European Research Council
Executive Agency

Established by the European Commission

European Research Council (ERC)

ERC Data Management Plan

Project Acronym	Project Number
WHIRLS	101118693

Administrative Details

Name of the project	The impacts of ocean fine-scale whirls on climate and ecosystems
Duration	01.06.2024 – 31.05.2030
Project Management	Anja Reitz GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany Contact: whirls-office@geomar.de
Principal Investigators	Arne Biastoch GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany ORCID: 0000-0003-3946-4390 Sabrina Speich École normale supérieure - PSL, France ORCID: 0000-0002-5452-8287 Sebastiaan Swart University of Gothenburg, Sweden ORCID: 0000-0002-2251-8826 Sarah Fawcett University of Cape Town, South Africa ORCID: 0000-0002-0878-6496
Project funding information	European Research Council (ERC): ERC-2023-SyG - ERC Synergy Grants Grant agreement ID: 101118693 DOI:10.3030/101118693
Data manager of the project	Klaus Getzlaff, GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany Contact: whirls-office@geomar.de ORCID: 0000-0002-0347-7838
Contact for data supply / data delivery for Work Packages (WPs)	WP1: Fine-scale processes impact the large-scale circulation, heat & salt transports, and water masses Sabrina Speich (École normale supérieure – PSL) WP2: Climate-critical air-sea fluxes of heat & carbon are regulated by the fine-scale ocean dynamics Marcel du Plessis (University of Gothenburg) WP3: Fine-scale processes shape the biogeochemical environment and the marine biome Sarah Fawcett (University of Cape Town) WP4: Future ocean & climate models need to implement fine-scale processes for improved predictions and projections Klaus Getzlaff (GEOMAR Helmholtz Centre for Ocean Research Kiel)

Project Description	A closer look at the impact of oceanic eddies Small-scale processes wield immense influence on our climate, yet their intricacies remain poorly understood. Oceanic whirls, or eddies, serve as critical conduits for heat and carbon exchange between the ocean and atmosphere. These exchanges profoundly impact regional and global climate patterns, shaping everything from temperature and precipitation to marine biodiversity. In this context, the ERC-funded WHIRLS project focuses on the Agulhas Current System near South Africa, a hub of eddy activity and marine biodiversity. Through interdisciplinary research spanning scales from 1 km to 100 km, the project aims to unravel the complexities of these processes. By using advanced observation techniques and high-resolution modelling, WHIRLS seeks to enhance our understanding of fine-scale dynamics.
Relevant guidelines for handling research data	- EC-Europe: Open Data, software and Code Guidelines (https://open-research-europe.ec.europa.eu/for-authors/data-guidelines and https://open-research-europe.ec.europa.eu/about/policies#dataavail) - Getzlaff, K., & WHIRLS Operational Management Committee. (2024). WHIRLS Research Data Policy. Zenodo. https://doi.org/10.5281/zenodo.14193288 - The data are affected by legal requirements and conventions: - Nagoya protocol
History of changes	Version 1 [22.11.2024] initial version

1. Data summary

This document describes the Data Management Plan (DMP) of the WHIRLS project in alliance with the project specific Research Data Policy [1]. It aims to describe which data types will be collected and generated during the project and how the compliance to FAIR [2] and open research will be granted. The format of the DMP follows the guidelines suggested for Horizon Europe projects [3]. The DMP is intend to be a *living document* and will be regularly updated, as a minimum in the context of the periodic evaluation/assessment of the project.

The overarching goal of WHIRLS is to improve our understanding and modelling of the ocean's fine-scale dynamics and how they steer ocean-atmosphere interactions and upper-ocean biogeochemistry and biology. Through synergies between the different experts and their teams, we will develop novel methodological approaches that will yield a step-change in our knowledge of the ocean fine scales, improve the modelling of these processes, and enhance the predictability of weather and climate. WHIRLS structures its scientific objectives in four work packages, each introduced by a hypothesis:

WP1: Fine-scale processes impact the large-scale circulation, heat & salt transports, and water masses

WP2: Climate-critical air-sea fluxes of heat & carbon are regulated by the fine-scale ocean dynamics

WP3: Fine-scale processes shape the biogeochemical environment and the marine biome

WP4: Future ocean & climate models need to implement fine-scale processes for improved predictions and projections

Data will be collected or generated in all WPs of the project. The following text gives an overview on the methodology and data types collected, while a more detailed compilation of the observations and data generating activities in the WPs is given in Appendix table 1, with details on specific parameters to be collected as well as file formats, size and file counts in the Appendix table 2.

During the planned multi-ship based observational approach underway data as well as station profiles will be collected such as hydrographical (temperature, salinity, pressure, sound velocity, fluorescence, turbidity), meteorological (air temperature, humidity, rain, wind speed and cloud cover derived from atmospheric radiosounding with balloons) and biogeochemical data (nutrients and components of the carbonate system are measured such as fluorescence, nitrates concentration, pCO2, dissolved inorganic carbon and pH). An array of free-drifting Biogeochemical Argo Floats (BGC-Argo floats) will be deployed during the cruise to collect high-resolution hydrographical and biogeochemical. In addition, various types of uncrewed surface vehicles (USVs) such as Saildrones, Sailbuoys, Wave Gliders will collect high-resolution data at the air-sea interface and the lower atmosphere with different level of resolution. A set of high-resolution numerical simulations for the ocean and the atmosphere will complement the observational approach by providing high-resolution temporal and spatial data. The model hierarchy will comprise so called “ocean-only” configurations (hydrographic data only) as well as coupled climate configurations including the feedback mechanisms from an interacting atmosphere (additionally meteorological and air-sea interface data) with the ocean. Both types of configurations may include modules to simulate the oceanic biogeochemistry (biogeochemical data).

2. FAIR data

This section outlines how WHIRLS will foster dissemination and exploitation of data and results by implementing and applying FAIR data principles: making research data findable, accessible, interoperable and re-usable.

2.1. *Making data findable, including provisions for metadata*

Data will be identified by persistent identifiers (PID) [4,5] provided as digital object identifiers (DOI) for repositories such as WDCC [6], PANGAEA [7], SEANOE [8], SND [9] and Zenodo [10]. Data deposited at institutional repositories and connected data servers like ERDDAP or THREDDS will be assigned DOIs or PID handles for identification depending on the implementation of the repository.

Data sets will be enriched with metadata to enable and foster re-use and interoperability. INSPIRE compliant standards such as ISO 19115 or accordingly enriched simpler formats such as extended Dublin Core are the preferred metadata profile to be easily harvested. Following these standardized formats will allow easy integration of information into national or international portals and data services with well-known and widely used access protocols such as OAI-PMH or OGC-CSW.

In addition, specific metadata are provided using disciplinary specific controlled vocabularies and metadata scheme such as CF-Conventions [11], Attribute Convention for Data Discovery 1-3 [12], SeaDataNet vocabularies [13] and Global Change Master Directory (GCMD) keywords [14].

The project data management will make use of the existing Ocean Science Information System (OSIS) [15] hosted at GEOMAR. OSIS merges information on field data, numerical simulations or experiments with peer review publications and available research data. The metadata will be made publicly accessible immediately. The direct publication of the metadata promotes communication with scientist outside of the project without endangering the

safety of the scientific data prior to publication. OSIS will therefore serve as a one-stop information for research data generated in WHIRLS and linked through the WHIRLS webpage https://whirls.eu. Contact information for access to these large volume data files and standardized metadata information will be provided in OSIS.

2.2. *Making data openly accessible*

Data will be deposited and published in trusted long-term repositories such as WDCC, PANGAEA, SEANOE, SND, depending on the type of the data. Persistent identifiers are mandatory to be used for data deposited in archives or repositories. Depending on the repository these may be resolved as DOI or PID handle. The data will be published with undue delay after necessary processing steps such as calibration and validation are completed.
The large size of model output for individual simulations (each several 100 TBs) prevents a standard use of repositories. Hence, institutional repositories as well as storage and exchange solutions available on High Performance Computing (HPC) systems will be used for this type of data. However, suitable reference simulations at reasonable resolution will be published at WDCC. In any case, specific details on accessing the model output will be available and findable through its metadata deposited in OSIS. Metadata will be made openly available and will contain information to enable the user to access the data. At least subsets and derived data will be published at institutional repositories through data handles when model data are used in publications.
All project partner institutions have clear guidelines and workflows for data handling and publication and follow international standards in obtaining and disseminating the data. Data managers in the project ensure the processing workflow of the data from the instruments or generation to local and international archives.
All data will be stored to a minimum of 10 years. However, most of the above-named repositories, in particular for observational data, provide a longer storage.
No personal or sensitive data will be generated, shared and published.

2.3. *Making data interoperable*

Community-endorsed metadata vocabularies, standards and formats such as SeaDataNet vocabularies, CF-Conventions and Attribute Convention for Data Discovery ensure full interoperability. No project-specific vocabularies are needed.
The data formats used and shared can be opened/processed using public domain software such as Python.

2.4. *Increase data re-use*

The data will be licensed according to the latest version Creative Commons Attribution International Public License (CC BY), preferably using the most “open” CC BY 4.0 license, thus permitting unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.
Data quality assurance will be the responsibility of the individual institution collecting or generating the data. Detailed information on methodology, processing or analysis will be described in data type specific standardized technical reports or standard operating procedures. These will be referenced in the metadata and openly available to document data governance.

3. Other Research output

Digital research outputs (scripts and code) will be made publicly available in technical reports or respective scientific publications and their supplementary files. What is stated above regarding FAIR principles (see 2.), also applies here.

4. Allocation of resources

The task of central data management is covered and paid by the project WHIRLS, while data services will be covered through institutional cooperation agreements with individual data providers and repositories.
Resources for the larger volume of model simulation output is allocated through computing and data application at the HPCs. Additionally, long term storage of large volume data is allocated through institutional cooperation agreement between GEOMAR and the HPC facility at Kiel University.

5. Data security

Data submitted and uploaded to trustworthy long-term repositories are covered against data loss by the repositories data security and storage policy, same accounts for institutional repositories.

6. Ethics

No Ethics issues exist in the project, as proved by the grant agreement.

7. Other issues

No other issues need to be addressed here.

References

[1] Citation for WHIRLS Research Data Policy
[2] Wilkinson, M.D., et al. (2016): The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3: 160018. doi:10.1038/sdata.2016.18
[3] Horizon Europe Data Management Plan template version 1.0, 2021 https://www.openaire.eu/images/Guides/HORIZON_EUROPE_Data-Management-Plan-Template.pdf
[4] ePIC – Persistent Identifiers for eResearch; https://www.pidconsortium.net
[5] Ferguson, C., McEntrye, J., Bunakov, V., Lambert, S., Sandt, S. van . der ., Kotarski, R., Stewart, S., MacEwan, A., Fenner, M., Cruse, P., Horik, R. van ., Dohna, T., Koop-Jacobsen, K., Schindler, U., & McCafferty, S. (2018). D3.1 Survey of Current PID Services Landscape (Version 1). Zenodo. https://doi.org/10.5281/zenodo.1324296
[6] World Data Center for Climate (WDCC) Data Publisher for Earth System model data. German Climate Computing Center, Hamburg; https://www.wdc-climate.de
[7] PANGAEA Data Publisher for Earth & Environmental Science. Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research (AWI); https://pangaea.de
[8] Sea Scientific Open Data Publication (SEANOE); https://www.seanoe.org
[9] Swedish National Data Service (SND); https://snd.se/en
[10] Zenodo. CERN Data Centre & Invenio; https://zenodo.org
[11] CF Metadata Conventions; https://cfconventions.org
[12] Attribute Convention for Data Discovery 1-3; https://wiki.esipfed.org/Data_Discovery_(ACDD)
[13] SeaDataNet – Metadata Service; https://www.seadatanet.org/Metadata
[14] Global Change Master Directory (GCMD). 2024. GCMD Keywords, Version 19.8, Greenbelt,

MD: Earth Science Data and Information System, Earth Science Projects Division, Goddard Space Flight Center, NASA. URL (GCMD Keyword Forum Page): https://forum.earthdata.nasa.gov/app.php/tag/GCMD+Keywords

[15] Ocean Science Information System (OSIS) [Software]. GEOMAR Helmholtz Centre for Ocean Research Kiel; https://osis.geomar.de

- **Table 1: General information about expeditions, experiments, numerical simulations – example data sets**

WP No.	Expedition name	Platform	Sampling Date Start/End	Location Start/End	Research Area	Principal Investigator / Chief Scientist	Organizing Institution
1.1	RV Marion Dufresne II	Ship	2026Q3	Agulhas region	Full water depth, air-sea interface, lower atmosphere	S. Speich	ENS - PSL
1.2	BGC-Argo floats	Argo float	2026Q3	Agulhas region	upper 1000m	S. Speich	ENS-PSL
2.1	Uncrewed underwater glider	Seaglider underwater glider	2025Q4	Agulhas region	upper 1000m	S. Swart	UGOT
2.2	Uncrewed Surface Vehicles (USV)	Sailbuoy	2025Q4	Agulhas region	Air-sea interface	S. Swart	UGOT
2.4	USV	Wave Glider	2025Q4	Agulhas region	Air-sea interface	S. Swart	UGOT
3.1	RV SA Agulhas II	Ship	2026Q3	Agulhas region	upper 1000-2000 m	S. Fawcett	UCT, South African Department of Forestry, Fisheries and Environment
3.2	RV Algoa	Ship	2026Q1/Q2	Agulhas region	upper 1000 m (glider recoveries if necessary)	S. Fawcett	UCT, South African Department of Forestry, Fisheries and Environment
WP No.	**Experiment name**	**Subject/ Theme**	**Date Start/End**	**Location / Laboratory**	**Description**	**Principal Investigator / Chief Scientist**	**Organizing Institution**
3.3	UCT-Marine Biogeochemistry Lab	Isotope ratio mass spectrometry,	2024Q3/ 2030Q2	UCT laboratory	Continued development of analytical methods for the analysis of nitrogen isotopes	S. Fawcett	UCT
WP No.	**Numerical simulation**	**Subject/ Theme**	**Production date**	**Period covered by simulation**	**Description**	**Principal Investigator / Chief Scientist**	**Organizing Institution**
4.1	eINALT20.Z120	ocean-only	2025Q3	1958-2023+	eddy-rich ocean-only simulation	A. Biastoch	GEOMAR
4.2	eINALT100.Z120	ocean-only	2025Q3	2010-2023+	ocean-only simulation resolve down to the km-scale	A. Biastoch	GEOMAR

4.3	FOCI-eINALT20	Coupled (ocean and atmosphere)	2025Q4	1958-2023+	eddy-rich climate model	A. Biastoch	GEOMAR
4.4	eINALT100.Z120-PISCES	Ocean-only with ocean biogeochemistry	2026	2010-2023+	Oceanic biogeochemistry at the km-scale	A. Biastoch	GEOMAR

- **Table 2: Description of data gathered during the project with planned processing and publishing timeline, responsibilities and data access - examples**

WP No.	**Data description** (e.g., expedition, measuring device, sample, method, analysis, model)	**Parameter(s)**	**Responsible person**	**Internal data availability** (planned due date)	**Data format, estimated size and number of files**	**Data publication**		**Optional: Archiving of used software** (repository, name, planned due date)
						Planned due date	repository, DOI, label, licensing, (embargo)	
1.1.1	Ship's underway data	Underway CTD, ADCP, pCO2, wind, cloud, humidity, rain, temperature	S. Speich	6 to 12 months after the cruise [2026Q1/Q3]	NetCDF, 1 GB, order of 10 files	2027	SEANOE, DOI, CC-BY4	
1.1.2	Moving Vessel Profiler (MVP300)	Pressure, conductivity, temperature, fluorescence, backscatter, oxygen	S. Speich	6 to 12 months after the cruise [2026Q1/Q3]	NetCDF, 1 GB, order of 10 files	2027	SEANOE, DOI, CC-BY4	
1.1.3	CTD rosette	Pressure, conductivity, temperature, fluorescence, backscatter, PAR, SUNA	S. Speich	6 to 12 months after the cruise [2026Q1/Q3]	NetCDF, 1 GB, order of 100 files	2027	SEANOE, DOI, CC-BY4	
1.2	BGC-Argo float	Temperature, salinity, PAR, fluorescence, backscatter, SUNA nitrate, oxygen, pH	S. Speich	6 to 12 months after the cruise [2026Q1/Q3]	NetCDF, 1 GB, order of 10 files	2027	SEANOE, DOI, CC-BY4	
2.1	Seaglider Underwater Glider	NO3, current, turbulence microstructure, temperature, salinity, oxygen, light (PAR) chlorophyll, backscatter	M. du Plessis	2026Q3	NetCDF, 10 GB, 10^3 file count	2027	Zenodo, ERDDAP, DOI, CC-BY4	Github
2.2	Sailbuoy (USV)	momentum, micro-meteorology, currents, temperature, salinity	M. du Plessis	2026Q3	NetCDF, 1 GB, order of 10 files	2027	Zenodo, ERDDAP, DOI, CC-BY4	Github
2.3	Wave Glider (USV)	pCO2, radiation, momentum, micro-meteorology, currents, waves, temperature, salinity, chlorophyll, oxygen	M. du Plessis	2026Q3	NetCDF, 1 GB, order of 10 files	2027	Zenodo, ERDDAP, DOI, CC-BY4	Github
3.1.1	Ship's underway data (from discrete seawater samples)	Nutrients (incl ammonium and urea), nitrate and DON isotopes (derived Δ(15,18)), chlorophyll, POC and PON biomass and isotopes, phytoplankton community composition	S. Fawcett	6 to 12 months after the cruise(s) [2027Q1/Q3]	xlsx, csv	2028	Zenodo, SAPRI data centre (South Africa), DOI, CC-BY4	

3.1.2	CTD rosette (from discrete seawater samples)	Nutrients (incl ammonium and urea), oxygen (derived AOU), nitrate and DON isotopes (derived Δ(15,18)), chlorophyll, POC and PON biomass and isotopes, phytoplankton community composition	S. Fawcett	6 to 12 months after the cruise(s) [2027Q1/Q3]	xlsx, csv	2028	Zenodo, SAPRI data centre (South Africa), DOI, CC-BY4	
3.2.1	Biogeochemical rates (bulk and size-fractionated)	Primary production, nutrient uptake, nitrification, N_2 fixation rates	S. Fawcett	12 to 18 months after the cruise(s) [2027Q3/2028Q2]	xlsx, csv	2029	Zenodo, SAPRI data centre (South Africa), DOI, CC-BY4	
3.2.2	McLane pump [Large volume in situ pumps]	Suspended and sinking particle composition (C,N,P, bSi, $CaCO_3$), isotopes (C and N), stoichiometry	S. Fawcett	12 to 18 months after the cruise(s) [2027Q3/2028Q2]	xlsx, csv	2029	Zenodo, SAPRI data centre (South Africa), DOI, CC-BY4	
3.2.3	Particle interceptor sediment traps	Sinking particle composition (C,N,P, bSi, $CaCO_3$), isotopes (C and N), stoichiometry, C and N export flux	S. Fawcett	12 to 18 months after the cruise(s) [2027Q3/2028Q2]	xlsx, csv	2029	Zenodo, SAPRI data centre (South Africa), DOI, CC-BY4	
3.3	Nitrate fluxes	Derived from nitrate concentration and isotope data, with vertical velocity data	S. Fawcett	18 to 24 months after the cruise(s) [2028Q2/Q4]	xlsx, csv	2029	Zenodo, SAPRI data centre (South Africa), DOI, CC-BY4	
4.1	eINALT20.Z120-KFS001	temperature, salinity, currents	F. Schwarzkopf	2025/Q3	NetCDF; up to 95GB; about 3300 files	2025/Q4	WDCC, DOI, CC-BY4	OSIS, simshare json, 2025/Q3
4.2	eINALT100.Z120-KFS001	temperature, salinity, currents	F. Schwarzkopf	2025/Q3	NetCDF; up to 180GB; about 700 files	2025/Q4	WDCC, DOI, CC-BY4	OSIS, simshare json, 2025/Q3
4.3	FOCI-eINALT20.Z120-KRS001	ocean: temperature, salinity, currents – atmosphere: temperature, humidity, wind speed, cloud cover, radiation parameter	R. Schubert	2025/Q4	NetCDF; Up to 95GB; about 4500 files	2026/Q1	WDCC, DOI, CC-BY4	OSIS, simshare json, 2025/Q4

4.4	eINALT100.Z120-PISCES-KLP001	ocean: temperature, salinity, currents – biogeochemistry: 24 tracers (2D and 3D fields)	L. Patara	2026	NetCDF; up to 180 GB; about 1000 files	2026/Q3	WDCC, DOI, CC-BY4	OSIS, simshare json, 2026