up to two years). When the requested embargo exceeded 2 years, the assigned RDC asked to reduce to up to two years. All concerned PIs (of cruises Focus-AUV, iMAR, BENCHMARK, FIGURE) accepted to align to the 2 years as prescribed except for OASIS which needs to be contacted to request the downgrade. From the completed DMP, only data of 4 cruises (GSHARK, GRASSMAP, ISOMED, SYNERGY) are not subject to embargo. 7 of the cruises embargoes end before the end of the project where 15 cruises data embargoes extend beyond the end of the project.

FAIRness implementation

The DMP is used to plan the data management in order to achieve the FAIR principles. Table 2 defines the monitoring indicators for its implementation. It is shown that, in the scope of EF+, the collected datasets are Findable and Accessible when metadata are published on EMODnet DIP, SeaDataNet SEANOE or SeaDataNet CDIs service (as open or restrictive access). Data are Interoperable and Re-usable when they are openly published on SeaDataNet CDIs service and harvesting EMODnet lots and EurOBIS.

Monitoring indicator	Description	Achieved when
1.1 Data is complete	Completeness of datasets is checked	Number of datasets and volume announced in DMP match with published ones. If numbers and/or volumes do not match, discrepancies should be justified.
2.1. Data is findable	Required metadata elements are present. If complete, the dataset can be pushed to phase 1 and further to phase 2 and published on metadata infrastructures.	Metadata are published on EMODnet DIP or on one of the SDN services (CDIs or SEANOE) or EUROBIS
2.2. Data is openly accessible	Data is openly accessible including data placed under embargo after embargo period. They should be eventually published on open repositories	Data are published openly on EMODnet DIP or one of the SDN services (CDIs, or SEANOE) or EUROBIS
2.3. Data is interoperable	Required meta-information are present and described with common vocabularies. Publication to phase 2 required all meta-information to be completed with standardized vocabularies. When all meta-information are completed, data can be published on interoperable infrastructures.	Data are published on SDN CDIs infrastructure or EUROBIS
2.4. Data is re-usable	As soon as data are published openly with CCO or CC-BY licenses, the data are re-usable	Data are published openly on one of the SDN services (CDIs, or SEANOE) or EUROBIS with CCO or CC-BY licenses

Table 2: FAIRness implementation indicators and how to achieve them.

The RDCs are now supporting the implementation of the DMPs while managing the collected data. The monitoring of their effective application is being checked with, as first indicator, the completeness of the generated datasets and associated files. When cruise data are available, assigned RDCs will follow-up the other FAIR indicators and proceed to their publication to the dedicated repositories.

3. Cruise Summary Report

The CSR needs to be submitted within 2 weeks after the campaign. No CSR needs to be submitted for the Co-PI projects since CSRs are reported by cruise. The former CSR submission interface hosted by BSH has been migrated in March 2022 on a new instance managed by IFREMER. PIs have been invited to submit their CSR on SeaDataNet CSR back-office instead of BSH following a well-documented procedure. The EF+ data management guidelines have been updated accordingly and all PIs have been informed by the RDCs of the new submission procedure. Support has been given by the RDCs and the sdn-userdesk when issues were encountered.

Most of the CSRs were submitted on time. Reminders for submission have been sent to the others and support given to assist the submissions. In particular, for GSHARK, iMAR, GRASSMAP, CALYPSO, CABLE, SEAQUAKE, MYRTOON and SYNERGY cruises, the assigned RDCs proceeded to the CSR submissions on behalf of the PI. In some cases, CSRs were not submitted on time because of a confusion between the CSR and the Cruise Report. The narrative Cruise Reports were submitted to AWI but no CSR were submitted on the dedicated instance. RDCs had to explain how to proceed.

All 2021, 2022 CSRs have been submitted successfully and are accessible on the SDN CSR back office (<https://csr.seadatanet.org/>). For already departed cruises of 2023, only CSR of Hydee-Obs has been submitted. CSRs for VISIT, POSEIDON, OASIS-UNSEEN cruises are still expected. The PI of VISIT cruise has submitted the CSR to the SDN system and currently is implementing changes that requested by the back office during the validation process before the CSR publication. CSR of ERODOTO cruise is only due for mid-July after the planned cruise beginning of July.

A dedicated web page on EUROFLEETS+ EVIOR website (<https://evior.eurofleets.eu/csr-intro.html>) contains the different links to the CSR interfaces (for submission and for browsing).

4. Cruise data handling

Data submission timeline

The data submission timeline is described in the Data Management guidelines as follows. It is first recommended to the data submitters (the person responsible for the data) to perform the necessary QC on the raw data so that it is ready for publication in the reference data repositories. The data submitter gathers all the raw data collected during the campaign and within 2 months after the campaign or as soon as the data becomes available (when lab results are ready), make the data available either by submitting it directly to EMODnet DIP or by transferring the datasets directly to their assigned RDC depending on the data ingestion pathway.

Data Ingestion workflow

For the scientific cruise management, ingestion of cruise data is processed by datasets. A cruise can result in many datasets with different availability times and embargo periods depending on the data types and data processing needs. The data submission stream depends on the presence of an embargo and on the data type.

- For “common” data types without embargo, data are openly published on EMODnet DIP (Phase 1) and then on the CDIs SeaDataNet infrastructure (Phase 2) where data are harvesting the corresponding EMODnet thematic infrastructure such as the EMODnet lots and EuroBIS.
- For “common” data types where an embargo period is requested, data are sent to the assigned RDC who will generate SeaDataNet CDIs with a restrictive access. The access will be changed to open licence at the embargo’s end which will allow data to be harvesting the thematic lots.
- For data types which do not fit into SeaDataNet CDIs infrastructure, data will only be published on EMODnet DIP but in case of embargo the data will be submitted to SeaDataNet SEANOE with a restricted access first. When the embargo ends, the SEANOE submission is automatically harvested into EMODnet DIP repository and follow the CDIs SeaDataNet workflow.
- Data can in any case be submitted to SeaDataNet SEANOE to get a DOI.

Data transfer and storage

Data transfer can be done through different channels. During data submission to EMODnet DIP, the data files are uploaded and available for the RDC for further processing. PI can also send directly to his/her RDC the files by mail or, for large files such as bathymetry, transfer the files via a cloud link which can be generated by the RDC. Transferred datasets are then downloaded and stored locally and/or replicate on a RDC cloud account such as Google Drive. For bathymetry data, the volume are often very large which can reach gigabytes with the largest single file for SYNERGY cruise of 800 GB. Storage space is therefore necessary and cloud storage is opted for its large capacity. In some cases, PI submit the data on other data repositories (e.g. PANGEA, SQUIDLE+) or on their own institute infrastructure where RDC can download the data.

It is expected that data providers perform the necessary QC on raw data and transfer processed and quality flagged data. Having the PIs performing these steps is justified by the fact that he/she is the one who knows best his/her data and its context and can therefore process the data in a suitable manner, given his/her expertise. Furthermore, processed data are more valuable and re-usable. For geophysical data, providing processed data instead of raw data is even more recommended since managing big data volumes of raw data is time and resources consuming.

Data curation by RDCs

Gathering complete information

The completeness and consistency of the transferred datasets are checked by the RDCs based on the full DMP as well as on the Cruise report which also gives useful information on used devices, methods, analysis and processing procedures. Missing or incomplete metadata and data are queried with the data providers.

Phase 1

Submission to EMODnet DIP can be performed by the PI him/herself or by the assigned RDC on his/her behalf. This was the case for FOCUS-AUV, PORO-CLIM, iMAR, GSHARK, GRASSMAP, SENERGY. In practice, when the cruise datasets are submitted to EMODnet DIP, the DIP master assigns the RDC as curator for further follow-up. The RDC reviews the datasets submission and update it if necessary by adding additional meta-information. Next, the data submitter approval is requested before publishing the dataset as Phase 1 on EMODnet DIP.

Phase 2

The data processing steps performed by the RDCs for ingestion into SDN CDIs service consists in:

- performing the 1st-line quality checks on provided data,
- formatting the data to one of the SDN exchange formats such as SDN odv format and to standardize the data according the NERC common vocabulary.
- Generate the Common Data Index (CDI) with all required metadata describing the dataset.

Depending on the data types, SDN software NEMO and MIKADO can be used to format the data.

When the CDIs are generated and published on the SDN CDIs service, the publication links are referred to in EMODnet DIP, achieving the Phase 2 of the DIP ingestion process.

For data under embargo, the PIs are expected to transfer the dataset directly to their assigned RDC as soon as available for the RDC to already start the data preparation for their submission to the CDIs service.

Phase 2 is heavier as the generation of CDIs requires significant work from the RDCs.

Progress follow-up

The data management progress is monitored cruise by cruise by their assigned RDC(s). A share excel file is accessible on the Project Share files to support the follow-up of the data management process. Each RDC is also keeping track of their contacts with the PIs. RDCs are discussing progress follow-up and good practices during the Project meetings as well as by emails. A detailed progress follow-up per cruise and per dataset is presented in the next chapter.

5. Case-by-case overview

The following overview gives the data management progress cruise by cruise. The data ingestion status is described dataset per dataset with the dataset phases indicated in green (transferred, submitted to DIP, CDIs published or published to SEANOE) which are already achieved. The 'R' in the green cells are related to 'Restrictive' access when the ingested datasets are under embargo. The cells coloured in grey represent a 'not applicable' phase.

Cruise Name	iMAR				
Cruise No.	1				
Cruise year	2021				
RDC	OGS				
	HCMR (bathymetric data management)				
Cruise dates	May 18th - June 3rd				
PI	Telmo Morato				
DMP					
Status	Full DMP reviewed				
Evaluation	Complete				
CSR					
Status	Submitted on time				
CSR link	https://csr.seadatanet.org/report/20213136				
Data management					
Embargo	No embargo for CTD and SVP data. Embargo of 2 years for ADCP and multibeam data.				
End of embargo	31/12/2024				
License	Public and Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical data	CTD				
	List of stations				
	List of deep-sea video stations				
	SVP				
	ADCP				R

Bathymetric data	Multibeam		R
Number of published CDIs	46 + 40 for multibeam data		
Progress follow-up	<p>16/06/2021 – Invitation to submit CSR</p> <p>17/06/2021 – Creation of CSR on behalf of PI</p> <p>22/09/2021 – Reminder to send the data</p> <p>23/09/2021 – PI sent the list of acquired data</p> <p>09/11/2021 – Reminder to send the data</p> <p>24/11/2021 – Reminder to send the data</p> <p>01/12/2021 – Reminder to send the data</p> <p>31/01/2022 - Virtual meeting with PI on data management</p> <p>23/02/2023 - HCMR asked PI to send the processed multibeam data</p> <p>02/06/2023 - HCMR sent to PI a reminder about multibeam data</p> <p>02/06/2023 - PI asked HCMR which would be the desirable format</p> <p>05/06/2023 - HCMR replied that geotiff would fit</p> <p>09/06/2023 - PI uploaded the bathymetry data for the 2021 and 2022 campaigns to wetransfer. Transferred files included geotiffs and the metadata (xml).</p> <p>12/06/2023 - HCMR asked PI why there are 2 cruises since the plans include only a 2021 campaigns</p> <p>15/06/2023 - HCMR uploaded 24 CDIs of multibeam data for the 2021 cruise</p> <p>19/06/2023 - HCMR uploaded 16 CDIs of multi-beam echosounders for the 2022 cruise</p>		
Planning	No further action for HCMR		

Cruise Name	BENCHMARK
Cruise No.	2
Cruise year	2021
RDC for DMP and CSR	RBINS
RDC for data handling	RBINS (physico-chemical and biological data)
	HCMR (bathymetric data management)
Cruise dates	August 1 – August 10
PI	Burgos Julian
DMP	
Status	Submitted on time by PI and reviewed by RDC
Evaluation	Very complete
Review	Overall full DMP completed on time and well detailed. RDC suggests to use more community standards and vocabularies and to add more lineage information.
CSR	
Status	Submitted on time
CSR link	https://csr.seadatanet.org/report/20213165
Data management	
Embargo	No embargo for physico-chemical data. No embargo for multibeam echosounders and backscatter data. Embargo of 2 years for biological data.
End of embargo	10/10/2023
License	CC-BY

Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical data	CTD				
	ROVCTD				
	bottle				
	TSG				
	ADCP				
Biological data	Taxonomic identification				
	DNA sequencing				
Bathymetric data	Multibeam echosounder				
	acoustic backscatter sensors				
Number of published CDIs	81 + 32 (16 CDIs for Multibeam echosounder and 16 for acoustic backscatter)				
Progress follow-up	PIs submitted physico-chemical and geophysical data on EMODnet DIP. Additional metadata were requested to the PI in order to push data to phase 2 (start and end time of transects, variable description and units, QC). Data originator performed quality checks and added quality flags to all physico-chemical data. Data have been standardized and formatted according SeaDataNet reference vocabularies (using NEMO and MIKADO) including the flags. Data have been published on the SeaDataNet CDIs service. Multibeam echosounder and backscatter data were downloaded from EMODnet Ingestion as geotiff grids. 16 CDIs for multibeam echosounders on a 10m resolution and 16 CDIs for backscatter data on a 5m resolution were prepared by HCMR and submitted to SDN/CDI on Nov. 2022. The process of bathymetric data for the cruise is completed.				
Planning	ADCP data needs supplementary formatting using new release of NEMO 2.0.1. PI informed in February 2023 that the analyses of biological samples and pictures are mostly finished and that data will be sent to RDC as soon as completed.				
Difficulties, lessons learned	Description of the variables and units as well as measuring instruments have been requested to the data originator as well as quality checks and flags				

Cruise Name	FOCUS-AUV
Cruise No.	3
Cruise year	2020
RDC	HCMR
Cruise dates	September 30 – October 28
PI	Dr Joshu Mountjoy
DMP	
Status	Submitted on time by PI and reviewed by RDC.
Evaluation	Clear and informative DMP
Review	PI updated the full DMP concerning the embargo time and reduced it from 3 to 2 years following the project specifications. RDC suggested to use more community standards and vocabularies and to improve data interoperability.

CSR					
Status	Submitted on time				
CSR link	https://csr.seadatanet.org/report/20213027				
Additional remark	CSR has been created and submitted by PI using BSH CMS				
Data management					
Embargo	2 years (it was 3 years before the DMP review)				
End of embargo	29/10/2022				
License	Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Bathymetric/geological data	multi-beam echosounders			R	
	Acoustic seafloor Backscatter			R	
	Sub-bottom Profiler			R	
	SideScanSonar			R	
Number of published CDIs	53				
Progress follow-up	<p>18/01/2022 - HCMR started the download of multibeam and backscatter data</p> <p>21/01/2022 - HCMR started the download of the remaining files. Due to the large size of the files (700 Gb) and missing files that were identified by HCMR, the downloading was delayed</p> <p>11/10/2022 - Downloading completed. Files are kept in google Drive</p> <p>08/03/2023 - HCMR upload to SeaDataNet 12 CDIs for backscatter data</p> <p>11/03/2023 - HCMR upload to SeaDataNet 20 CDIs for multibeam data</p> <p>23/05/2023 - HCMR upload to SeaDataNet 9 CDIs for Sub-bottom Profiler data</p> <p>31/05/2023 - HCMR upload to SeaDataNet 12 CDIs for SideScanSonar</p>				
Planning	HCMR ask PI if embargo period can expire.				
Difficulties, lessons learned	It should be more clear to the PIs that the Project needs processed instead of raw data. Managing big data volumes of raw data is time and resources consuming.				

Cruise Name	GRASSMAP
Cruise No.	4
Cruise year	2021
RDC	OGS
Cruise dates	September 14 – September 20
PI	Miguel Massot Campos
DMP	
Status	Full DMP reviewed
Evaluation	Complete
CSR	
Status	Submitted on time
CSR link	https://csr.seadatanet.org/report/21000608
Data management	
Embargo	No embargo

License	CC-BY				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
	Turbot AUV Photographic survey				
Number of published CDIs					
Progress follow-up	24/09/2021 - Invitation to submit CSR 09/11/2021 - Invitation to submit CSR 02/12/2021 - Invitation to submit CSR 29/11/2021 – PI submitted CSR to AWI 21/12/2021 – RDC created and loaded CSR on SDN 11/01/2022 - Reminder to send the data 11/01/2022 – PI informed that data will be available in https://soi.squidle.org and on Google Bucket 11/01/2022 - EMODnet Data ingestion Portal was suggest for data submission				

Cruise Name	GSHARK				
Cruise No.	5				
Cruise year	2021				
RDC	RBINS				
Cruise dates	July 31 – August 12				
PI	Bernal Diego				
DMP					
Status	Submitted with some delay. Reviewed by RDC.				
Evaluation	Sufficient				
Review	RDC suggested to give more detailed answers.				
CSR					
Status	Submitted				
CSR link	https://csr.seadatanet.org/report/20213297				
Data management					
Embargo	No embargo				
License	CC0				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical data	CTD				
Biological data	sharks catch				
Physiological data	shark tagging				
	sharks feeding				
	In vitro blood experiment for DNA resiliency to damage				
	Reproductive biology	partially			
	in vivo cardiac responses to temperature				
	whole heart experiments				
	in vitro cardiac mitochondria physiology				

	electroretinograms		
Number of published CDIs	1		
Progress follow-up	<p>Since PI was not responding RDC contacted directly the scientific crew to enquire about data availability, missing data and supplementary information on already transferred data. 9 Scientists have been reached who could provide useful information or additional data. A total of 29 emails have been exchanged in 2023 to get the requested information.</p> <p>As physiological data are 'exotic data' which cannot be ingested into SeaDataNet CDIs service, they will only be published in EMODnet DIP 'as is' as Phase 1.</p> <p>Regarding biological data of sharks catch, mapping variables to SDN standard vocabularies is not straightforward even with the support of EMODnet Biology coordinator since some variables are not mappable to any P01 terms. Publication of the dataset to Phase 2 is therefore not possible. Definitions of the missing variables have been asked to the data-originators, response is pending.</p> <p>Data of sharks feeding and in vitro blood experiment are being processed and are currently missing. Scientists have been reminded to transfer data as soon as available. According to the Cruise Report, there are some data about reproductive biology that have not been transferred. CTD data of one cast has been published in phase 2.</p>		
Planning	Follow-up with missing datasets. The Physiological dataset needs to be updated on EMODnet DIP for phase 1 with additional data the scientists have provided in March 2023. As the data file is more than 1 Gb, DIP master needs to upload the data file manually.		
Difficulties, lessons learned	<ul style="list-style-type: none"> - Amongst the transferred physiological data, there are several datasets which can be only opened with proprietary softwares which does not allow RDC to open and check the submitted data. - No publication to phase 2 possible if no mapping terms for variables is found. This is the case for the biological dataset. - In absence of response from PI, scientific crew at the origin of the data needed to be contacted 		

Cruise Name	MYRTOON
Cruise No.	6
Cruise year	2021
RDC	HCMR
Cruise dates	September 30 – October 09
PI	Dr Andreas Koutsodendris
DMP	
Status	Submitted on time and reviewed by the RDC
Evaluation	Clear and informative full DMP
Review	RDC asked for specifications of the data formats, if they plan to share the analysis software, metadata standards to be used. Full DMP updated with RDC comments.
CSR	
Status	Submitted on time
CSR link	https://csr.seadatanet.org/report/21000531

Data management					
Embargo	CTD are open, 2 years embargo for the rest				
End of embargo	09/10/2023				
License	Public and Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physical oceanographic data	CTD				
Bathymetric data	Multibeam bathymetric			R	
Geophysical	Airgun profilings			R	
	Sparker			R	
	Sediment corers			R	
Number of published CDIs	18				
Progress follow-up	<p>18/01/2022 - HCMR upload to SeaDataNet 2 CDIs for CTD data</p> <p>16/02/2022 - HCMR upload to SeaDataNet 2 CDIs for multibeam</p> <p>16/02/2022 - HCMR upload to SeaDataNet 2 CDIs for airguns</p> <p>23/02/2022 - HCMR upload to SeaDataNet 1 CDI for sparker</p> <p>28/02/2022 - HCMR upload to SeaDataNet 11s CDI for corers</p> <p>30/01/2023 - HCMR asks PI for the complete data of sediment samples analysis</p> <p>31/01/2023 - PI replies that in 2-3 months the publication will be ready and then the final data will become open</p> <p>02/06/2023 - HCMR sends a reminder</p> <p>05/06/2023 - PI replies that the paper is not yet ready but agrees to send the data.</p> <p>07/06/2023 - HCMR gets the complete data set of geochemical and radiochronological data. CDI to be updated. After the publication, data will become open and CDI will be updated again.</p>				

Cruise Name	PHYCOB
Cruise No.	7
Cruise year	2021
RDC	RBINS
Cruise dates	9th - 15th Sept
PI	Bernd Krock
DMP	
Status	Submitted on time. Reviewed by RDC.
Evaluation	Sufficient
Review	DMP almost complete. Some questions require, however, complementary information such as a complete list of collected data types, data QC and data harmonization process, the use of common vocabularies.
Additional remark	In addition to the reviewed DMP and suggested updates, RDC provide to PI additional explanations and an example of full DMP.
CSR	
Status	Submitted with short delay
CSR link	https://csr.seadatanet.org/report/20213265

Data management					
Embargo	CTD and nutrients data: no embargo Biological data: under embargo of 2 years				
End of embargo	15/09/2023				
License	CC0				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical data	CTD				
	Bottle nutrients				
Biological data	Phytoplankton: species determination and quantification				
	Dinoflagellate cyst counts in sediments				
	Phycotoxin measurements				
	Isolates of HAB taxa and their taxonomic and genetic description				
Number of published CDIs	53				
Progress follow-up	CTD and nutrient data were submitted to DIP by PI and pushed to phase 1. For the corresponding CDIs submission, data originators have been contacted to asked for further metadata, quality checks and flags. They provided the required information and CDIs have been published on the SDN CDIs service in May 2023. Biological data were requested first for 3 years of embargo which were downgraded to two years. Lab analyses are almost ready and PI informed in February 2023 they should be provided by end of March 2023.				
Planning	To enquire about availability of biological data				
Difficulties, lessons learned	Quality flags on provided data were missing. RDC contacted directly the scientific crew at the origin of the data collection to ask for quality flags. Direct contact with data-originators is efficient.				

Cruise Name	PORO-CLIM
Cruise No.	8
Cruise year	2021
RDC	HCMR
Cruise dates	May 5-30
PI	Dr Stephen Jones
DMP	
Status	Submitted on time by PI and reviewed by RDC
Evaluation	Complete and detailed full DMP
Review	RDC asked for specifications of the data formats, some more details and references to the QC and harmonization.
CSR	
Status	Submitted on time

CSR link	https://csr.seadatanet.org/report/20213132				
Data management					
Embargo	XBT, Underway data, ADCP are open, 2 years embargo for the rest				
End of embargo	30/05/2023				
License	Public and Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physical oceanographic data Bathymetric/Geophysical data	XBT/raw				
	Underway data (AIS, EUcaws, fluorometer, Gill wind direction, SBE21 thermosalinograph)/raw				
	ADCP/raw				
	Acoustics			R	
	Multi beam bathymetric			R	
	Multi-Channel Seismic reflection (MCS)			R	
	Ocean Bottom Seismometer (OBS)			R	
	Magnetometer			R	
Number of published CDIs	106 under preparation, not yet published, 9 have been published				
Progress follow-up	<p>21/01/2022 - PI started to upload the data (besides the magnetics) on a cloud link</p> <p>15/02/2022 - HCMR completed the download (170 Gb)</p> <p>19/05/2022 - HCMR sent a reminder for the magnetics</p> <p>22/06/2022 - HCMR sent a reminder for the magnetics</p> <p>23/09/2022 - HCMR sent a reminder for the magnetics</p> <p>01/06/2023 - HCMR sent a reminder for the magnetics and the status of the rest (about their process and if they can become open)</p> <p>02/06/2023 - PI replies that would prefer to keep the seismic data (MCS and OBS) restricted until the two PhD students have submitted their papers. PhD students need to be trained first, the process is expected to take 3 to 4 years. It will not happen before 2025, including the training of students and the papers' publications. One PhD started in Oct 2021 and the other in Oct 2022.</p> <p>06/06/2023 - HCMR gets the processed and georeferenced magnetics data. Also HCMR gets a list of conference presentations (EGU, AGU, UK conferences):</p> <p>1) HK Knight, SM Jones, JR Hopper, BM O'Reilly, T Funck (2022), Project PORO-CLIM initial results: A new oceanic crustal record of magma productivity throughout initiation of the North Atlantic Igneous Province at Volcanic and Magmatic Studies Group Annual Meeting 2022, 10-12 January</p> <p>2) HK Knight, SM Jones, JR Hopper, BM O'Reilly, T Funck (2022), Project PORO-CLIM initial results: Towards a new oceanic crustal record of magma productivity throughout initiation of the North Atlantic Igneous Province, EGU22-12380 presented at 2022</p>				

	European Geosciences Union General Assembly, 23-27 May 3) HK Knight, SM Jones, JR Hopper, BM O'Reilly, T Funck (2022), Testing the Relationship Between the North Atlantic Igneous Province Initiation and the Paleocene-Eocene Thermal Maximum, V42F-0125 presented at 2022 American Geophysical Union Fall Meeting, 12–16 December.
--	--

Cruise Name	CALYPSO				
Cruise No.	9				
Cruise year	2022				
RDC	OGS				
Cruise dates	March 16th - March 10th				
PI	Shaun Johnson				
DMP					
Status	Full DMP reviewed				
Evaluation	complete				
CSR					
Status	Submitted on time				
CSR link	https://csr.seadatanet.org/report/20223082				
Data management					
Embargo	2 years				
End of embargo	11/03/2024				
License	CC BY-NC 4.0				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical data	CTD			R	R

	ADCP			R	R
	uCTD			R	R
	Float				R
	Drifter				
	Gliders				
Number of published CDIs	3209				
Progress follow-up	05/04/2022 – Invitation to submit CSR 08/04/2022 – Creation of CSR on behalf of PI 30/05/2022 – Reminder to send the data 06/07/2022 – Reminder to send the data 02/08/2022 – Reminder to send the data 03/08/2022 -Submission data on SEANOE 30/08/2022 - Data sent by mail. Data are imported in SDN as restricted data				

Cruise Name	CABLE leg 1 + leg 2				
Cruise No.	10				
Cruise year	2022				
RDC	OGS				
Cruise dates	22-29th April & 10-13th October				
PI	Taavi Liblik				
DMP					
Status	Full DMP reviewed				
Evaluation	complete				
CSR					
Status	Submitted on time				
CSR link	https://csr.seadatanet.org/report/21025640 https://csr.seadatanet.org/report/21027804				
Data management					
Embargo	2 years				
License	by negotiation				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical data	CTD			R	
Number of published CDIs	93				
Progress follow-up	09/03/2022 – Invitation to Full DMP 05/04/2022 – Invitation to Full DMP 10/04/2022 - Creation of full DMP 16/05/2022 – Invitation to submit CSR 30/05/2022 – Creation of CSR on behalf of PI 06/07/2022 – Reminder to send the data 02/08/2022 – Reminder to send the data 16/01/2023-Reminder to send the data 07/02/2023-Reminder to send the data 27/02/2023-Reminder to send the data 07/03/2023-Reminder to send the data 02/05/2023-Reminder to send the data (Project Coordination Office)				

Difficulties, lessons learned	PI decided to submit data to SDN by himself.
-------------------------------	--

Cruise Name	DOMUSe (CABLE leg 2)				
Cruise No.	10				
Cruise year	2022				
RDC	OGS				
Cruise dates	10-13th October				
PI	Alexandra Loginova				
DMP					
Status	Full DMP reviewed				
Evaluation	complete				
Additional remark	Late, after cruise				
CSR					
Status	Submitted on time				
CSR link	https://csr.seadatanet.org/report/21027804				
Data management					
Embargo	2 years				
End of embargo	29/12/2024				
License	Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANO E
Physico-chemical data	https://doi.org/10.48457/iopan-177a2d0e-489b-42f8-b1d3-4ac17ce7ba19				
Number of published CDIs					
Progress follow-up	09/03/2022 -> mail full DMP 05/04/2022 -> mail full DMP 26/10/2022 -> mail full DMP 08/11/2022 -> mail full DMP 30/11/2022-> mail full DMP+CSR 14/12/2022-> mail full DMP+CSR 15/12/2022-> DMP 28/12/2022-> CSR https://csr.seadatanet.org/report/21027804 16/01/2023-> mail Data 07/02/2023-> mail Data 27/02/2023-> mail Data 07/03/2023-> mail Data 02/05/2023-> mail Data				
Difficulties, lessons learned	PI decided to submit data to IOPAN data infrastructure.				

Cruise Name	GRACE
Cruise No.	11
Cruise year	2022
RDC	HCMR
Cruise dates	28 April - 11 May

PI	Dr Carmen Juan				
DMP					
Status	Submitted on time and reviewed by the RDC				
Evaluation	Complete				
Review	No updates needed				
CSR					
Status	CSR submitted				
CSR link	https://csr.seadatanet.org/report/21025651				
Data management					
Embargo	2 years				
End of embargo	11/5/2024				
License	Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physics	CTD			R	
Bathymetric data	Multi beam echosounders			R	
Geophysical	Seismics			R	
	ROV images			R	
	Sediment cores			R	
Number of published CDIs	1				
Progress follow-up	<p>23/09/2022 - CSIC informed HCMR that a grid of about 100m resolution has been uploaded to EMODnet ingestion (phase 1) but considering that they will be under restricted access.</p> <p>23/09/2022 - HCMR delete the submission as EMODnet Ingestion accepts only fully open data</p> <p>02/06/2023 - HCMR asks PI the bathymetric data</p> <p>08/06/2023 - CSIC is partner in EMODnet Bathymetry and wanted to publish directly the CDI at Bathymetry. After some discussions, CSIC agreed HCMR to publish the multibeam CDI</p> <p>12/06/2023 - HCMR upload to SeaDataNet one CDI for multibeam echosounders</p>				
Planning	HCMR to ask PI the rest of the GRACE geophysical data				

Cruise Name	SEAQUAKE
Cruise No.	11
Cruise year	2022
RDC	HCMR - OGS
Cruise dates	April 28th - May 11th
PI	Sara Martinez Loriente
DMP	
Status	Full DMP to review
Evaluation	complete
CSR	
Status	Submitted on time
CSR link	https://csr.seadatanet.org/report/21025650
Data management	

Embargo	2 years				
License	by negotiation				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Data type 1					
Physico-chemical data (OGS)	CTD			R	
Bathymetric data (HCMR)	Multibeam/raw			R	
	TOPAS (Sub Bottom Profilers)/raw				
	AUV (Side Scan Sonar)/raw				
Number of published CDIs	3 +1				
Progress follow-up	<p>14/12/2021 – Invitation to Full DMP</p> <p>04/01/2022 - Creation of full DMP</p> <p>30/05/2022 – Invitation to submit CSR</p> <p>30/05/2022 – Creation of CSR on behalf of PI</p> <p>01/06/2022 - All raw data (CTD, Multibeam, TOPAS (Sub Bottom Profilers), AUV (Side Scan Sonar), received and transferred by HCMR to google drive</p> <p>20/07/2022- Restricted CTD data received by email</p> <p>29/07/2022 - CTD data are duplicates, PI will check them</p> <p>01/09/2022 - PI sent corrected CTD data. Data will be imported in SDN as restricted data</p> <p>02/06/2023 - HCMR asks the PI for the status of processed bathymetries and the rest geological data.</p> <p>02/06/2023 - PI replied that: a) not all multibeam tracks have been processed, b) AUV data are not processed, they are being processed by VLIZ technicians, and c) TOPAS data is of very poor quality and is not worth spending time processing. They cannot be used as they do not provide useful information.</p> <p>09/06/2023 - after some email exchanges (PI is a partner in EMODnet Bathymetry willing to submit the CDIs directly to Bathymetry), it was agreed between HCMR and the PI, that instead of the RDC the PI prepares and submits the CDI(s) to SeaDataNet. On 09/06, submitted one CDI with the raw bathymetric data as restricted data. The CDI will be updated in the near future with the processed grid.</p>				
Planning	<p>1) Project coordinators contact with VLIZ for the AUV data</p> <p>2) HCMR will monitor the updating of the multibeam CDI(s)</p>				

Cruise Name	SYNERGY
Cruise No.	12
Cruise year	2022
RDC	OGS
Cruise dates	18th - 25th June
PI	James Waggitt
DMP	
Status	Full DMP reviewed
Evaluation	complete

CSR					
Status	Submitted on time				
CSR link	https://csr.seadatanet.org/report/21026618				
Data management					
Embargo	No embargo				
End of embargo					
License	CC0				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical data (OGS)	CTD				
BIOLOGICAL OCEANOGRAPHY	Birds				
MARINE GEOLOGY	Trawl				
MARINE GEOLOGY (HCMR)	Acoustic data (echosounders)/raw				
Number of published CDIs	107				
Progress follow-up	<p>15/07/2020 – Invitation to update DMP</p> <p>22/09/2021 - Updated embargo period on DMP</p> <p>28/07/2022 – Invitation to submit CSR</p> <p>28/07/2022 – Creation of CSR on behalf of PI</p> <p>26/08/2022 - HCMR asks the PI about the acoustic data if they were not submitted to EMODnet Ingestion because of their size (raw data ~ 800Gb) or because they have not processed yet. (DMP mention that data from echo sounders will be retained as .raw files with no processing).</p> <p>26/8/2022 - PI replies that data have not been processed, it is not decided yet if or when they process them.</p> <p>26/08/2022 - HCMR asks for the raw data then</p> <p>29/08/2022- Data available on Ingestion Portal</p> <p>02/06/2022 - HCMR send a reminder for the raw acoustic data</p> <p>03/06/2023 - PI gives HCMR a cloud link (OneDrive) to download the data of total size 800Gb. In addition, PI sends to HCMR an updated trawl file</p> <p>06/06/2023 - HCMR send to OGS the updated trawling spreadsheet</p> <p>14/06/2023 - after several unsuccessful attempts, HCMR finally manage to get the raw acoustic data at a cloud link (also OneDrive)</p>				
Planning	<p>HCMR to transfer the acoustic raw data (800Gb) from OneDrive of Microsoft to the google Drive where the rest of the received raw or processed bathymetric and geological data EF+ data are kept.</p> <p>No further activity is foreseen by the end of EF+</p>				

Cruise Name	CARBO-ACID
Cruise No.	13
Cruise year	2022
RDC	HCMR - OGS
Cruise dates	03-12 August
PI	Dr Emilia Salgueiro

DMP					
Status	Submitted on time by PI and reviewed by RDC				
Evaluation	Complete				
Review	RDC asked PI to change the embargo from 3 to 2 years. PI updated the full DMP				
CSR					
Status	Submitted on time				
CSR link	https://csr.seadatanet.org/report/21026663				
Data management					
Embargo	2 years				
End of embargo	12/8/2024				
License	Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Bathymetric data/geophysical data (HCMR)	Multi beam bathymetry			R	
	Multibeam backscatter			R	
	Sub-bottom profiler			R	
	CTD and water sampling			R	
	Multinet tows			R	
	Sediment data			R	
				R	
Number of published CDIs					
Progress follow-up	<p>26/05/2023 - RBINS asks PI for the data</p> <p>29/05/2023 - PI replies that they finished the reprocessing of the bathymetry and CTD data and have sent it directly to their EMODnet pivot for quality control and for upload of the datasets in the EMODnet portals. The status of this process and the EMODnet DIP references will be sent to RBINS</p> <p>07/06/2023 - RBINS asks for an update on the submission status of the datasets</p> <p>20/06/2023 - No reply received yet</p>				

Cruise Name	TAIPRo2022
Cruise No.	14
Cruise year	2022
RDC	RBINS
Cruise dates	17th – 26th May
PI	Katrin Schroeder
DMP	
Status	Submitted on time, reviewed by RDC, updated by PI
Evaluation	Very complete
Review	DMP very well detailed except for lineage information which is missing. RDC asked for updates of lineage. PI updated it accordingly.
CSR	

Status	Submitted on time				
CSR link	https://csr.seadatanet.org/report/20223094				
Data management					
Embargo	No embargo for all data measured on the ship. 6 months embargo for other physical data. 1 year embargo for all other variables except isotopes and tracers. Isotopes and tracers: 2 years				
End of embargo	26/05/2023 except for isotopes and tracers which should be later				
License	CC BY 4.0				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical sensors data	CTD				
	LADCP				
	VM-ADCP				
	Ferrybox				
	Thermosalinograph				
	Weather				
Chemical measurements	Bottle data from CTD sensors				
	Inorganic nutrients and carbonate parameters (embargo)				
	Radionuclides and Barium (embargo)				
Biological data	Underwater Vision Profiler				
	Environmental DNA				
Number of published CDIs					
Progress follow-up	A first batch of physical data (LADCP, CTD and bottle data) has been submitted to DIP by PI in July 2022. LADCP has been published as phase 1 in October 2022. A second batch of missing physical data (TSG, ferrybox, weather, VM-ADCP) has been enquired by RDC and submitted to DIP by PI in February 2023. Remaining datasets from batch 1 and 2 have been published in phase 1 in May 2023 by RDC. Preparation of the data for CDIs submission is in progress but quality flags were missing and have been requested and sent to RDC in June 2023. VM-ADCP data were submitted in netCDF format and have been asked to be converted in ASCII format for the CDIs submission. PI sent a converted file in May.				
Planning	<ul style="list-style-type: none"> - PI informed that radionuclide and barium data will be ready and quality-checked by the end of 2023. To be followed-up after the end of the project. - Enquire about underwater Vision Profiler and Environmental DNA. - Finalize processing and submission of physico-chemical data to CDIs service. 				
Difficulties, lessons learned	The VM-ADCP data is in a netCDF format which is not exploitable by the CDIs service and need to be converted and provided by PI in ASCII format.				

Cruise Name	IsoMed (RTA during TAIPro2022)				
Cruise No.	14				
Cruise year	2022				
RDC	RBINS				
Cruise dates	17th – 26th May				
PI	sarah.magozzi@szn.it				
DMP					
Status	no DMP required for RTA cruise				
CSR					
Status	CSR published for TAIPro2022				
Data management					
Embargo	no embargo				
License					
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Isotopes	Zooplankton stable carbon (d13C) and nitrogen (d15N) isotope				
Number of published CDIs					
Progress follow-up	IsoMed project has been agreed in June 2022. It has been agreed with the EF+ data managers that no DMP would be required for RTA cruise but an overview of the data (data type, embargo, availability time) would be provided. After many email exchanges, PI replied that isotopes data on zooplankton will be generated, no embargo is foreseen, availability time is not known.				
Planning	As it is not known when the data will be available, no planning can be further foreseen.				
Difficulties, lessons learned	PI informed that since the collected samples only represent a tiny part of a larger collaborative project, the processing of the EF+ samples is not a priority and therefore, it is not known when the data would be available.				

Cruise Name	IOPD
Cruise No.	15
Cruise year	2022
RDC	RBINS
Cruise dates	28th June - 10th July
PI	Wieter Boone
DMP	
Status	Submitted in time, reviewed by RDC.
Evaluation	Very complete
Review	As the DMP is very complete, it is accepted by reviewer as such. No updates required.
CSR	
Status	Submitted with some delay

CSR link	https://csr.seadatanet.org/report/21026654				
Additional remark	Support provided by SDN helpdesk to proceed to the CSR submission				
Data management					
Embargo	2 years				
End of embargo	10/07/2024				
License	CC-BY				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical data	CTD				
	Underway Hydro Acoustic				
	Weather				
	Turbulence profiler				
Biochemical data	Bottles				
Biological data	plankton samples				
	Video plankton recorder + CTD				
Number of published CDIs					
Progress follow-up	Physical and imaging data (15 Gb) have been transferred in October 2022 to RDC who keeps it locally. In February 2023, RDC enquires about missing metadata in the transferred files and the availability time for the other datasets. In March 2023, PI informed that the results/datasets are in different stages of the process: some data is still being processed or quality-checked, while others are basically ready. Hence, PI let know that there were problems with the CPICS during the campaign which prevented it from collecting useful data. They will therefore not be able to supply data for the CPICS.				
Planning	To Process and publish CDIs of already available data. To remind PI to send remaining data as soon as possible.				

Cruise Name	FIGURE - CARING (Co-PI)
Cruise No.	16
Cruise year	2022
RDC	RBINS
Cruise dates	21th – 28th July
PI	Mar Benavides (FIGURE) and Lidia Carracedo (CARING)
DMP	
Status	Both DMPs were submitted in time, reviewed by RDC and updated by PIs
Evaluation	Both very complete
Review	Both DMPs were very well documented despite some missing information on data format, data versions (raw vs processed data), common standards and license. PIs updated the DMP according the RDC reviews.
CSR	
Status	Submitted in time
CSR link	https://csr.seadatanet.org/report/21026645
Data management	

Embargo	2 years on biogeochemical data				
End of embargo	30/07/2024 (or after publication aimed by 2d semester of 2023)				
License	CC0 for FIGURE, CC-BY for CARING				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physical data	CTD and bottle				
	Navigation, weather, underway data				
Biogeochemical data	SADCP				
	N2 fixation (FIGURE)				
	NO3, NH4 and AA uptake measurements (FIGURE)				
	DNA discrete and underway sampling (FIGURE)				
	pCO2, pH, alkalinity (CARING)				
	discrete inorganic nutrients (CARING)				
	underway pCO2 and pH (CARING)				
Number of published CDIs					
Progress follow-up	<p>In October 2022, RDC reminds PIs to provide their data. SADCP data were submitted by PI to EMODnet DIP in October 2022. CTD, bottle and Navigation, weather and underway data were submitted to DIP by PI in January and February 2023. All Physical data are now published as Phase 1 on DIP. A conversion of the SADCP netCDF file to ASCII format has been requested but is pending until the return of the PI from temporary leave.</p> <p>Biogeochemical data were separately sent to RDC as data are under embargo until the results publication.</p> <p>FIGURE: As FIGURE PI needed a DOI for the N2 fixation rate, PI submitted the dataset to SEANOE in March 2023.</p> <p>CARING: In order to process nutrient data for CDIs submission, missing headers and units variables have been asked to CARING PI but status is pending the return of the PI end of June.</p>				
Planning	CDIs preparation for all available datasets				
Difficulties, lessons learned	Conversion of provided netCDF data have been requested by RDC				

Cruise Name	GLICE
Cruise No.	17
Cruise year	2022
RDC	RBINS
Cruise dates	10th -24th August
PI	Mark Hopwood
DMP	
Status	Submitted in time, reviewed by RDC with no updates required

Evaluation	Very complete				
CSR					
Status	Submitted with some delay				
CSR link	https://csr.seadatanet.org/report/21026980				
Data management					
Embargo	1 year for biogeochemical data				
End of embargo	24/08/2023				
License	CC0				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physics data	CTD				
	ADCP				
Biogeochemical data	Drone high resolution imagery				
	underway data				
	bottle data				
	in situ data				
	Iceberg macronutrients				
	Iceberg metals and suspended particles				
Number of published CDIs					
Progress follow-up	In January, PI informed that there is a major issue with the CTD and underway probes that are not yet delivered back to him for post-processing calibration checks. Risk is that post-calibration is not possible and raw data are delivered only. Biogeochemical data are being processed.				
Planning	To reach back PI to enquire about his data submission.				
Difficulties, lessons learned	Major issue with the instruments post operation which can impair the quality of the probes data. Uncertainty on the availability of the data and its quality.				

Cruise Name	SINES
Cruise No.	18
Cruise year	2022
RDC	OGS
Cruise dates	11 - 20 September 2022
PI	Marcos Fontela
DMP	
Status	Full DMP reviewed
Evaluation	complete
CSR	
Status	Submitted on time
CSR link	https://csr.seadatanet.org/report/21027629
Data management	
Embargo	CTDs and EARS data without embargo Other datasets under embargo of 2 years

End of embargo	20/09/2024				
License	CC-BY				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Physico-chemical data	CTD				
	Multinet CTD data				
	AUV				
	Satellite derived data				
	ADCP				
MARINE GEOLOGY	Sediment sampling				
Number of published CDIs	31				
Progress follow-up	DMP was completed. CSR is published. The data are on phase 2.				

Cruise Name	VISIT				
Cruise No.	19				
Cruise year	2023				
RDC	HCMR				
Cruise dates	31 March - 11 April				
PI	Dr Rebecca Bell				
DMP					
Status	Submitted on time and reviewed by the RDC				
Evaluation	Clear and informative full DMP				
Review	Full DMP updated with RDC comments				
CSR					
Status	Under validation by the SeaDataNet/CSR Back Office entry data base. The PI of VISIT cruise has submitted the CSR to the SDN system and currently is implementing changes that requested by the back office during the validation process before the CSR publication.				
CSR link					
Data management					
Embargo	2 years				
End of embargo	11/4/2025				
License	Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Bathymetric data	Multi beam bathymetry			R	
Geophysical	Seismics			R	
Number of published CDIs					
Progress follow-up	02/06/2023 - HCMR asks PI about the CSR submissions 02/06/2023 - PI replies that they are close to complete the Eurofleets+ report. Not clear that they had to fill another CSR, no such reference in the DMP 02/06/2023 - HCMR sends the project guidelines along with additional explanations about the data submission themselves 05/06/2023 - PI started the completion of the SDN/CSR				

	<p>07/06/2023 - HCMR provides an HCMR cloud for uploading of the seismic (raw and processed) data of about 22 Gb.</p> <p>13/06/2023 - PI uploaded the seismic data to the HCMR cloud and informed HCMR of issues with uploading the ship track (navigation files and images).</p> <p>13/06/2023 - HCMR suggested PI to contact the SeaDataNet/CSR Back Office to solve the problem</p> <p>13/06/2023 - SeaDataNet/CSR Back Office replies that the extension of the track image should be in small and not capital letters (.jpg instead of .JPG). CSR system will fix that issue. Concerning the navigation files, CSR accepts geographical coordinates (in degrees) and not projected ones and asked PI to update the file with the coordinates</p>
Planning	HCMR to ask PI to send the bathymetric also.

Cruise Name	Hydee-Obs				
Cruise No.	20				
Cruise year	2023				
RDC	HCMR				
Cruise dates	21 - 31 March				
PI	Dr Joerg Bialas				
DMP					
Status	Submitted on time and reviewed by the RDC				
Evaluation	Clear and detailed full DMP				
Review	Full DMP updated with RDC comments				
CSR					
Status	CSR submitted				
CSR link	https://csr.seadatanet.org/report/21029901				
Data management					
Embargo	2 years				
End of embargo	31/3/2025				
License	Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Geophysical	Ocean Bottom Seismometers (OBS)			R	
	Multi Channel Seismics (MCS)			R	
Number of published CDIs	32				
Progress follow-up	<p>11/03/2023 - HCMR asks PI when the processed data will be received as the embargo extends beyond the end of the project</p> <p>02/06/2023 - PI replies about how the transfer will take place</p> <p>02/06/2023 - HCMR provides an HCMR cloud link</p> <p>02/06/2023 - PI uploads the Ocean Bottom Seismometers (OBS), and the Multi Channel Seismics (MCS)</p> <p>02/06/2023 - PI submits the CSR with the support of HCMR</p> <p>12/06/2023 - HCMR uploads to SeaDataNet 20 CDIs for OBS and 12 for MCS</p>				

Cruise Name	ERODOTO				
Cruise No.	22				
Cruise year	2023				
RDC	OGS				
Cruise dates	1-12 July				
PI	Silvia Ceramicola				
DMP					
Status	not submitted yet				
Evaluation					
Review					
CSR					
Status	Cruise scheduled in July 2023				
CSR link	-				
Data management					
Embargo					
End of embargo					
License					
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Number of published CDIs					
Progress follow-up	Cruise postponed from 19-30 June to 1-12 July.				

Cruise Name	OASIS				
Cruise No.	23				
Cruise year	2023				
RDC	OGS				
Cruise dates	27 March - 7 April				
PI	Claudio Lo Iacono				
DMP					
Status	Full DMP reviewed				
Evaluation	complete				
CSR					
Status	Not submitted yet				
CSR link	-				
Data management					
Embargo	3 years				
End of embargo					
License	restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Number of published CDIs					

Progress follow-up	
Planning	RDC to ask PI to downgrade the embargo from 3 to 2 years

Cruise Name	UNSEEN (Co-PI on OASIS cruise)				
Cruise No.	24				
Cruise year	2023				
RDC	RBINS				
Cruise dates	26 March - 7 April 2023				
PI	Martina Pierdomenico				
DMP					
Status	Submitted in time, reviewed by RDC and updated by PI				
Evaluation	Complete				
Review	DMP complete and well described. Missing information on data harmonization and foreseen data processing have been updated by PI.				
CSR					
Status	OASIS CSR - submission pending				
CSR link	-				
Additional remark					
Data management					
Embargo	2 years				
End of embargo	27/04/2025				
License	CC-BY 4.0				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Microplastics	Microplastics in sediment				
	Microplastics in water				
	Microplastics in biota				
Number of published CDIs					
Progress follow-up	Embargo of 3 years requested first and downgraded to 2 years upon request of RDC. In February 2023, PI informed that considering the time required to process the samples and that the cruise has been postponed from summer 2021 to spring 2023 she still does not know whether she will be able to submit all processed data by autumn 2023. She said, surely all the metadata associated with the samples and the available results will be sent before the end of the project. She will keep the RDC updated on her timeline.				
Planning	To revert to PI to ask about any updates on data availability.				
Difficulties, lessons learned	Long time required for the analyses the samples				

Cruise Name	Poseidon
Cruise No.	25
Cruise year	2023

RDC	HCMR				
Cruise dates	12 - 22 June				
PI	Dr César Rodríguez Ranero				
DMP					
Status	Full DMP not submitted yet				
Evaluation					
Review	HCMR contacted PI for the full DMP late				
CSR					
Status					
CSR link					
Data management					
Embargo	3 years				
End of embargo	22/3/2026				
License	Restricted				
Datasets ingestion status:		Transferred	DIP1	CDIs	SEANOE
Bathymetry	Multibeam bathymetry			R	
Geophysical	Ocean Bottom Seismometers (OBS)			R	
	Sediment cores			R	
Number of published CDIs					
Progress follow-up	07/06/2023 - HCMR asks PI for the full DMP late. Also reminds for the CSR in 2 weeks after the cruise and processed data by the end of the project. 07/06/2023 - PI replies that we discuss the points after the cruise				

6. Ingestion progress

As shown in Table 1, the majority of the cruises data providers has started to transfer their collected datasets. One third of the cruises has already transfer all their data. More than one third of the other cruises has already transfer partially their data. Pending datasets are for most due to lab analysis or data processing. They are for most data under embargo. Contacts have been made with PIs for the other missing or incomplete datasets to request their submission. All received datasets which are not under embargo have been or are to be submitted and published on EMODnet DIP at Phase 1. The data curation process has started and is in progress for most of the cruises which have provided data. Their Phase 2 are now in progress. This phase is heavier as the generation of CDIs requires significant work from the RDCs. Table 3 shows the number of CDIs already ingested in the SDN CDIs service. It includes both open and restricted data. 3810 CDIs have already been generated over 14 cruises, meaning that those data are harmonized and interoperable. These CDIs are mostly composed of physico-chemical data (mainly CTD and currents data) and geophysical data (bathymetric and seismic data). This is already a good progress given the significant work the data ingestion process represents and the encountered delays in receiving the data.

ID	Cruise Name	Geophysical CDIs	Geological CDIs	Physico- chemical CDIs	TOTAL CDIs	CDIs in preparation
3	Focus-AUV	53			53	
8	PORO-CLIM	9			9	106
1	iMAR	40		46	86	
5	GSHARK			1	1	
2	BENCHMARK	32		81	113	
7	PHYCOB			53	53	
6	MYRTOON	5	11	2	18	
9	CALYPSO			3209	3209	
10	CABLE			93	93	
11	GRACE	1			1	
11	SEAQUAKE (Co-PI on GRACE cruise)	1		3	4	
12	SYNERGY			107	107	
18	SINES			31	31	
20	Hydee-Obs	32			32	
TOTAL NUMBER OF CDIs		173	11	3626	3810	106

Table 3: Number of ingested CDIs in Phase 2

The ingestion progress is also reported cruise by cruise on the EVIOR portal in the EUROFLEETS+ Cruise Datasets Catalogue (<https://evior.eurofleets.eu/cds/search>).

7. Difficulties and lessons learned

Communication with PIs

Most of the PIs are very collaborative and endeavour to submit all their data at the due time as well as providing any required information to their RDC. However, some PIs are less and sometimes totally non responsive. RDCs need to send many reminders to get ultimately some answers and to be sent some datasets but sometimes there is no answer. The reasons can be due to; some misunderstanding of the EF+ cruise data management process, overloaded PI, missing coordination and little communication between PI and cruise scientists, PI absence for professional or personal reasons or reluctance to follow the prescribed procedures. For the first, RDCs took the time to explain the different DM steps and data submission pathways, sometimes via virtual meetings and gave additional support to pursue the data transfer and submission. When a PI is not responding after several reminders, EF+ coordinator is asked to intervene to ask for a better collaboration from the PI. Most of the time, the dialogue could start again based on a more collaborative approach. In some cases, it is easier to reach directly the data originator (crew scientists) to get complementary information on the submitted data rather than keeping the PI as intermediate contact. In one case, as the PI did not respond anymore, RDC had to find and contact directly other scientists of the cruise to seek for information. In conclusion, the communication with the data providers may take time, with many emails exchanges and waiting time to get answers. It is not always clear what is the process status of the datasets. The follow-up of the data submissions needs to be actively conducted by the RDC with regular emails sent to remind the data provider to make their data available and to keep RDC updated with their data processing timeline.

Transfer of complete sets of metadata and data

Based on both DMP and Cruise Report, the RDCs are checking the completeness of the submitted/transferred data and metadata. They contact PIs to enquire about missing datasets and make the follow-up of the data availability. Often, metadata information on the time reference, device model, data processing (lineage) and quality flags are incomplete or missing. Checking for completeness and gathering the data and related information takes time and, in few cases, does not end up with the requested information.

Where it is recommended to provide processed and quality flagged data rather than raw data, in some cases, PIs only provide raw data which may result in unnecessary large files, in particular for bathymetric data. Yet, it implies time and resources for the RDC to manage such raw data. It should be clearer to the PIs that the Project needs processed instead of non-processed data.

Checking for completeness and consistency of the transferred datasets are sometimes not possible because the data are produced and/or processed by proprietary software. Also, in some cases, the datatypes are so specific (e.g. organism physiology) that the RDC does not have the expertise to assess the dataset. In those cases, RDCs rely on the expertise of the data-provider and ensure the data is attested to be complete and sufficiently processed and documented for their publication.

In some cases, PIs do not provide the collected data even after many attempts and reminders from their assigned RDCs. Instead, they make the choice to dispense with RDC curation and to submit their CDIs themselves or to publish their data on their own institutional infrastructures.

Regarding the Cruise Report which includes useful information on the measurements, analyses and processing procedure, it has been suggested by some scientists to append it to the data submission on EMODnet DIP and SDN CDIs service next to the metadata/data as it documents well the sampling, measuring, processing procedures as well the scientific objectives, context and preliminary results of the project. This needs to be assessed with respect to the GDPR guidelines since the names of the cruise crew is mentioned.

Data curation

When transferred, the RDC performed a 1st-line quality check (spatial and temporal coverage, range checks). However, it does not replace the data provider knowledge in his/her field of expertise to assess the consistency and quality of the data. Quality flags need to be provided by the data providers but often, they are not and need to be requested by RDCs.

In order to make the data interoperable and to push data to Phase 2, data need to be harmonized. Data and metadata are formatted following SDN common ingestion format. Measured variables and sampling/analysis devices need to be mapped to common vocabularies. In one case (GSHARK), biological variables could not be mapped to existing PO1 vocabulary (i.e. NERC vocabulary for parameters) and therefore could not be pushed to Phase 2. Another issue was when PI submitted his/her data in a netCDF format else than the SDN one. The extraction of the relevant parameters and linked meta-information requires a good knowledge of the studied variables as well as the technical tools to perform the conversion to odv file. Therefore, when data are provided in netCDF, RDC asks the data provider to also provide a converted ASCII data file with the relevant variables. Regarding the data harmonization, in several cases depending on the data types, the RDCs rely on SDN softwares (NEMO and MIKADO) to format the data files according the SDN CDIs requirements. Before the last upgrade of NEMO, some ADCP data files could not be handled properly by NEMO. The new version of NEMO released in May 2023 should now allow such handling. MIKADO has also been upgraded in May 2023 which will certainly ease the metadata preparation process.

As the data curation process in order to make data interoperable via the SDN CDI service is a time and effort consuming task, it will not be possible to handle the data and to push them to Phase 2 in the case it takes too long for the data submitters to provide the RDCs with their dataset near or after the end of the EUROFLEETS+ project. In those cases, RDC can only ensure the data is published on EMODnet DIP or on SDN SEANOE for the datasets under embargo.

Timing: late data delivery because of lab work and other issues

PIs are expected to submit or transfer their data within 2 months after the cruise or as soon as the data becomes available (i.e. lab results). Where a large proportion of the cruise data have been transferred, there are still missing datasets that are expected to be provided. Most of those delayed datasets are due to lab work which require a significant amount of time and for which it seems often difficult to foresee how long it would take to provide resulting data. Longer delay is even expected for genomic or isotopic data acquisition. RDCs keep following-up the data availability with the PIs but it is often at the own initiative of the RDC that updates are given. Also, some data needs further processing that also impaired the data submission in due time (e.g. imagery analyses, bathymetry processing). In one case (GLICE), there was some major issues delivering back the probes to the PI's institute which impaired the post-calibration of the data. It results from these situations that RDCs need to deploy more effort in contacting the PIs to eventually get the data or to be informed of the submission progress.

Lessons learned

Overall, the cruise data management of the 28 projects are progressing well with almost all DMPs and CSRs submitted. Regarding the data handling, despite some difficulties already mentioned, the communication with most of the PIs is going well and data submissions is on track even though it can take some time. A close follow-up with the scientists will be performed until the end of the project with a special attention to communication and good sharing of information aspects.

8. Future data management strategy

The RDCs are and will continue the follow-up of the data submissions in close contact with the PIs and they will continue to progress as and when the datasets are supplied in processing the datasets for publication to Phase 1 and 2. As the project comes to an end, the challenge will be to process most of the remaining datasets by the end of the project. Therefore, the priority would go to the gathering and preservation of as many datasets as possible given the late availability of the data. For that reason, PIs have been contacted to fasten their samples and data analyses and to make their data available as soon as possible. Hopefully, the lab analysis should be ready before the end of the project. As mentioned, it may eventually be necessary, when close to the end of the project to leave the submission in Phase 2 and to focus on the data publication to DIP and SDN SEANOE which require less effort to preserve the collected data. It would ensure datasets with both open and restricted access to be Findable and eventually Accessible. As many datasets embargo extend beyond the end of the project, RDC have started to ask the data providers whether their embargo can expire earlier. If not, embargoed data will need to be followed-up beyond the end of the project to change the restricted access licenses to an open one. In any case, funded cruises should transfer to the project RDCs all their data by the end of the project.