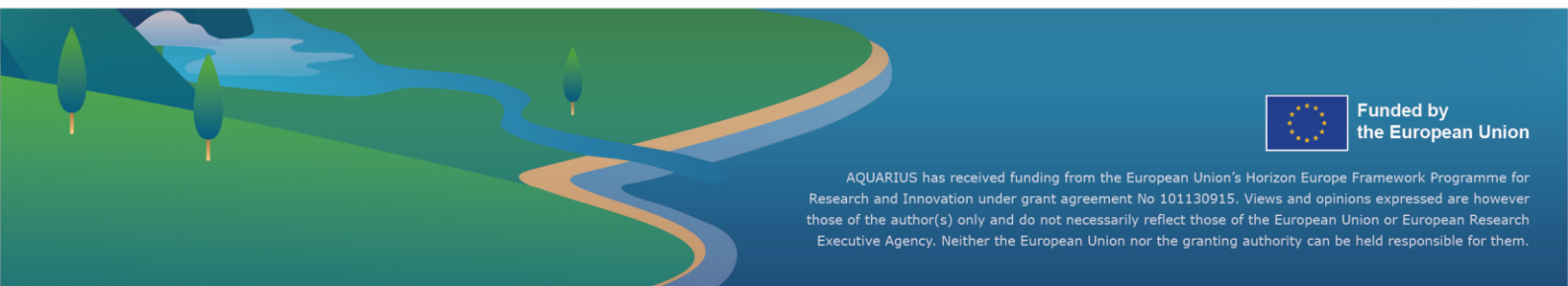




Data Management Plan

AQUAGAPS-AEGEAN

October 2026



Funded by
the European Union

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?

Data generated during the project will contribute to the global monitoring effort within the AQUA-GAPS/MONET worldwide monitoring network of aquatic organic contaminants. The data will be used for publications in peer reviewed scientific journals, addressing the main project objectives outlined in section 2 of the proposal. In addition, data will be used to support the implementation of the Stockholm Convention (SC) on POPs, specifically PFAS data will be reported to the Global Monitoring Programme of the SC.

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

The project will result in data reporting **aqueous concentrations** of selected aquatic **pollutants in surface seawater** and adjacent **atmosphere** at **two monitoring stations** during **two sampling seasons**.

The **metadata** will contain the following information:

Sampling metadata:

Sampler number - Unique sampler number
Sampling Location Code - Unique location number
Sampler preparation batch code
Sample type (*deployment, field blank, laboratory reference, solvent blank*)
Sampler deployment start date
Sampler deployment end date

Location Code metadata:

L_No: Location Code
L_Name: Sampling location description (*e.g. Heraklion Coastal Buoy (HCB), E1-M3A Buoy*)
P_No: Contact person responsible for sampling operations at the Sampling location
L_Country: Country
L_Latitude: Sampling location latitude
L_longitude: Sampling location longitude
L_Type: Water type (*Freshwater or Marine*)
L_Info: Coastal or Open Sea
L_Depth: Sampling depth
L_Elevation: Altitude of the sampling site

Sampling person/team metadata:

P_No P_Name P_Contact P_Institute P_Country P_Email
P_Telephone P_Remark

Sample analysis metadata:

Chemical compound metadata:

OrderNo Number code of each chemical number
GENASIS name Name of the chemical in the Genasis data repository
Par chemical abbreviation (*e.g. anthracene = ANT*)
Name chemical name
Group chemical group name (*e.g. PAH = polycyclic aromatic hydrocarbon*)
Type_comp purpose of the chemical in the analysis (*target compound, recovery internal standard, performance reference compound*)

CasNo	CAS number of the compound
MW	molecular mass of the compound

Compound specific passive sampler material metadata:

Par	chemical abbreviation (e.g. anthracene = ANT)
Name	chemical name
Group	chemical group name (e.g. PAH = polycyclic aromatic hydrocarbon)
Cas_No	CAS number of the compound
A_Mat	polymer material used for sampler construction
Kmw	polymer/water partition coefficient
Source	Reference to the literature source of Kmw

The **data from chemical analysis** in csv format will contain the following information:

A_Labno	RECETOX Labsystem registration number
A_batchNo	Sample processing batch number
A_Type	Sample type (<i>deployment, field blank, laboratory reference, solvent blank</i>)
Par	chemical abbreviation (e.g. anthracene = ANT)
CasNo	CAS number of the compound
Result	Value of the measured chemical amount in the sampler
Unit	Unit of Result (e.g. ng/sample)
A_Mat	polymer material code
SignTxt	Text, Above (>) or below (<) measurement LOQ
Sign	Number Above (1) or below (0) measurement LOQ
Rslt_ng	Value of the measured chemical amount in the sampler in ng/sample

The **data from chemical analysis**, converted to **aqueous concentration**, will contain the following information:

Par	chemical abbreviation (e.g. anthracene = ANT)
Sign	Number Above (+) or below (-) measurement LOQ
Cw_pgL	aqueous concentration in pg/L
LOQw	calculated limit of quantification in water, in pg/L
DEQ	degree of equilibrium of the sampler with the sampled water
LOQng	calculated limit of quantification in sampler, in ng/sampler
SamVol_L	equiv. volume of water, from which the sampler extracted the chemical (L)

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

Contaminants from water and air will be absorbed by passive samplers installed on the sampling site. Sampler deployment will be done in collaboration with the POSEIDON team, (HCMR), of passive water and passive air samplers in the Cretan Sea during two seasons at the two coastal buoys, according to an agreed **standard operation procedure**. Metadata on sampling will be recorded by the POSEIDON team in a standardised **Sample Information Form (SIF)**. Deployments will include various water samplers (sheets of silicone for sampling neutral legacy POPs, a porous polyethylene tube sampler for sampling PFAS, and Baker Bond Speedisk samplers for polar CECs, all mounted on stainless steel holders) and passive air samplers (polyurethane foam disks in metal bowls).

The exposed samplers will be shipped by courier to RECETOX research infrastructure laboratory (Trace analytical laboratory, **TAL**) for analyte extraction and **instrumental analysis**.

At **TAL**, samples will be registered using a unique sample code (A_Labno) in the RECETOX Labsystem. **Extraction and instrumental analysis** using gas or liquid chromatography/tandem mass spectrometry will be performed in the RECETOX central laboratories, accredited for standard analysis methods according to EN ISO/IEC 17025. This includes QA/QC.

The raw result data (Rslt_ng) will be exported from the Labsystem to the user group leader.

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).

A procedure routinely used within the AQUA-GAPS/MONET network will be applied. The user group leader will convert the raw result data (Rslt_ng) from **TAL** to sampling event and chemical specific. The processing includes checking the quality of data reported from Labssystem (extraction recovery, below LOQ values, outlier values) and then import to a uniform reporting format in a standardised dedicated MS ACCESS database for further processing. The conversion of mass/sampler to the corresponding aqueous concentration (pg/L) using an algorithm outlined in (Vrana et al., 2022) is performed automatically by a sequence of queries in MS ACCESS, using the metadata parameters above and the chemical analysis data (Rslt_ng) from the Labssystem.

Final data is exported to an Excel report sheet and submitted to the RECETOX infrastructure data repository GENASIS (Global ENVIRONMENTAL ASSESSMENT Information System), which provides a comprehensive information on contamination of the environment by chemicals (www.genasis.cz).

Final data will also be submitted to SeaDataNet Pan-European infrastructure for ocean & marine data management repository, for integration in the European data management infrastructures as described at the AQUARIUS data management approach.

For creating metadata of reported files, MIKADO tool will be used to create and edit standardized XML metadata records, including CDI and CSR. NEMO tool will be used to convert non-standard ASCII or CSV data into SeaDataNet formats.

See an example data of results from AQUA-GAPS/MONET:

<https://data.genasis.cz/#/outdoor/spatial-distribution#project@212>

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

The data processing can be started only when all samplers are received at RECETOX from POSEIDON after field exposure, processed by the TAL and reported in the Labssystem. According to the project research plan, samples of water and air will be collected by the POSEIDON team before June 2027. The samples will then be transported to Masaryk university, analysed in the TAL laboratory and reported to Labssystem within 3 months after physically receiving the samples in TAL, i.e. by September 2027. Data processing will be ready in October 2027 and reporting finalised in October 2027 so that they can be published in AQUARIUS data repositories before the end of 2027.

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.

The data reports concentrations of contaminants in the water column and adjacent atmosphere. The size of the reported data is not expected to be larger than several megabytes.

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

The data most relevant users are organizations responsible for marine pollution assessment, ocean protection, seafood safety, and regulation of persistent chemicals, namely:

1. Regional seas conventions and marine commissions; The data can support assessments of contaminant distribution, identification of transboundary pollution, comparison of coastal and offshore sites, and evaluation of whether regional pollution-control measures

are working. The particular benefit is a harmonized reference dataset extending beyond an individual regional sea.

2. The Stockholm Convention and UNEP. Marine measurements can contribute evidence on the global distribution, long-range transport, and temporal trends of POPs and candidate POPs. Dissolved marine concentrations are useful for assessing whether restrictions on production and emissions eventually produce lower concentrations in receiving waters and global oceans.
3. Marine Strategy Framework Directive authorities. In the European Union, marine contaminant monitoring results could complement assessments under MSFD Descriptor 8, contaminants.
4. Oceanographic and marine research institutes. Researchers can use the data to investigate contaminant transport by currents, latitudinal gradients, coastal-to-open-ocean gradients. The project was designed to address the lack of globally consistent information on dissolved POP concentrations and includes coastal marine sites in its monitoring concept.
5. Scientists studying air-sea exchange. Combined water and atmospheric passive sampling can determine whether a marine area acts as a sink or secondary source of contaminants. This becomes important after atmospheric emissions decline, because historically accumulated contaminants may volatilize from ocean water back into the atmosphere.

Vrana, B., Smedes, F., & Hilscherová, K. (2022). Passive Sampling of Waterborne Contaminants. In T.-B. Seiler & M. Brinkmann (Eds.), *Methods in Pharmacology and Toxicology Series. In Situ Bioavailability and Toxicity of Organic Chemicals in Aquatic Systems*. (pp. 1–35). Springer. https://doi.org/10.1007/7653_2020_43

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**

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>

Submission service of EMODnet Data Ingestion⁶ and/or SeaDataNet SeaNoe service⁷ 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 files will follow a consistent naming incorporating, where applicable, information such as project name, data type, platform or station name, temporal reference, version.

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

YES. Metadata coding for the AQUARIUS Data Management needs to be adjusted.

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

The processing procedure is documented in section 1. All steps of the procedure, including sampling, sample preparation, instrumental analysis, data processing, data conversion are well documented and traceable, incl. QA/QC.

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 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.

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

NO temporary embargo is required.

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?

⁶<https://www.emodnet-ingestion.eu>

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

Data will be made publicly available at the data repository GENASIS (Global ENvironmental ASsessment Information System), which provides a comprehensive information on contamination of the environment by chemicals (www.genasis.cz).

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.

*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?*

YES. The parallel documentation of all steps, that is done anyway using the AQUA-GAPS/MONET workflow will mean an extra workload to the sampling team (POSEIDON team) and sample processing team (RECETOX laboratory and the user group leader).

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

The observations information will be transferred from the sampling team to the PI for the input required for the AQUARIUS TA Data Summary Log

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?

NO.

If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

No embargo on publication will be applied on the collected research data.

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?

NO.

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)** 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 of the project will be the person responsible for updating DMP, taking part in training activities, using the app during TA activities to complete the TA Data Summary log, completing a Cruise Summary Report (CSR) and for transforming the metadata and data to the AQUARIUS prescribed standards as required in point 4 of this DMP.

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?

Data will be stored using the resources provided by the RECETOX Center or Masaryk University will be monitored and managed in a controlled environment. Network storage includes safe storage servers. Backup of these storages is done by the owner of said storages. The base standard is to make a backup every day and save deleted files for 30 days.

6. Ethical aspects

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

NO.

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?

NO.