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	31/01/2022
Deliverable No	4.9
Lead Beneficiary	RBINS
Dissemination Level	Public

D4.9: Data management follow-up report for funded cruises

Document information	
Document Name	D4.9: Data management follow-up report for funded cruises
Document ID	D4.9 Data management follow-up report for funded cruises v1.4 RBINS.docx
Revision	1.4
Revision Date	2021-01-27
Author	Iona Sissy, Le Hong Minh, Partescano Elena, Vandenberghe Thomas
Security	Public

Approvals			
	Name	**Organisation**	**Date**
Coordinator	Aodhán Fitzgerald	Marine Institute	2022-01-31
Activity Coordinator	Anneli Strobel	AWI	2022-01-24
WP Leader	Nicole Biebow	AWI	2022-01-24

History			
Revision	**Date**	**Modification**	**Author**
v1.0	2021-02-03	First creation	Vandenberghe Thomas
V1.1	2021-02-17	Correction: only REGIONAL got reminders	Vandenberghe Thomas
V1.2	2022-01-19	Updated to postponed Deliverable date	Le Hong Minh, Vandenberghe Thomas
V1.3	2022-01-24	Merged input	Iona Sissy, Le Hong Minh, Partescano Elena, Vandenberghe Thomas
V1.4	2022-01-27	Merged coordinator's comments into the document; added license where missing; summarized difficulties and lessons learned into one section.	Iona Sissy, Le Hong Minh, Partescano Elena, Vandenberghe Thomas

TABLE OF CONTENTS

Contents

1. COVID-19 situation and deliverable postponement

This deliverable had an original submission date of 31/07/2020, which was postponed following a COVID-19 impact assessment: on 30/03/2020 and a revised date of 29/01/2021 was proposed. Unfortunately, in January 2021 only one SEA Programme campaign had been undertaken due to the severe disruption to research vessel operations schedules throughout 2020 and into 2021. A draft deliverable was submitted in January 2021 reporting on the FOCUS-AUV campaign in New Zealand (deployment of UGOT Hugin AUV by R/V Tangaroa from Sept 30th 2020 - Oct 29th 2020). This update deliverable incorporates all implemented TNA since the beginning of the project.

2. Campaign allocation and timelines

Campaigns have been allocated evenly among the Reference Data Centres (RDCs: BMDC, HCMR and OGS) taking into account regional and thematic affinity. Campaigns involving seismics or multibeam are assigned to HCMR. In some cases, different datasets from the same campaign have been assigned to different data centres. In addition, a data centre has been assigned to review the DMP.

The data management guidelines (https://raw.githubusercontent.com/tvandenberghe/roadmap/ef-dmp/public/files/EF%2B_data_management_guidelines_PI.pdf) and the data handling chain state the following timelines:

1. Follow the webinar on FAIR, DMPs and Data management via https://www.eurofleets.eu/2019/07/26/eurofleets-data-management-plan-dmp-webinar-how-to-prepare-a-dmp/

2. Prepare your proposal

3. Create an account on the EF+ DMP Roadmap (http://dmp.ef-ears.eu) and complete the preliminary DMP

4. Submit your proposal on the official Eurofleets+ submission website, together with a pdf of the DMP generated from DMP Roadmap)

(1-4 are not relevant in this context)

5. Await confirmation that you are awarded the project

6. Receive additional feedback from the reference data centres on improving the DMP

7. Two months at the latest before the campaign, complete the final DMP on http://dmp.ef-ears.eu

8. Participate in the campaign

9. Within 2 weeks after the campaign, provide the SeaDataNet campaign Summary Report

10. Within 2 months after the campaign, provide the campaign Report to the EUROFLEETS+ Evaluation Office at AWI

11. Perform the necessary QC on your raw data so that it is ready for publication in a data repository.

12. The data submitter (the person responsible for the data) gathers all the raw data collected during the campaign.

13. Within 2 months after the campaign, the data submitter provides the data files and metadata of the dataset in the EMODnet Data ingestion Portal.

14. Depending on the sea area the campaign takes place in, a data centre has been assigned to your dataset(s). A data manager will contact you if there are any questions about making your data interoperable.

The campaigns have been assigned to the RDCs as follows. Please note the 2023 DOMUSe Co-PI campaign still needs to be assigned to a data centre.

Year	R/V	Campaign	Topic	Location	Date	DMP DC	Data DC
2020	UGOT Hugin AUV	Focus-AUV	Geology, Marine Biology, Sedimentology	Kaikōura Canyon, New Zealand	Sept 30th - Oct 29th	HCMR	HCMR
2021	Celtic Explorer	PORO-CLIM	Climate dynamics, Geophysics, Sedimentology, Training	NE Atlantic (S Rockall Plateau)	05-30th May	HCMR	HCMR
2021	Pelagia	iMAR	Biological Oceanography, Geology, New technologies	Mid-Atlantic Ridge near the Azores	May 18th - June 3rd	OGS	HCMR (dataset 1) OGS (datasets 2 and 3)
2021	DANA	GSHARK	Marine Biology	Bredefjord, Greenland	July 30 - 12 August	RBINS	RBINS
2021	GO Sars_ROV Aegir	BENCHMARK	Marine Biology Physical Oceanography	Denmark Strait	01-10 August	RBINS	HCMR (dataset 1) RBINS (datasets 2 and 3)
2021	Tubitak Marmara	PHYCOB	Biological Oceanography	Western Black Sea	15th - 22nd Sept	RBINS	RBINS
2021	Aegeo	MYRTOON	Climate dynamics	SW Aegean Sea, Myrtoon Basin	29 Sept - 08 Oct	HCMR	HCMR

Year	R/V	Campaign	Topic	Location	Date	DMP DC	Data DC
2021	SOCIB	GRASSMAP	New technologies, Marine Biology	Mallorca and Cabrera islands	14th - 20th September	OGS	OGS
2022	Pelagia	CALYPSO	Physical Oceanography	Alboran or Balearic Sea	16 Feb - Mar 10	OGS	OGS
2022	Aranda	CABLE	Biological Oceanography	Gulf of Finland, Baltic Proper	22-29 April & October	OGS	OGS
2022	Aranda	DOMUSe (Co-PI on CABLE)	Biogeochemistry, biological oceanography	Gotland Deep Baltic Sea	22-29 April & October	*DC TBC*	*DC TBC*
2022	Belgica II_AUV Barabas	GRACE	Geology Geophysics Physical Oceanography Sedimentology	Ceuta Canyon and adjacent areas	May 2022	HCMR	HCMR
2022	Belgica II_AUV Barabas	SEAQUAKE (Co-PI on GRACE campaign)	Geology, Geophysics, New Technologies, Sedimentology	Ceuta Canyon and adjacent areas	May 2022	OGS	HCMR (dataset 1) OGS (dataset 2)
2022	Mare Nigrum	SLOGARO-II	Geology, Geophysics, Sedimentology, Training	Danube Fan	10-17 May	HCMR	HCMR
2022	Belgica II	TAlPro2022	Physical Oceanography	Algero-Provencal Basin, Sicily Channel	17-26th May 2022	RBINS	RBINS
2022	Arni Friedrikson	SENERGY	Marine Biology	North Western Iceland	18-25 June	OGS	HCMR (dataset 1) OGS (datasets 2 and 3)
2022	Sanna	IOPD	Biogeochemistry Climate dynamics New technologies	Godhabfjord, Ameralik fjord	28th June - 10th July	RBINS	RBINS

Year	R/V	Campaign	Topic	Location	Date	DMP DC	Data DC
			Marine Biology Polar Biology Training				
2022	Atlantic Explorer	FIGURE	Microplastics in Cold Water Coral reef	Gulf Stream (NW Atlantic)	21-28th July	RBINS	RBINS
2022	Ramon Margalef	CARBO-ACID	Biogeochemistry, Marine Chemistry, Physical Oceanography	Iberian margin	3-12th August	HCMR	HCMR (dataset 1) RBINS (datasets 2 and 3)
2022	Sanna	GLICE	Biological Oceanography Biogeochemistry Physical Oceanography	Disco Bay, West Greenland coastline	10-24 Aug	RBINS	RBINS
2022	Sarmiento de Gamboa	SINES	Biogeochemistry	Northeast Atlantic, Western Iberian Margin	11-20th September 2022	OGS	OGS
2023	Thalassa_ROV Ariane	OASIS	Geology New technologies Marine Biology Physical Oceanography	SE Alboran Sea		OGS	OGS
2023	Thalassa_ROV Ariane	UNSEEN	Microplastics in Cold Water Coral reef	SE Alboran Sea		RBINS	RBINS
2023	Tangaroa	VISIT	Geophysics	East coast of North Island, New Zealand	Spring 2023	HCMR	HCMR (dataset 1) OGS (dataset 2)
2023	Laura Bassi	ICON	Climate dynamics Deep Sea Research	Hillary Canyon, Ross Sea, Antarctica	Spring 2023	HCMR	HCMR

Year	R/V	Campaign	Topic	Location	Date	DMP DC	Data DC
			Geology Geophysics				
2023	Laura Bassi	IDEA (Co-PI on ICON campaign)	Geophysics	Hillary Canyon, Ross Sea, Antarctica	Spring 2023	HCMR	HCMR
2023	Tangaroa	Hydee-Obs	Geophysics	Hikurangi Margin, North Island, New Zealand	Spring 2023	HCMR	HCMR
2023	Laura Bassi	ROSEBUD	Climate Dynamics, Deep Sea Research, Geology, Physical Oceanography, Sea Ice, Sedimentology	Ross Sea, Antarctic	Spring 2023	RBINS	HCMR dataset 1 RBINS dataset 2
2023	RV Aegaeo_ROV Max Rover_AUV Barabas	ERODOTO	Deep Sea Research, Geology, Geophysics	Squillace Canyon, Italy	Summer 2023	OGS	HCMR dataset 1 OGS dataset 2

3. DMP completion and follow-up

Principal Investigators (PIs) complete a preliminary phase DMP during proposal submission and append a full DMP to it in the DMP Online tool (http://dmp.ef-ears.eu). The awarded campaigns of the first OCEAN call were reminded to complete the full DMP before June 2020. With the re-scheduling of their campaigns to the next years due to the COVID pandemic, three DMPs completions and reviews (Calypso, GLICE and TalPro2022) were requested to be put on hold by their PIs; TalPro2022 has been completed in the meantime. They have been or will be asked for completion and reviews in due time according to the EF+ Provisional TA schedule.

At the moment, 9 out of 29 do not have the full DMP completed. The earliest one of these (CABLE campaign of 2 April) should have a completed DMP by mid-February.

Review of the other Ocean, Regional and Co-PI call DMPs were completed and were all of high quality. Reoccurring remarks were: missing references to standard vocabularies (the PIs were encouraged to include these mappings in the data files and update the DMP accordingly); issues with embargoes; and limited lineage information.

4. Issues with embargoes

For the FOCUS-AUV (3), FIGURE (partial 4), SENERGY (3) and OASIS (3) campaigns, embargo periods beyond the allowed 2 years had been requested. The <u>data policy</u> is very clear on this. The Principal Investigators have been asked to confirm/change this information on 10th August 2020 for the first campaign and on June 10th, 2021 for the other ones. The DMPs have been updated accordingly by the SENERGY and FIGURE principal investigators (*resp 3 years to none and 4 years to "after publication"*). PIs with 2-year embargo dates are requested to also provide the data, but by email and not via the EMODnet Data Ingestion Portal (as this is only for fully open data). The OASIS campaign has been postponed to Spring 2023. Further confirmation on the availability of the data is still needed. The embargo for the FIGURE campaign is due to the fact that the -omics data cannot be derived and published from the samples within 4 years. The PI agreed to release all other data within the required timelines (physics, biogeochemical). Note that the SeaDataNet system does not explicitly recognize -omics data as a data type; in order to make the data available, a link can be made from the CSR to an open data repository (i.e. the EMBL-EBI database).

For the PHYCOB campaign, all data derived from the live cell cultures would be released after publication; when asked to specify this, the PI answered this would take 3 years. The RDC has asked the PI to consider shortening the embargo down to 2 years, to which he agreed on January 28th, 2022.

If the Principal Investigator asked for an embargo, the RDCs will still perform the necessary data management and disseminate the data to SeaDataNet within the Eurofleets+ project. The data will receive the "By negotiation" license, displaying its existence but making it available only when the PI would agree. An important consideration is that post-project these datasets would need to have their license changed to CC-BY.

5. Standard procedures and documents

The RDCs use a **standardized follow-up template** to review the DMPs and any related DM issues: https://marineinst.sharepoint.com/sites/EurofleetsPlus/_layouts/15/Doc.aspx?sourcedoc=%7BA1BF 13BF-985D-47AA-8B7D-3D2AD86F622F%7D&file=TNA_campaign_dc_assignnment_and_DMP_reviews.xlsx&action=default& mobileredirect=true&CT=1642415658995&OR=ItemsView.

The follow-up template is an extensive checklist and among others also includes the deadlines for the metadata and data after the campaign (i.e. CSR, dataset submission by Principal Investigator), and to keep track when the final data submission to SeaDataNet is done by the RDC. There is no set period to perform the final submission, but it should be done as soon as possible.

Quality Control guidelines for the RDCs for the handling of the datasets are enriched on an ad hoc basis in this document: https://marineinst.sharepoint.com/:w:/r/sites/EurofleetsPlus/_layouts/15/Doc.aspx?sourcedoc=%7 B5122E7A9-9A92-4C20-94B1-EA786235386C%7D&file=QC_DM_procedures.docx&action=default&mobileredirect=true. As the RDC encounter common issues with regard to metadata (CSRs, EMODnet DIP metadata) and data (enough data meta-info per data value) they are noted down in this document. If someone encounters an issue, he informs the RDC colleagues with the problem and solution and (possibly) after discussion the guideline is amended.

The **guidelines for PIs** (https://raw.githubusercontent.com/tvandenberghe/roadmap/ef-dmp/public/files/EF%2B_data_management_guidelines_PI.pdf) have been clarified that the full DMP should be completed and reviewed by the Reference Data Centre within 2 months before the campaign leaves. It has also been clarified that the DMP RDC is responsible to issue submission reminders to the PI (metadata and data). It has been made clear that there can be more than one submission in the EMODnet DIP if not all data is ready within the first 2 months after the campaign. As it is possible that additional lab work is required on samples and this may take time. In that case the datasets in the DIP should clearly indicate this ("Campaign name" Part 1: ...)

Finally, to coordinate the installation of the EARS server by CSIC, to remind the scientists of sending their device list (in order to prepopulate the EARS event logger) and to send training invitations, another document is used for the RDCs and WP3 partners.

6. Case-by-case overview

The BENCHMARK, Focus-AUV, GRASSMAP, GSHARK, iMAR, MYRTOON, PHYCOB and PORO-CLIM campaigns (8 in total) have been completed. The FIGURE and TalPro22 campaigns are included as well to give an overview concerning the DMP review.

The Principal Investigators have received a training (webinar) regarding the EARS software and the metadata and data submission guidelines. The attendants were informed about the general deadlines concerning CSR and data submission. Webinars were given on:

- 23/03/2020 (invited: BENCHMARK, GSHARK, iMAR, OASIS, PORO-CLIM, MYRTOON, SENERGY, SLOGARO-II).

- 23/07/2020 (attendance by GSHARK)

- 28/07/2020 (attendance by SENERGY)

- 07/09/2021 (attendance by GRASSMAP, MYRTOON and PHYCOB). The attendants were informed about the specific deadlines concerning CSR and data submission.

Campaign Name	**BENCHMARK**
Campaign year	2021
RDC	RBINS
Campaign dates	August 1 – August 10
PI	Burgos Julian
DMP follow-up	
Full DMP completed on time	Yes on time
Summary DMP review and progress follow-up	Overall full DMP completed on time and well detailed. RDC suggests to use more community standards and vocabularies and to improve lineage information.
Status	Full DMP needs to be updated with RDC comments
Logs	14/05/2020 – First invitation to complete full DMP. 12/06/2020 – Full DMP completed

	17/07/2020 – RDC sent DMP review to PI, some updates are requested
Data management follow-up	
Embargo	None except for observations from videos where 2 years embargo is applied and will be released after scientific publication.
License	CC0
Progress follow-up	CSR submitted on time. Awaiting full datasets.
Logs	24/08/2021 – Invitation to submit CSR and invitation to send data for October 10 25/08/2021 – PI informed CSR is submitted 17/01/2022 - Reminder to submit data to DIP 27/01/2022 - Reminder to submit data to DIP
Summary of data curation progress	
Status Metadata submission	RBINS
CSR on BSH	yes
CSR on SDN	yes
Status data handling	RBINS
EMODnet DIP status	No data submitted yet (expected on 10th October 2021)
Submission covers all datasets	
Dataset 1	RBINS
Data types	ROV dives: Video footage and photographs, preliminary video annotation files Hydrography: CTD profile, measurements of nutrients, etc. from water samples, acoustic current profiler.
EMODnet DIP Phase I	
EMODnet DIP Phase II	
Total number of CDIs	
Dataset 2	HCMR
Data types	Multibeam echosounder: raw bathymetric, water column, sound velocity and absorption profiles.
EMODnet DIP Phase I	
EMODnet DIP Phase II	
Total number of CDIs	

Campaign Name	**FIGURE**
Campaign year	2022
RDC	RBINS
Campaign dates	21th – 28th July
PI	Mar Benavides
DMP follow-up	

Full DMP completed on time	Yes (on time)
Summary DMP review and progress follow-up	DMP very well documented despite information on common standards and license which should be added.
Status	DMP reviewed and needs to be updated by PI
Logs	05/07/2021 - full DMP completed by PI 18/01/2022 - full DMP review and sent to PI for updates
Data management follow-up	
Embargo	Physics data: none Biochemical data: embargo until publication
License	'open'
Progress follow-up	There are genomic data types which will be made available on the EMBL-EBI database. The SDN infrastructure is not equipped for such data yet; a viable option is to link in the CSR to the EMBL-EBI database (which has openly available urls for datasets). The PI requested first an embargo period of 4 years for which the EF+ project coordinator asked a reduction to 2 years. This 4 year embargo request has been withdrawn (18/01/2022) since the researcher requesting the 4 years embargo will not participate in the campaign anymore. RDC asked whether the other datasets will be under embargo and whether genomics data will still be collected. The PI confirmed that DNA will be gathered, but not for -omics research (it will still allow quantification of the original research question). The lab to dissemination duration is set at 6 months for this type of research.
Logs	June 2021 - EF+ Project Coordinator wrote to PI to ask to improve embargo time. Sept 2021 – RDC exchanges mails with Niamh Flavin regarding embargo improvement and how to link genomics repository from within our systems. 18/01/2022 - RDC asked whether PI has considered to revise the embargo period. 18/01/2022 - PI informed embargo request of 4 years intended for one specific researcher using genomics data will be withdrawn because the researcher will not participate to the campaign eventually.
Data types	Physics data: ADCP/LADCP, turbulence, CTD, underway hydrography and weather data. Biogeochemical data: nitrogen fixation rate, flow cytometry plankton counts, nutrients, diazotroph abundance, amplicon sequencing and metatranscriptomics.

Campaign Name	FOCUS-AUV
Campaign year	2020

RDC	HCMR
Campaign dates	September 30 – October 28
PI	Dr Joshu Mountjoy
DMP follow-up	
Full DMP completed on time	Yes, on time
Summary DMP review and progress follow-up	Clear and informative DMP. 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.
Status	PI still has to include the RDC suggestions in the DMP
Logs	12/06/2020 – Full DMP completed by PI 10/08/2020 – RDC review was sent to PI 09/10/2020 – Full DMP updated by PI
Data management follow-up	
Embargo	2 years (it was 3 years before the DMP review)
License	Creative Commons
Progress follow-up	CSR has been created and submitted by PI using the BSH CMS. However, some information remains outstanding (such as ship track or some of the measured parameters). RDC has not yet asked the PI to update the CSR. If needed, the RDC will assist on this. campaign report submitted to AWI 2 months after the campaign. Data was sent to RDC
Logs	14/01/2021 - Invitation was sent to PI to submit CSR plus the processed data 27/01/2021 – CSR was published on line (SeaDataNet interface). PI informed the RDC that they had some issues with the navigation from one AUV dive so the processing is not finished yet. 26/05/2021 – Question to PI on the progress of data processing, reminder to send part of the finalized data to start their integration into the EF+/SDN repository. 22/09/2021 – Reminder to send the data, request for the navigation files. 25/09/2021 – Question for existence of underway data 30/09/2021 – PI sent the navigation files. No underway data for this campaign. 25/11/2021 - Reminder to send the data 06/12/2021 - Reminder to send the data 06/12/2021 – PI replied that he is still waiting on the final metadata attribution from their technician and that in one week should be able to send the first data. 13/01/2022 - Reminder to send the data 18/01/2022 – First campaign data sent to RDC (Multibeam bathymetric and Backscatter grids)

	20/01/2022 – PI started the uploading of the rest of the campaign data (Sub-bottom profilers, Sidescan Sonar). Large data set (~700GB total). 21/01/2022 – PI sent the link to RDC. RDC started the data download.	
Summary of data curation progress	RDC stores the data on the cloud and portable hard disks. RDC has started the preparation of the CDIs for the multibeam bathymetric data. The data will be labeled as restricted until their embargo finish.	
Status Metadata submission		NIWA
CSR on BSH	Yes	
CSR on SDN	Yes	
Status Data handling		HCMR
Data types	Multibeam bathymetric grids, Sidescan Sonar files, Seafloor backscatter grids, subbottom profiler files	
EMODnet DIP status	No submission through DIP (there is embargo of 2 years).	
Submission covers all datasets	Acoustic water column backscatter data are not specified in the DMP. Data screening does not show anything of interest and there is no intension to do any processing. These data will be made available on request.	
EMODnet DIP Phase I		
EMODnet DIP Phase II		
Total number of CDIs		

Campaign Name	**GRASSMAP**
Campaign year	2021
RDC	OGS
Campaign dates	September 14 – September 20
PI	Miguel Massot Campos
DMP follow-up	
Full DMP completed on time	Yes
Summary DMP review and progress follow-up	
Status	Full DMP to review
Logs	
Data management follow-up	
Embargo	One month after the campaign.
License	CC-BY
Progress follow-up	Cruise Report has been created and submitted by PI to AWI. The RDC had to create and load CSR on SDN.
Logs	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	
Summary of data curation progress		
Status Metadata submission		OGS
CSR on BSH	no	
CSR on SDN	yes	
Status data handling		OGS
Data types		
EMODnet DIP status	No data submitted yet	
Submission covers all datasets		
EMODnet DIP Phase I		
EMODnet DIP Phase II		
Total number of CDIs		

Campaign Name	**GSHARK**
Campaign year	2021
RDC	RBINS
Campaign dates	July 31 – August 12
PI	Bernal Diego
DMP follow-up	
Full DMP completed on time	Yes but late
Summary DMP review and progress follow-up	DMP completed short before the campaign, DMP review not possible before campaign. The DMP review was delayed because of absence of the DMP reviewer. DMP review: Brief DMP that needs more details. It seems that practical details of data management not fully foreseen or too simplified.
Status	1st DMP review sent (late, after the campaign) Awaiting PI updates.
Logs	14/05/2020 – First invitation to complete full DMP. 16/07/2020, 31/07/2020 – Reminder to complete full DMP. 21/07/2021 – Full DMP completed 13/01/2022 – Full DMP reviewed and sent to PI for update
Data management follow-up	
Embargo	none
License	CC0
Progress follow-up	CSR has been created and submitted by RDC. Data has been sent by PI to RDC but are not digitally re-usable (scans of papers). RDC will submit the data to EMODnet DIP on behalf of PI when digital version of data received. Completeness of the dataset is under question.

Logs	24/08/2021, 27/09/2021, 03/11/2021 and 24/11/2021 – Reminders to submit CSR 24/11/2021 – campaign report received 30/11/2021 – creation of CSR on behalf of PI 02/12/2021 – RDC required clarifications about CSR and remind PI for data submission. RDC offers to submit data on behalf of PI. 09/12/2021 – RDC requested to receive the data 14/12/2021: Data received: 12 print screen of navigation panel, 1 CTD profile, 1 pdf file of pictures of manual notes of 16 pages of a notebook, 1 pdf with pictures of 11 pages of measures results done on sharks + PI notified that EARS was not set up during the campaign, so no underway data acquisition data, data logged in excel format instead. CTD malfunctioning, only one vertical profile collected. 13/01/2022 – RDC asked PI for digital versions of the data 17/01/2022 – RDC informed Eurofleets+ Coordinator of the delay. 28/01/2022 – Formal request for submission of data sent by coordinator 28/01/2022 – PI agrees to send data as quickly as possible
Summary of data curation progress	

Status Metadata submission		RBINS
CSR on BSH	Yes	
CSR on SDN	No	

Status Data handling		RBINS
Data types	CTD, physiological data	
EMODnet DIP status	Unusable data received (late of 2 months)	
Submission covers all datasets	doubtful	
EMODnet DIP Phase I		
EMODnet DIP Phase II		
Total number of CDIs		

Campaign Name	**iMAR**
Campaign year	2021
RDC	OGS
Campaign dates	May 18th - June 3rd
PI	Telmo Morato
DMP follow-up	
Full DMP completed on time	Yes
Summary DMP review and progress follow-up	
Status	Full DMP to review
Logs	
Data management follow-up	

Embargo	2 years
License	Creative Commons
Progress follow-up	CSR has been created and submitted by PI using BSH CMS.
Logs	16/06/2021 – Invitation to submit CSR
	17/06/2021 – Creation of CSR on behalf of PI
	22/09/2021 – Reminder to send the data
	23/09/2021 – PI sent the list of acquired data
	09/11/2021 – Reminder to send the data
	24/11/2021 – Reminder to send the data
	01/12/2021 – Reminder to send the data
	26/01/2022 – Reminder to send the data
	26/01/2022 – PI answered mail asking to use Pangea as data sharing
	27/01/2022 – Data management WG confirmed to use EMODnet/SeaDataNet to share data
	28/01/2022 – After confirmation to use EMODnet Ingestion for data sharing the PI asked more information/clarification about the process
	28/01/2022 – RDC proposed a virtual meeting

Summary of data curation progress		
Status Metadata submission		OGS
CSR on BSH	yes	
CSR on SDN	yes	
Status Data handling		OGS
Data types		
EMODnet DIP status	No data submitted yet	
Submission covers all datasets		
EMODnet DIP Phase I		
EMODnet DIP Phase II		
Total number of CDIs		

Campaign Name	**MYRTOON**
Campaign year	2021
RDC	HCMR
Campaign dates	September 30 – October 09
PI	Dr Andreas Koutsodendris
DMP follow-up	
Full DMP completed on time	Yes, on time
Summary DMP review and progress follow-up	Clear and informative full DMP. RDC asked for specifications of the data formats, if they plan to share the analysis software, metadata standards to be used.
Status	Full DMP updated with RDC comments
Logs	28/02/2020 - Full DMP completed by PI
	24/03/2021 – RDC review was sent to PI

	01/04/2021 – PI updated the DMP
Data management follow-up	
Embargo	CTD are open, 2 years embargo for the rest of the campaign data
License	Creative Commons
Progress follow-up	- CSR has been created and submitted by RDC to SDN. - CTD data published to SeaDataNet - Marine geophysical data under process at HCMR
Logs	25/11/2021 – Invitation to submit CSR and submit the open data (CTD) to EMODnet Data Ingestion 02/12/2021 – PI reports issues to retrieve the correct EARS events export 02/12/2021 – RDC request assistance by RBINS to retrieve the correct CSR from the EARS server 06/12/2021 – EARS manager replied 07/12/2021 – PI submits CTD data to EMODnet Data Ingestion 14/12/2021 – CSR was published on line (SeaDataNet interface) 18/01/2022 – CTD data published to SeaDataNet
Summary of data curation progress	CDIs for multibeam data are under preparation. Multibeam data have been processed. Subbottom profiling data are under process at HCMR. Sediment data are under laboratory analysis.

Status Metadata submission		HCMR
CSR on BSH	No	
CSR on SDN	Yes	

Status data handling		HCMR
Data types	Multibeam, Airgun, Sparker, box corer, CTD data	
EMODnet DIP status	CTD submitted	
Submission covers all datasets		
EMODnet DIP Phase I	30/12/2021	
EMODnet DIP Phase II	19/01/2022	
Total number of CDIs		

Campaign Name	**PHYCOB**
Campaign year	2021
RDC	RBINS
Campaign dates	11th - 17th Sept
PI	PI: Bernd Krock PhD: Kristof Möller (in charge of data submission during 3 months of PI absence (from December 2021)
DMP follow-up	
Full DMP completed on time	Yes on time
Summary DMP review and progress follow-up	DMP almost complete. Some questions remain, however, mainly on embargo, complementary information or details such as a complete list of collected data types, data QC and data harmonization process,

This project has received funding from the EU
H2020 research and innovation programme
under Grant Agreement No 824077

15

	and the use of common vocabularies. The preliminary DMP reported that data could be made public after publication; yet the full DMP estimated this to be 3 years. Being asked for clarification, the PI will send the data to the RDC as soon as it is available and will put all data on EMODnet DIP when the embargo ends. PI agrees with 2 year embargo.
Status	DMP was reviewed after the campaign due to absence of the DMP reviewer.
Logs	03/02/2021 – PI asked support regarding DMP Roadmap connection– Support provided by RDC – PI asked support for DMP question – Support provided by RDC. Full DMP completed 18/01/2022 – RDC sent reviewed DMP to PI 24/01/2022 – PI replied 27/01/2022 – RDC asked for clarification on embargo issue 27/01/2022 – PI replied 29/01/2022 – RDC confirms answer. Embargo is reduced to 2 years
Data management follow-up	
Embargo	CTD data directly open access Other data types open up after publication (2 years).
License	CC0
Progress follow-up	With RDC support, PI promptly proceeded to submissions. CSR submitted and first part of the data (CTD and bottle data) submitted to DIP by PI. Second part of data including 4 different data types. Two of them will hopefully be made available in the first quarter of 2022. Availability of the third one unknown. For the fourth, it concerns taxonomic and genetic description. For genomics, link to PANGAEA from within our systems (likely CSR) to be foreseen.
Logs	03/11/2021 – Invitation to complete CSR and to submit data 24/11/2021 – RDC enquires about CSR submission status and data submission schedule. PIs refers to a colleague regarding CSR submission and says he didn't know he had to submit data to DIP portal. RDC remind him his duty and the procedure. 02/12/2021 – PI asked support regarding License significations – RDC answered – PI asks how different data types collected onboard should be gathered for DIP submission – RDC answered they can all fit in one dataset. RDC enquires when lab measurement will be available. They will constitute a second submission. RDC advised PI to name properly the two separate datasets submission. PI detailed the remaining data types and availability dates. 9/12/2021 – RDC asked PI to proceed to data submission. PI submits data part 1. 18/01/2022 - RDC asked in the DMP review whether PI can consider improving the embargo down to 2 years. 19/01/2022 – RDC asked clarifications to move submission to Phase I 24/01/2022 – PI replied with clarifications. He has asked instrument names to RV staff; response is pending
Summary of data curation progress	

Status Metadata submission		RBINS
CSR on BSH	Yes (almost on time 24/11/2021)	
CSR on SDN	No	
Status Data handling		RBINS
Data types	Dataset 1: CTD related data i.e. sensors of the CTD and pH, oxygen and nutrient data derived from bottle water samples that were determined on board. Dataset 2: Biological data measured in lab: cell counts, dinoflagellate cyst counts in sediment, phycotoxin measurements, isolates of HAB taxa and their taxonomic and genetic description.	
EMODnet DIP status	PI submitted data part 1	
Submission covers all datasets	Dataset 1 submitted	
EMODnet DIP Phase I		
EMODnet DIP Phase II		
Total number of CDIs		

Campaign Name	**PORO-CLIM**	
Campaign year	2021	
RDC	HCMR	
Campaign dates	May 5-30	
PI	Dr Stephen Jones	
DMP follow-up		
Full DMP completed on time	Yes, on time	
Summary DMP review and progress follow-up	Complete and detailed full DMP. RDC asked for specifications of the data formats, some more details and references to the QC and harmonization.	
Status	Full DMP needed to be updated with RDC comments	
Logs	09/06/2020 - Full DMP completed by PI 09/08/2020 – RDC review was sent to PI 09/02/2021 – Reminder sent to PI to finalize the updates 10/09/2021 – PI reply that hopefully the updates will be finished by the end of next week.	
Data management follow-up		
Embargo	2 years	
License	CC0	
Progress follow-up	CSR has been created and submitted by PI using BSH CMS. Data was sent to RDC.	
Logs	15/06/2021 – CSR was published on line (SeaDataNet interface) 24/09/2021 – Question if CTD, XBT, ADCP are still under restriction, otherwise to send them 25/09/2021 – Question for existence of underway data	

	06/12/2021 – Reminder for underway data, request for the navigation files, clarifications requested for some types of collected data and their availability 16/12/2021 – PI gave an update of the seismic, magnetic campaign data archive preparation. Data archives will be send to the RDC and the sponsoring company. There are underway data in this campaign. Question addressed to the RDC about the duplicate data repositories 12/01/2022 – Reply to the PI, reminder to send the cleaned and QC'ed so far data and not wait until the complete data archive is ready. 12/01/2022 – PI replied that will keep the RDC posted over the next fortnight on what is happening. 21/01/2022: PI assembled the entire campaign data package and started the uploading to the cloud (>170 Gb). The uploading has finished for: ADCP, Echosounders, Multibeam bathymetry, part of multi-channel seismic data, Ocean Bottom Seismometer and Ship's Computer System data. RDC has started the downloading of the above data. Queued for uploading are the remainder of the MCS data, Magnetometer, and XBT data.
Summary of data curation progress	RDC stores the data on the cloud and portable hard disks. Further handling will start soon.

Status Metadata submission		Univ. of Birmingham
CSR on BSH	Yes	
CSR on SDN	Yes	
Status data handling		HCMR
Data types	Ocean Bottom Seismometer, Multi-channel seismic data, Magnetometer data, Multibeam, Echosounders, ADCP	
EMODnet DIP status	No submission through DIP yet (there is embargo of 2 years).	
Submission covers all datasets		
EMODnet DIP Phase I		
EMODnet DIP Phase II		
Total number of CDIs		

Campaign Name	**TAIPro2022**
Campaign year	2022
RDC	RBINS
Campaign dates	17th – 26th May
PI	Katrin Schroeder
DMP follow-up	
Full DMP completed on time	Yes (on time)
Summary DMP review and progress follow-up	Overall full DMP completed on time and very well detailed except for lineage documentation which seems to be missing.
Status	Full DMP reviewed, waiting for updates by PI

Logs	14/05/2020 - Invitation to complete DMP in full 08/06/2020 - Request to put on hold DMP completion because change of schedule 18/03/2021 - full DMP completed by PI 18/01/2022 - full DMP reviewed and sent to PI for updates
Data management follow-up	
Embargo	Some early, others at 1y. - Preliminary data set released within 6 weeks (e.g. all data measured on the ship) - 6 months for final physical data - 1 year for final data of all other variables (except for isotopes or tracers)
License	CC BY 4.0
Data types	CTD, biochemical data, chemical data, particles data, zooplankton data

Overview of difficulties

Of the 8 completed campaigns, all have submitted a CSR. Some have been entered in the BSH system and do not yet appear in the SeaDatanet system. For three, no data has been submitted yet, notwithstanding regular reminders.

In general, the following difficulties have been encountered:

- Some difficulties setting up communication, especially in the beginning of the pandemic.

- Embargoes:

  ◦ Overhead of asking for a reduction in the embargo duration to max 2 years, which in all cases succeeded.

  ◦ Small overhead in double submissions (one public, the other by mail) when an embargo is present.

- Making CSR available: Delays of CSR submission of up to 4 months, with regular reminders. In some cases, this was due to confusing the difference between the Cruise Report (sent to AWI, due after 2 months), and the Cruise Summary Report (due after 2 weeks).

- Making data available:

  ◦ In most cases with short delays beyond the 2 months to in a few cases very long delays (from cruise to this moment, up to 8 months), with regular reminders. In one case, data was made available but not in a usable format.

  ◦ Technical issues during the campaign that led to longer data finalization times.

  ◦ Data standardization taking longer than expected due to complex data.

Lessons learned:

- PIs have to be contacted often and assisted as best as possible to ensure a timely submission of the CSR. If the PI submitted, a modification of the CSR is in some cases needed.

- For the actual data submission assistance is less critical as the dataset metadata is reviewed by the RDC anyway as part of the review process. At this stage the PI is contacted. For the actual submission, frequent reminders (both of future deadlines and any delays afterwards) remain required.