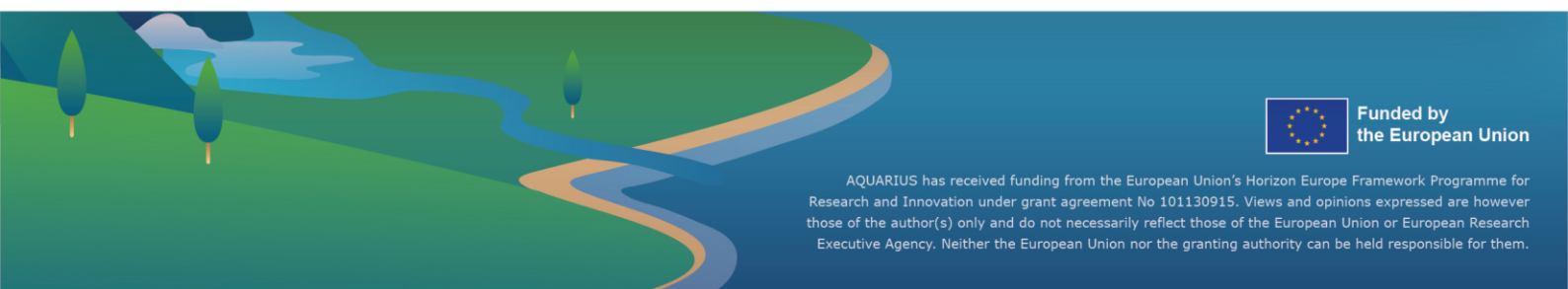




Multifaceted Impacts of Eddy Activity on the East Greenland Current System E(ddies)GC

AQUARIUS initial Data Management Plan template for TA Call applications

September 2024



AQUARIUS has received funding from the European Union's Horizon Europe Framework Programme for Research and Innovation under grant agreement No 101130915. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

AQUARIUS initial Data Management Plan template for TA Call applications

As part of the Transnational Access (TA) projects in AQUARIUS, many new data sets in a large variety of data types will be collected by the TA scientific teams, making use and combining multiple and different observation installations as provided. There will be a strong effort in AQUARIUS to get the maximum return of investment from the TA activities towards serving the EU Mission and AQUARIUS Partnership targets and associated initiatives and projects with new data, data products, and scientific knowledge. Therefore, AQUARIUS has adopted an **open data policy**, which will be implemented with a **dedicated Data Management approach**, to ensure that all gathered and generated metadata and data will be managed in line with the FAIR principles (Findable, Accessible, Interoperable, Reusable). The metadata and data should become part of the archives managed and operated by leading European data management infrastructures for quality assurance, long term stewardship, and wide access and use.

TA scientific teams are advised to take note of the AQUARIUS Data Management approach document and guideline which can be found at:
<https://zenodo.org/records/13642869>

The AQUARIUS data management flow scheme includes a number of steps from planning to training to deployment to publishing, and a number of instruments, which should be applied in those steps by the TA scientific teams.

An initial step is that TA Call applicants are requested to complete an **initial Data Management Plan** for their proposed project, following the **AQUARIUS DMP template phase 1**, and to submit this as part of their TA Call application

Once the TA proposal has been awarded, the TA project team will be requested to review and refine the DMP once the TA team has a better understanding of the TA project that will be deployed and the AQUARIUS data management approach. AQUARIUS expert data centres can assist with the Call process and respond to questions about the AQUARIUS data management approach, for as far as answers cannot be found in the AQUARIUS Data Management approach document. In the phase after the TA project awarding there will be a kick-off meeting for selected TA projects with the AQUARIUS expert data centres to give support for completing and refining the TA project DMP template.

The questions in the DMP template are derived from the Horizon Europe DMP template and adapted for use as part of the AQUARIUS data management approach.

AQUARIUS initial Data Management Template - to be completed and submitted as part of the TA Call application

1. Data Summary

What is the purpose of the data collection/generation and its relation to the objectives of the project?

The purpose of data collection is to investigate the multifaceted impacts of eddy activity on the East Greenland Current (EGC) system, specifically focusing on how eddies influence physical, chemical, and biological processes. This contributes to the objectives of understanding the variability in water column structure, nutrient cycling, dissolved organic matter (DOM) pools, and phytoplankton biomass within the EGC. This research aligns with the broader project objectives of advancing our understanding of the EGC's role in regional and global climate systems, improving predictions of oceanic changes, and supporting sustainable management of Arctic marine ecosystems.

What types and formats of data will the project generate/collect?

The project will generate and collect the following types and formats of data:

- **Ocean Currents:**
 - **Data:** Ocean velocities from surface to 400 m depth (via VM-ADCP), enhanced vertical resolution profiles at specific stations (via L-ADCP), and high-resolution vertical profiles from surface to seafloor spanning eddy radius (via glider-mounted ADCPs).
 - **Format:** NetCDF.
- **Water Column Structure:**
 - **Data:** Temperature, salinity, pressure, depth, nitrate concentrations and fluorescence measurements (CDOM and Chl-a).
 - **Format:** NetCDF, CSV.
- **High-Resolution Eddy Sampling:**
 - **Data:** Vertical profiles of CTD, turbidity, dissolved oxygen and ADCP data collected by gliders.
 - **Format:** NetCDF.
- **Dissolved Nutrients:**
 - **Data:** Concentrations of dissolved nitrate, phosphate and silicate.
 - **Format:** CSV, Excel files.
- **Dissolved Organic Matter (DOM):**
 - **Data:** Bulk dissolved organic carbon (DOC), coloured dissolved organic matter (CDOM) absorption and fluorescent dissolved organic matter (FDOM) characterization.
 - **Format:** CSV, spectral data files.
- **Phytoplankton Biomass:**
 - **Data:** Concentrations of chlorophyll-a (Chl-a), used as a proxy for phytoplankton biomass.
 - **Format:** CSV.
- **Greenhouse Gases (GHG):**
 - **Data:** Concentrations of dissolved CO₂ and CH₄ in surface waters.
 - **Format:** CSV.
- **Carbonate System Parameters:**
 - **Data:** Dissolved Inorganic Carbon (DIC), alkalinity, pH and concentrations of dissolved CH₄ and N₂O.
 - **Format:** CSV.
- **Microplastics:**
 - **Data:** Microplastic concentrations and classification.
 - **Format:** ODV, following guidelines from EMODnet Chemistry (<https://nodc.ogs.it/catalogs/doidetails;jsessionid=04B7C13EFF>)

A974DC4CAACA0D046F4EC4?0&doi=10.6092/d3e239ec-f790-4ee4-9bb4-c32ef39b426d)

All collected data will be accompanied by detailed metadata to ensure proper documentation. Metadata will follow international standards such as INSPIRE, CF Conventions (<https://cfconventions.org/cf-conventions/cf-conventions.html>) and ISO and will include mappings to controlled vocabularies such as the British Oceanographic Data Centre (BODC) Identifiers (https://vocab.nerc.ac.uk/search_nvs/) and EMODnet processing levels (https://emodnet.ec.europa.eu/sites/emodnet.ec.europa.eu/files/public/Brochure/Data%20Portfolio/EMODnet%20Portfolio_16012023.pdf?t=1673878841), ensuring accessibility, traceability and adherence to global best practices such as those published in <https://www.oceanbestpractices.org/>.

Data Category	Description	Collection Method	Format
Ocean Currents	Ocean velocity profiles from surface to 400 m and near the seafloor.	VM-ADCP (continuous), L-ADCP (enhanced resolution), glider-mounted ADCP (vertical profiles).	NetCDF
Water Column Structure	Temperature, salinity, pressure, depth, nitrate concentrations, and fluorescence data.	CTD system with SUNA (nitrate), CDOM fluorometer, and Chl-a fluorometer.	NetCDF, CSV
High-Resolution Sampling	Vertical profiles of CTD, turbidity, dissolved oxygen, and ADCP data spanning eddy radii.	Gliders Yoko and Tsuno with CTD, turbidity, JFE Rinko AROD-FT (oxygen), Nortek AD2CP (ADCP).	NetCDF
Dissolved Nutrients	Concentrations of nitrate, phosphate, and silicate.	Samples from Niskin bottles, preserved frozen for later analysis with a nutrient analyser at DTU Aqua.	CSV, Excel
Dissolved Organic Matter	Bulk DOC, CDOM absorption, and FDOM characterization.	Filtered water samples (0.2 µm), stored in acid-washed amber vials, analysed at DTU Aqua.	CSV, spectral data files
Phytoplankton Biomass	Concentrations of photosynthetic pigment chlorophyll-a (Chl-a) as a proxy for biomass.	Samples filtered onto 0.7 µm glass fiber filters, shock-frozen, stored at -80°C, analysed with fluorometry.	CSV
Greenhouse Gases (GHG)	Dissolved CO ₂ and CH ₄ concentrations in surface waters.	Continuous surface water sampling with CRDS system (Picarro G2201-i).	CSV
Carbonate System	DIC, alkalinity, pH, dissolved CH ₄ , and N ₂ O.	Water samples from surface supply and Niskin bottles, analysed at Flanders Marine Institute and Liège.	CSV
Microplastics	Collection of microplastics/microlitter in multiple size	FerryBox device with sieves connected to the vessel's seawater supply;	Laboratory analysis results;

	fractions (>50 µm, >100 µm, >300 µm), with filtering capability down to 20 µm.	digital flow meter records volume and flow rate; samples processed in a laboratory.	Flow rate data logs.
Bathymetry	Recording of underwater topography using vessel-based sensors.	Multi-beam, sub-bottom profiling, echo sounder, sonar sensors on vessel.	Bathymetric maps; Data delivered to EMODnet.

How is the original data gathered on the TA installation(s) and how do you plan to transfer it to you?

The original data for the project will be gathered using a combination of ship-based systems, glider-based systems, and laboratory analyses conducted during the research campaign. Ship-based systems include vessel-mounted ADCPs (VM-ADCP) for continuous data collection and L-ADCPs mounted on the CTD rosette to obtain enhanced vertical resolution at specific stations. CTD measurements will be carried out to characterize the water column structure and collect water samples, using the ship's CTD system equipped with additional sensors such as SUNA (nitrate), CDOM, and Chl-a fluorometers. Surface water for greenhouse gas and carbonate system analyses will be obtained from the ship's seawater supply system.

Microplastics data will be collected using a FerryBox device installed on the vessel, which captures microlitter in various size fractions (e.g., >50 µm, >100 µm, >300 µm) with filtration down to 20 µm. A digital flow meter records water volume and flow rate. The collected samples will be stored and processed at the VLIZ installation in Oostende, while digital data will be transferred via onboard storage and secure data channels. Bathymetry data will be gathered using the vessel's multi-beam echo sounders, sub-bottom profilers, and sonar systems, which continuously collect depth and seafloor profile data. Processed data will be delivered to EMODnet in formats such as GeoTIFF and NetCDF via secure online transfers.

Glider-based systems, including the Alseamar SeaExplorer gliders (Yoko and Tsuno), will be employed to gather high-resolution vertical profiles of physical (CTD), chemical (dissolved oxygen, turbidity), and biological (Chl-a fluorescence) data. These gliders are also equipped with ADCPs to collect detailed measurements of ocean velocities.

Water samples for laboratory analysis will be collected from Niskin bottles attached to the CTD rosette. These samples will be preserved and transported to specialized laboratories such as DTU Aqua, Flanders Marine Institute (VLIZ), and the University of Liège, where detailed analyses of dissolved nutrients, organic matter, phytoplankton biomass, greenhouse gases, and carbonate system parameters will be conducted.

The data transfer process includes daily backups of digital data onboard the vessel, downloading glider data upon recovery, and ensuring completeness. After the campaign, raw data files in formats such as NetCDF and CSV will be securely transferred to a central server at the lead institution via external storage devices or encrypted online systems. Results from laboratory analyses will be shared electronically by participating laboratories using secure data-sharing platforms.

What processing on the raw data do you plan? Please differentiate between data quality assurance (handling of outliers, missing and suspect values, null observations) and data harmonisation (code, label and QC flag explanations, consistent use of headers and units, data formatting, conforming to standards).

The raw data collected during the project will undergo processing to ensure both quality and harmonization.

For data quality assurance, to ensure the reliability and accuracy of the data, we will implement robust quality control measures. Outliers will be identified through both automated and manual checks and will be flagged or removed based on predefined

thresholds and statistical analysis. Missing or suspect values will be addressed by interpolation methods where appropriate or flagged if no reliable interpolation is feasible. Null observations will be systematically managed and flagged using quality codes to distinguish between instrument errors, physical constraints, or other causes of data gaps.

The data harmonization process will ensure consistency and compliance with established standards. Standardized codes and quality control flags, such as those from SeaDataNet, will be applied to facilitate understanding and interoperability. Consistent headers and SI units will be used across all datasets, ensuring uniformity in measurements. Raw data will be converted into standardized formats such as NetCDF or CSV to ensure compatibility with global databases and analysis tools. Metadata and data will adhere to international standards, such as BODC Identifiers and EMODnet processing levels, to ensure proper documentation and integration with existing repositories.

In addition to these processes, data interpolation will be applied to fill gaps, using spatial or temporal interpolation techniques when appropriate, while ensuring physical and biological plausibility. By combining these quality assurance and harmonization efforts, we aim to create a reliable, well-documented, and standardized dataset suitable for analysis and dissemination.

Based on the training to be provided by Aquarius in March 2028, and for selected datasets, we will consider using the Blue-Cloud services, including its analytical services and federated repositories (in situ and satellite data, as well as derived data products). We will also explore the use of the Virtual Labs within the Blue-Cloud VRE to analyse newly acquired data and combine them with other datasets to generate new products and insights relevant to the Mission and the Partnership.

When do you plan to perform these processing steps by means of a concise planning?

The processing steps will be carried out in the following timeline: Initial quality assurance, including the detection and removal of outliers, will be performed within six months of data collection. Harmonization steps, such as applying standard codes, flags, and formatting, will begin after the end of the cruise deployment and be completed within twelve months post-collection. Advanced processing tasks like interpolation and data integration for specific analyses will be finalized within fourteen months after collection to ensure the dataset is ready for modelling and interpretation. After two years the data can be made publicly available.

What is the expected size (Megabyte to Terabyte range) of the data? Please note that in case of multibeam or imagery data, the derived products (such as grids) should be submitted instead of the raw measurements.

To be collected datasets are in the Terabyte range.

Multibeam data collection will be performed if there is available system in the research vessel, provided that such a collection will not interfere with other available acoustic systems, such as vessel mounted ADCPs, lowered ADCP, and echosounder. No further data processing within the scope of E(ddies)GC is expected in case multibeam data is collected. Therefore, raw data will be available.

To whom might your data be useful, outside your project?

The datasets to be collected are highly specialized, and their primary users are expected to be researchers working in the field. We will actively communicate and promote both the datasets and scientific findings to the wider community, encouraging their reuse.

Simplified versions of the datasets, translated into accessible metrics and indicators, will also be made available for policymakers, the media, and the general public.

2. Making data FAIR

2.1 Making data findable, including provisions for metadata

The **AQUARIUS Data Management approach** is aimed at inclusion of your collected/generated data in leading European aquatic data management infrastructures such as **SeaDataNet**¹ (physics, chemistry, geophysics, geology, and biology), **EuroBIS**² (marine biodiversity), **ELIXIR-ENA**³ (biogenomics), **ICOS-Ocean**⁴ (carbon), and **Copernicus INSTAC**⁵ (Near- Real-Time data). Inclusion will be mediated through expert data centres which are partners in the AQUARIUS project. They will give training and education to the TA teams about planned data management practices as part of AQUARIUS WP5. In addition, as part of AQUARIUS WP6, they will give support to the TA teams for the adoption of the preferred metadata and data standards, and ingestion of the resulting metadata and data in repositories that are nodes of the indicated European data management infrastructures. This way, the metadata and data from the TA projects will become FAIR and also available for EMODnet, Copernicus Marine, DTO developments, EOSC, and global data exchanges such as GEOSS and the digital ocean ecosystem which is being developed in the framework of the UN-IOC-Ocean Decade programme. In the AQUARIUS Data Management approach, it is recommended to make use of the online **Data Submission service of EMODnet Data Ingestion**⁶ and/or **SeaDataNet SeaNoe service**⁷ for transfer of your metadata and data packages to the AQUARIUS expert data centres.

Paragraph 5.4 of the **AQUARIUS Data Management approach** document lists the minimum set of metadata that EMODnet Ingestion requests you to enter in order to document your data packages submission. In addition, Annex 1 of the document gives a list of items that should be documented during the TA data acquisition and processing activities as these are later needed, when transforming data to the standard metadata and data formats as required.

What naming conventions for your data files will you follow?

Data file names will follow the convention:

CAMPAIGN_YYYYMMDD_SENSOR/PARAMETER_STATION_NUMBER

where:

¹<https://www.seadatanet.org>

²<https://www.eurobis.org>

³<https://www.ebi.ac.uk/ena/>

⁴<https://www.icos-cp.eu/observations/ocean/otc>

⁵<https://marine.copernicus.eu/about/producers/insitu-tac>

⁶<https://www.emodnet-ingestion.eu>

⁷<https://www.seadatanet.org/Software/SEANOE>

- **CAMPAIGN** is EddiesGC.
- **YYYYMMDD** is the sampling date, formatted as year (four digits), month (two digits), and day (two digits).
- **SENSOR/PARAMETER** identifies the instrument or measured parameter (e.g., CTD).
- **STATION_NUMBER** is the sequential CTD station number, beginning with 01.

CTD stations will serve as the reference for the **STATION_NUMBER**, as they represent the most comprehensive set of discrete measurements. Instruments sampling at CTD stations but not necessarily deployed at every station (e.g., the spectrophotometer) will nevertheless use the corresponding CTD station number to maintain consistency across datasets.

For underway systems (e.g., VM-ADCP), filenames will retain the **CAMPAIGN** prefix and use the UTC timestamp generated by the data acquisition system in place of the station number.

Will you in principle be able to follow the AQUARIUS Data Management approach for metadata? What possible challenges do you foresee?

We will be able to follow the AQUARIUS Data Management approach for metadata as outlined in the guidelines. The required metadata elements, such as title, abstract, keywords, file formats, provenance, lineage, processing methodology, dataset documentation, and spatial and temporal extents, align with our planned metadata documentation practices. We are prepared to provide detailed records on data acquisition, quality assurance, harmonization processes, and software used, ensuring compliance with EMODnet submission requirements. Additionally, our datasets will include appropriate coordinate reference systems, bounding boxes, and taxonomic coverage, as specified.

A potential challenge may arise from ensuring the metadata and data formats align seamlessly with the standards and vocabularies of the European data management infrastructure. To address this, we plan to maintain detailed documentation throughout the project, enabling smooth transformation of data and metadata to conform to these standards during the submission phase.

How will you document all this information before submission, especially lineage information (i.e. processing and QC steps)?

We will document all relevant information, including lineage, processing, and QC steps, by maintaining comprehensive records throughout the project. This will include detailed logs for data acquisition methods, sensor specifications, and any calibration procedures. Each data processing step will be documented, including any interpolation, outlier removal, and quality control actions taken, as well as the tools and software used. Additionally, we will keep track of the metadata associated with each dataset, ensuring that all transformations, formatting adjustments, and harmonization steps are thoroughly explained. This documentation will be compiled and organized in a format compatible with the EMODnet standards, enabling a smooth submission process. By maintaining this level of detail, we ensure that the lineage and quality assurance processes are transparent and traceable.

2.2. Making data openly accessible

The AQUARIUS project follows an **open data access policy**. Therefore, the **AQUARIUS Data Management approach** is aimed at inclusion and publishing of your collected/generated data as open metadata and data in trusted repositories, within the **AQUARIUS | Initial DMP template**

lifetime of the AQUARIUS project. However, temporary embargos (max 2 years after collection) are possible for specific data that require additional processing and to allow for scientific analyses that will lead to scientific papers and PhD theses.

Will you collect/generate data that might require a temporary embargo? Indicate which type of data and motivate its temporary embargo for each case.

The data collected during the project will be subject to a temporary embargo period of two years. This embargo is necessary to allow adequate time for quality control, processing, and analysis, as well as to ensure that the research team can publish findings derived from the data. After this period, all data will be made publicly available in accordance with open data access policies. This approach ensures both the integrity of the research process and the eventual accessibility of high-quality, well-documented data for the broader scientific community.

Do you also plan to make the data and metadata available on another repository next to the AQUARIUS repositories, for instance an institutional, national or general data repository?

In addition to submitting the data to the AQUARIUS repositories, we intend to make the data and metadata available through other established platforms for wider accessibility and long-term storage. Specifically, we plan to archive the datasets on the Marine Data Archive (MDA) hosted by VLIZ (<https://mda.vliz.be/>), EMODNet Chemistry (<https://emodnet.ec.europa.eu/en/marine-litter>), or other equivalent repositories. This approach will help ensure the data is easily accessible to a broad scientific community while adhering to high standards of data management and preservation.

In case of use of exotic data formats for original data, what methods or software tools are needed to read and understand the data and could you provide documentation about that (including possible open-source code)?

Not applicable.

2.3. Making data interoperable

The **AQUARIUS Data Management approach** is aimed at making the collected/generated data interoperable by adopting common standards from leading European data management infrastructures. Use will be made of standard formats, QA-QC procedures, and controlled vocabularies for many metadata tags. For this process the TA teams will be given support by the AQUARIUS expert data centres, who also can help with use of common tools and services. To be able to do these transformations from original metadata and data to standardised metadata and data, it is important that relevant information about the data collection/generation is noted down during the actual events.

This consists of the **AQUARIUS TA Data Summary Log** which should be maintained by the PI of the TA project scientific team to keep an overview and index of the data collection events. It should be shared afterwards with the AQUARIUS data centre experts as it will provide an index to the collected/generated data and relevant documentation, such as who, when, where, how, etc. It contains only metadata and no data. For keeping this log, the PIs of the TA projects will receive a tablet from WP6 with a preloaded app, that will facilitate entering, browsing, and exporting information. The tablet should be returned to the WP6 team after finalisation of the TA acquisition activities. The **AQUARIUS TA Data Summary Log** will serve as a list for the expert data centres to know what data to expect from where and who and as a checklist for the following steps in the data management workflow.

In addition to the TA Data Summary Log it is recommended that the researchers of the TA teams document several aspects of the data acquisition and following processing, so that the later transformation to the common standards of the European marine data management infrastructures will be easier. Annex 1 of the AQUARIUS Data Management

approach document gives a checklist which documentation TA researchers are recommended to maintain for this purpose.

How are you going to organise gathering the input required for the AQUARIUS TA Data Summary Log during the TA activities?

All information relevant to data collection and generation will be documented throughout both the cruise and subsequent data processing phase in the AQUARIUS TA Data Summary Log file. Metadata will also include, where applicable, details on who collected the data, when and where the measurements were made, how they were acquired (instrumentation, protocols, and settings), and any observations relevant to data quality or interpretation.

Metadata collection will follow the structure and mandatory fields defined in the AQUARIUS App and Annex 1 of the AQUARIUS Data Management Approach, ensuring consistency with the project's data management standards. Datasets will be registered in the AQUARIUS App as close as possible to the time of acquisition using the project tablet provided for this purpose.

To ensure consistent and complete metadata capture, one or two designated members of each working shift will be trained and assigned responsibility for maintaining the data log and entering information into the AQUARIUS App. The completeness and consistency of the metadata will be reviewed regularly during the cruise.

Summaries of the logged datasets, together with their associated metadata (but not the datasets themselves) will be transmitted to the AQUARIUS Data Management Team in separate file(s), in accordance with the project's reporting procedures.

Will you in principle be able to organise your TA team during the TA activities in order to gather the input required for the AQUARIUS TA Data Summary Log? What possible challenges do you foresee?

The TA team will be organized during the activities to ensure the accurate collection of all necessary input for the AQUARIUS TA Data Summary Log. A well-structured approach will be adopted, assigning specific responsibilities to team members based on their areas of expertise to cover essential aspects such as data collection, documentation, and metadata recording. This structured methodology will ensure the comprehensive documentation of essential data acquisition details, including personnel involved, timing, locations, and methodologies employed. The use of the preloaded tablet app provided by AQUARIUS will facilitate real-time data entry, ensuring consistency, accuracy, and compliance with established standards.

A coordinated strategy will be implemented to manage various technical components, including physical oceanography, hydrography, and the deployment and operation of instrumentation for data collection. Emphasis will be placed on glider deployments, hydrographic measurements, and water sampling to achieve high-quality data acquisition. The systematic operation and maintenance of specialized equipment, such as CTD and glider systems, will be carried out with a focus on efficiency and precision.

To further enhance the process, standardized protocols will be strictly adhered to for water sampling and filtration activities, ensuring that data is collected and analysed following best practices. Collaboration with remote team members will be facilitated to seamlessly integrate collected data into predictive frameworks and model-based simulations, ensuring full alignment with European marine data management standards.

By implementing clear workflows, robust communication channels, and a proactive approach to potential challenges such as coordinating across various expertise areas, maintaining data consistency, and integrating remote contributions, the TA team will ensure smooth execution of all activities. This comprehensive and structured approach will

guarantee that the requirements for the AQUARIUS TA Data Summary Log are met effectively, contributing to the overall success of the project.

2.4. Increase data re-use (through clarifying licences)

The AQUARIUS project follows an **open data access policy**, and the AQUARIUS Data Management approach is aimed at inclusion and publishing of your collected/generated data as open metadata and data in trusted repositories, within the lifetime of the AQUARIUS project. The preferred licence for use of AQUARIUS data is CC-BY-4.0.

Do you foresee any difficulty for your collected/generated data to follow this policy?

We do not foresee any significant difficulty in following the AQUARIUS project's open data access policy and adopting the CC-BY-4.0 license for our collected/generated data. We are committed to making our data publicly available, ensuring transparency, and promoting data re-use.

For a few datasets, embargo will apply for a maximum period of two years after data collection. This provides us the right timing to process the datasets ensuring rigorous QC/QA methods, interpretation of results, and first scientific publication. This is key for early career scientists working the datasets (e.g., PhD students), where publications are a key part of their thesis. Nevertheless, as soon as papers are published, datasets will be made immediately available through the publication process.

3. Other research outputs

Do you plan to generate other research outputs next to the collected/generated data such as digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.) results?

We aim to generate impactful research outputs beyond the collected/generated data to engage with the scientific community and broader audiences. Our primary focus will be on publishing our findings in open-access journals and presenting at international conferences and workshops to foster scientific discussions and stimulate collaborations. Members of APECS (Association of Polar Early Career Scientists), including Rita Novo and Coraline Leseurre (members), and Leandro Ponsoni and Wieter Boone (mentors), will leverage this global network to disseminate our results and encourage interdisciplinary partnerships.

To reach a wider scientific and stakeholder audience, we plan to participate in events such as Greenland Science Week (<https://www.scienceweek.gl/>) and engage in outreach via specialized blogs hosted by divisions of the European Geosciences Union (e.g., 'Atmospheric Sciences', 'Cryosphere', 'Natural Hazards'). We also plan to develop accessible outreach materials, such as infographics and summary reports aimed at early-career scientists, which will be shared through institutional websites and social media platforms.

Additionally, we aim to increase public awareness of environmental challenges, particularly oceanic and climate-related phenomena, through public talks in schools, online platforms, and community events. These efforts aim to communicate our research findings effectively and ensure they reach relevant stakeholders at local, national, and international levels. Activities will be shared through our outreach channels and by engaging directly with stakeholders.

4. Allocation of resources

Terms & Conditions of AQUARIUS Funding require TA teams to follow the AQUARIUS Data Management Approach. This includes efforts for:

- updating their initial **Data Management Plan (DMP)** per TA project, as included in their original proposal submission, after being awarded and getting ready for the actual TA project deployment
- taking part in WP5 data management training activities to learn about AQUARIUS data management standards and getting experience with tools and services
- using the app during TA activities to complete the **AQUARIUS TA Data Summary Log**
- completing a **Cruise Summary Report (CSR)**⁸ in case the TA project will involve scientific cruises with Research Vessels
- transforming the original metadata and data to the AQUARIUS prescribed standards, performing QA-QC, and making these FAIR and ready for inclusion in AQUARIUS repositories, thereby supported by AQUARIUS expert data centres.

Who will be responsible for data management in your TA project?

The PI, Leandro Ponsoni, is the responsible for coordinating the data management in our TA project. Specific partners teams, collecting specific datasets, will be responsible to providing datasets in the correct format and metadata, following Aquarius recommendations and best practices.

5. Data security

The operators of the repositories as used for the AQUARIUS archival and long-term stewardship have security provisions in use such as for backups, secure storage, firewalls, etc.

What provisions are in place by the TA teams for data security (including backups, secure storage and transfer) in the initial phase, before the data is transferred to the AQUARIUS expert data centres?

In the initial phase, data will be securely stored on internal servers with restricted access. Regular backups will be performed, and data will be stored in multiple locations for redundancy. Secure, encrypted file transfer methods will be used to move data from collection devices to storage. Once transferred, data will remain protected through encryption both in transit and at rest until it is sent to the AQUARIUS expert data centres.

6. Ethical aspects

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing?

⁸<https://www.seadatanet.org/Metadata/CSR-Cruises>

There are no immediate ethical or legal issues anticipated for data sharing, as the project focuses on environmental and oceanographic data that does not involve personal or sensitive information. However, we will ensure that data sharing complies with relevant regulations and agreements, particularly regarding data privacy and intellectual property. If any legal restrictions arise, such as embargoes on specific data types or consent requirements for data sharing, these will be carefully addressed before submission to repositories.

7. Other issues

Will you follow any other national/ sectoral/ departmental procedures for data management, on top of the infrastructure and procedures that AQUARIUS will provide as part of its data management approach?

In addition to the AQUARIUS data management approach, we will follow national and institutional procedures for data management, ensuring compliance with relevant standards and regulations. This includes adhering to specific guidelines set by the Marine Data Archive (MDA) (<https://mda.vliz.be/>), EMODNet Chemistry (<https://emodnet.ec.europa.eu/en/marine-litter>) or other equivalent repositories. We will also ensure data security, quality control, and documentation procedures align with the best practices of these repositories, facilitating seamless integration with AQUARIUS while maintaining consistency across different platforms.