BLUE CONNECT

Initial Data Management Plan

DELIVERABLE 1.3

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

Project No.	**101156759**
Project Acronym	**Blue Connect**
Project Title	**Strict protection, restoration and co-management of Marine Protected Areas to ensure effective ecosystem conservation and improved connectivity of Blue Corridors**
Call	**HORIZON-MISS-2023-OCEAN-01**
Start Date	**02.09.2024**
Duration	**36 Months**
Delivery Title	**Initial Data Management Plan**
Due Date Delivery	**28.02.2025**
Date of Submission	**28.02.2025**
Deliverable Lead Partner	**Flanders Marine Institute (VLIZ)**
Dissemination Level	**Public**

Recommended citation:
Whatley L., Kvanneid A.J. (2024) Initial Data Management Plan. Deliverable – D1.3 under the WP1 of the BLUE CONNECT project (GA n° 1011656759).

Document History			
Version	**Date**	**Note**	**Revised by**
01	04.02.25	General revision	Mercedes R. Wansart (SUBMARINER), Francisco R. Barboza (UTARTU), Barbara Bauer (UMA), Rita Trabulo (RBINS), Margarita Stancheva (CCMS), Laura Friedrich (NIVA)
02	14.02.25	Specific revision	Mercedes R. Wansart, Ivana Stojanovic (SUBMARINER), Francisco R. Barboza, Jonne Kotta (UTARTU)

Name	**Organization**
Lawrence Whatley	VLIZ
Aase J. Kvanneid	NIVA

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

Contents

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

1. Acronyms

AB	Advisory Board
BODC	British Oceanographic Data Centre
DMP	Data Management Plan
DOI	Digital Object Identifier
DTO	Digital Twin of the Ocean
DS	Demonstration Site(s)
EAB	Ethical Advisory Board
EMODnet	European Marine Observation and Data Network
EU	European Union
FAIR	Findable, Accessible, Interoperable, and Reusable (data)
GBIF	Global Biodiversity Information Facility
GDPR	General Data Protection Regulation
IMIS	Integrated Marine Information System
INSPIRE	Infrastructure for Spatial Information in Europe
IPR	Intellectual Property Rights
KPI	Key Performance Indicator
LSWG	Local Stakeholder Working Group
MDA	Marine Data Archive
MPA	Marine Protected Area
MSP	Maritime Spatial Planning
NDA	Non-Disclosure Agreement
OBIS	Ocean Biodiversity Information System
WoRMS	World Register of Marine Species
WP	Work Package

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

2. Executive summary

This document is the initial data management plan (DMP) of the BLUE CONNECT project. The DMP presents a summary of the data that will be collected, generated and (re)used in the project and explains the purposes of using this data. It also includes a definition of the data and outlines the steps necessary to maximise the FAIRness (Findability, Accessibility, Interoperability, and Reusability) of the project's outputs. Issues concerning ethics and the General Data Protection Regulation (GDPR) are also addressed. Key data management issues are discussed, including common (meta)data standards and formats, the data flow to be followed throughout the project, data quality control, (meta)data vocabulary, publication policy, and data licensing. Therefore, the DMP provides data management guidelines for the partners to follow. It should be considered a living document; updates will be made throughout the project's lifetime with more specific information on data collected, generated, and used in BLUE CONNECT. These updates will be used to write the Final DMP (D1.4) in month 42, which will also outline data management after the end of the project.

3. Data summary

3.1. Purpose and context of data collection and (re)use

BLUE CONNECT aims to support the effective achievement of marine conservation and restoration targets in the EU and beyond, thereby contributing to the implementation of the EU Mission "Restore our Ocean and Waters". The project's final output will be the Blueprint; a systematic ecosystem-based approach to marine conservation planning and management which will provide tools for the identification of conservation objectives, the design and implementation of conservation measures to achieve those objectives, and the monitoring and assessment of conservation effectiveness. This will result from a collaborative, science-based, data-driven systematic approach for the designation and management of Marine Protected Areas (MPAs) and a co-development approach in cooperation with stakeholders and citizens in 12 Demonstration Sites (DS) located in four EU sea basins. The DS will provide data, information, best practices, and lessons learnt to inform and guide the development of the Blueprint and will serve as testing grounds for tools and (co-)management schemes.

BLUE CONNECT will therefore reuse existing data and collect new data to support its activities and produce its expected outputs, which include reports, models, algorithms and digital applications, guidelines, and maps.

Data needed for internal project management will be mostly managed by **WP1**; much of this will be personal or financial data related to project partners and the Advisory Board

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

(AB). WP1 will not produce any relevant scientific data of use to other WPs or outside the project. BLUE CONNECT's data management will be carried out under T1.2.

WP2 will (re)use ecological, biogeochemical, hydrodynamic, and remote sensing data to develop and test tools related to biodiversity monitoring and assessment, spatial structural and functional connectivity, ecosystem functioning, and pressure management in MPAs. Much of this data will come from other BLUE CONNECT tasks, particularly T4.1, and data platforms including the European Marine Observation and Data Network (EMODnet) and Copernicus Marine. WP2 will also collect data to support their tasks through literature reviews and tailored requests to the Demos Sites.

WP3 will map pressures on the marine environment, review governance frameworks to identify barriers and enablers for conservation measures, assess stakeholders' ocean literacy, and compare (co-)management methods to create an MPA best practice governance toolbox, requiring a variety of ecological, socio-economic, and regulatory data.

The focus of **WP4** is to engage stakeholders and citizens through LSWGs, public surveys, and actions to showcase effective co-ownership and co-management, including sustainable financial models. The data needed for these tasks will consequently be mostly socio-economic. WP4 will also compile available baseline data and information from the DS under T4.1 which will serve as a knowledge base for many other tasks.

WP5 will identify best practices among MPA monitoring and evaluation systems for conservation effectiveness and test them in the DS using innovative techniques. In-situ monitoring data will therefore be produced, and, along with existing data, used to validate the tested systems and techniques and identify lessons learnt.

The project's final output, the Blueprint, will be produced under **WP6** using the data collected and compiled by WP2, WP3, WP4, and WP5.

WP7 is responsible for BLUE CONNECT's communication, dissemination, and exploitation, and will collect data on its activities to monitor its progress towards its KPIs.

This Initial Data Management Plan (DMP) outlines the ways in which data will be collected, generated, processed, archived, and made openly accessible (when appropriate) throughout the lifespan of the project. The DMP will be updated at the end of BLUE CONNECT (D1.4, month 42) with more detailed information about the project's data collection and management processes and long-term data archiving after the end of the project.

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

3.2. Broad-scale types of data

Data can be defined as units of information which, in a scientific context, are derived from research outputs and provide evidence for knowledge claims without having truth-value in and of themselves (Leonelli, 2015). Examples include experiment results, in-situ measurements, observations from fieldwork, remote observations, results of surveys and questionnaires, interview recordings, information extracted from documents, images, and maps. Data can be classified in different ways, e.g., the Maritime Spatial Planning Data Framework (Abramic et al., 2023). This section gives an overview of the broad-scale types of data expected to be reused and generated in BLUE CONNECT.

3.2.1. Biological data

Biological data consist of observations of organisms and their attributes, including biotic measurements and life-history traits. These may include measures of abundance, body size or shape, age, sex, biomass, growth, movement, and biological interactions. They can be collected from field observations (e.g., visual surveys, photo quadrats, and video transects), sample collections (e.g., trawls, nets, corers), acoustic surveys, imaging, molecular techniques, or remote sensing using drones/satellites. Data on species densities or abundance, taxonomic diversity, genetic diversity, functional diversity, species distributions, population dynamics and biogeochemical processes can be generated. These data are often also available from data repositories and European data platforms such as the Global Biodiversity Information Facility (GBIF), EMODnet, and the Ocean Biodiversity Information System (OBIS). Biological data are often related to environmental data; habitat maps, for example, are derived from the distribution of habitat-forming species and physical data. In BLUE CONNECT, biological data will be collected by WP5 to establish ecological baselines in the selected Demo Sites. This data, along with existing biological data compiled by T4.1, will be used by WP2, WP3, WP4, and WP5, and will be integrated into the final Blueprint by WP6.

3.2.2. Environmental data

These data are measurements from the ocean-atmosphere boundary layer, the water column, and the sea floor. These can be physical data (e.g., temperature, salinity, water turbidity, currents, sea level, waves, air temperature, air pressure, wind speed and direction, rainfall, etc.) or biogeochemical data (e.g., nutrients, blue carbon, sediment C/N content, chlorophyll, dissolved organic matter etc.), which are often related to biological data. These variables can be measured in situ from ship-based measurements, buoys, robots, and sensors, or through remote sensing. Environmental data can also be obtained from data repositories and European data platforms such as EMODnet and Copernicus Marine Service. Environmental data collected by WP5 in selected DS will be used as an ecological baseline for the sites. The connectivity and ecosystem modelling

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

(WP2) and pressure analysis (WP3) will be mainly supported by existing environmental data compiled by T4.1. The results will be fed into the final Blueprint by WP6.

3.3.3. Quantitative socio-economic data

Quantitative socio-economic data consist of numerical data about human activities, views and opinions, economics, and the relations between them. Data concerning human activities at sea include information on the space and/or structures used to conduct human activities and their economic value. This data type may also include coastal data (e.g., demographics, main industries, agriculture, etc). Such data are used to identify and valuate ecosystem services and assess adaptation and vulnerability by studying population and human development, and economic conditions, amongst others. They can be collected in various ways, for example from the databases of commercial organisations and enterprises (e.g., fisheries logbooks and vessel monitoring systems), or quantitative surveys. Quantitative socio-economic data can also be obtained from data repositories and platforms, such as EMODnet Human Activities. If these data contain sensitive and/or personal data, the ethical and legal guidelines described in section 9 must be followed. Existing quantitative socio-economic data will be used to create pressure models (WP2), to identify hotspots of impacts (WP3), to analyse financial models (WP4), and ultimately validate the Blueprint (WP6). Some quantitative socio-economic data may also be collected from the DS using surveys.

3.3.4. Qualitative data

Qualitative data is non-numeric data generated from questionnaires, interviews, focus groups, workshops, observations, or textual analysis of documents and/or policies. These data usually appear in formats such as text, audio, or video. As sensitive and/or personal data may be present in qualitative data; BLUE CONNECT will ensure that ethical and legal guidelines are followed when collecting and sharing these data (Section 9). In BLUE CONNECT data related to the DS will be collected using surveys and interviews to assess anthropogenic pressures, ocean literacy, co-management approaches and governance frameworks (WP3) and compile a knowledge base which will support other tasks (WP4). Existing qualitative socio-economic data will be reused, through formats such as literature reviews and story maps. In some cases, it may be useful to convert qualitative data into quantitative values, for example, for visualisation or integration into a modelling framework.

3.3.5. Spatial governance and administrative data

This data type includes information on regional, national, and supranational governance structures, institutions, and frameworks. Combined with a spatial element, these data give information about the location of administrative and/or jurisdictional features across space (and sometimes also time), for example national or subnational borders and boundaries, MPAs, and MSP zones. These may be associated with biological or

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

environmental data, for example an MPA which aims to protect a certain species or undersea feature, or socio-economic data, for example an MSP zone focused on the economic exploitation of a particular resource. In BLUE CONNECT these data will be collected and used by WP3 to create a governance toolbox, and spatial MPA data will be compiled by T4.1 for use in WP2, WP3, WP4, WP5, and WP6.

3.3.6. Spatial data products

Spatial data products are derived data; they include raster matrices (gridded data and images) and vector features, such as points, lines, and polygons, and can include temporal data at many scales. Maps or spatial products are commonly produced from observations, interpolation of data, or from model outputs (both statistical and numerical). Examples of spatial data products include anthropogenic activities and pressures maps (e.g., fishing, shipping, aquaculture, marine litter, and pollution). They are often the result of multiple datasets and might be associated with more than one data type, including biological, environmental, and/or socio-economic data. Administrative spatial data can also be integrated or overlayed with such products to explore the interactions between governance, management, and ecological or environmental processes. Spatial data products to be produced include the outputs of pressure maps (T3.1) and connectivity models (T2.2) which will be further developed into ecosystem models (T2.3). Some of these data products may ultimately be integrated into the Blueprint (T6.3).

3.3. Data formats

BLUE CONNECT will produce data in a great variety of formats. We recommend adopting the data and metadata formats commonly used by the marine community for data submissions to increase the data's FAIRness. Table 1 gives an overview of the common data file formats which BLUE CONNECT partners should follow where possible. We also recommend avoiding the use of proprietary software formats (e.g., .docx, .xlsx, .shp) to increase the accessibility of the data. If the original data files are in a proprietary format, they should be converted to a more common, non-proprietary data format before making them publicly available and the transformation method employed should be documented in the metadata.

Table 1: Accepted common file formats by data format

Data format	Accepted common file formats
Text	• TXT • ODT • XML / HTML
Tabular data	• CSV • TSV

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

Images	• TIFF • JPG • PNG
Sound files	• WAVE (.wav) • FLAC • MP3
Videos	• MKV • MPEG / MPG
Spatial data products (maps, model outputs, etc.)	• NetCDF • GeoTIFF • OGC GeoPackage (.gpkg)
Code	• Plain text (usually with an extension that represents the source language)

3.4. Expected reuse of existing data

BLUE CONNECT will reuse existing data from a wide variety of sources to produce its outputs. Many of these data will be compiled by T4.1 with assistance from the DS to produce an inventory of baseline data to support other tasks. Existing data will also be sourced from open-access data repositories, platforms, and inventories, such as Zenodo, EMODnet, and Copernicus. Table 2 provides a summary of existing data envisaged to be reused by each task. Some of these data, such as ecological and environmental data derived from platforms including EMODnet and Copernicus, will be large (>5MB), while the remainder is mostly expected to be of relatively small sizes (<5MB).

Table 2: Summary of data to be reused by each work package and task

WP	Data to be reused	Sources
WP1	**T1.4**: Data and information on Demo Sites	Demo Sites, Demo Site board, other WPs
WP2	**T2.1**: Baseline data, including list of tools needed and currently in use at Demo Sites	WP4 (T4.1), other projects (e.g. MSP4BIO, CLIMAREST, BLUE4ALL…), scientific publications
	T2.2: Biological and environmental data relevant for the analysis of structural and functional connectivity	T2.1, T4.1, European data platforms (Copernicus, EMODnet, etc.), and other Horizon Europe projects (e.g. MARBEFES, MSP4BIO, BLUE4ALL)
	T2.2: Tools, models and algorithms for the analysis of structural and functional connectivity in marine system	WP4, other Horizon Europe projects and scientific publications (along with their openly distributed associated code).
	T2.3: Models and tools developed in WP2; hydrodynamic, biogeochemical, external pressure, and ecological data	Other WP2 tasks, T4.1, platforms such as Copernicus Marine, EMODnet, and Eionet.

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

WP	Data to be reused	Sources
	T2.4: Hydrodynamic, biogeochemical, external pressure, and ecological data (habitat distribution, species population and distribution)	Platforms such as Copernicus Marine, EMODnet, Eionet, data stored in national databases not integrated in EMODnet (national monitoring and mapping programs and research projects datasets)
	T2.5: Data for WP2 summary and compilation	T2.1 - T2.4
WP3	**T3.1**: Data on socio-economic drivers and pressures, habitat and species sensitivity maps	WP2, WP4
	T3.2: Data on governance frameworks at the Demo Sites	T4.1
	T3.3.1 Baseline data on Demo Sites related to Ocean Literacy, inclusive and just management/co-management	Demo Site leaders and stakeholders, T4.1, other projects (e.g., EU4Ocean Coalition), scientific literature
	T3.3.2: Data from other tasks, including baseline data on Ocean Literacy, and inclusive and just management/co-management in selected Demo Sites	T.3.3.2 (earlier phases), T3.1, T3.2, T3.3.1, T4.2, Demo Site leaders and stakeholders, scientific literature
	T3.4: Data produced by T3.2 and T3.3	T3.2, T3.3
WP4	**T4.2**: Data from LSWG interactions (workshops for co-development and validation), public surveys and consultations	WP2, WP3, T4.1, and WP5
	T4.3: Data on available funding for MPAs	Demo Site leaders and stakeholders, online sources
	T4.4: Best practices and lessons learnt at the Demo sites	T4.1 – T4.3, Demo Sites
WP5	**T5.1**: Best practices and criteria in biodiversity monitoring and evaluation from the Demo Sites	T4.1
	T5.2: Data on innovative monitoring techniques from the Demo Sites	T4.1
	T5.3: Data produced by T5.1 and T5.2 to assess ecological effectiveness in the Demo Sites	T5.1, T5.2
	T5.4: Data on successful conservation outcomes and lessons learnt in the Demo Sites	Demo Sites, T2.4, T2.5, T3.1, T3.4, T4.1, T4.4, T5.3
WP6	**T6.1**: Data on best practices and lessons learnt	WP2, WP3, WP4, WP5
	T6.3: Data for the Blueprint	Other WPs

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

3.5. Expected generation/collection of data

BLUE CONNECT will collect data from literature reviews, surveys and questionnaires with stakeholders in the DS, and in-situ monitoring using innovative techniques (T5.2). Table 3 provides a summary of data envisaged to be collected by each task, which will likely be of relatively small sizes (≤5MB) with the possible exception of the in-situ data produced by T5.3.

Table 3: Summary of data to be collected by each work package and task

WP	Data to be collected	Sources
WP1	**T1.1**: Partners' contact details and pictures (closed)	Partners
	T1.2: Partners' financial information (closed)	Project partners
	T1.3: AB members' details and NDAs (closed)	AB members
	T1.4: Challenges, good practices, information and lessons learnt from the Demo Sites	LSWGs and Demo Sites
WP2	**T2.1**: Data on methodologies for monitoring and assessing marine biodiversity, relevant indicators and existing models currently used to understand connectivity, ecosystem functioning and vulnerability	Scientific papers and grey literature
	T2.2: Necessary data to support the new models and tools specifically tailored to address the needs of Demo Sites	Tailored requests to DS in collaboration with WP4, collaboration with other Horizon Europe projects, literature review
	T2.3: Outputs of hydrodynamic, biogeochemical and ecological process-based models	Models developed by WP2
	T2.4: Outputs of WP2 models and tools	WP2 models and tools
WP3	**T3.1**: Data on socio-economic drivers and pressures in the Demo Sites	Literature review, surveys on expert knowledge
	T3.2: Data on institutional barriers and enablers for effective MPAs	Literature review
	T3.2: Data on governance frameworks of the Demo Sites	Literature review
	T3.2: Data on institutional barriers and enablers at selected DS	Literature review (and if needed Demo Site survey)
	T3.2: Data on details about barriers and enablers to effective MPAs at selected Demo Sites	Interviews with Demo Site leaders and/or stakeholders
	T3.3.1: Baseline Ocean Literacy data	Demo Site questionnaires
	T3.3.2: Data from literature review and context review of conservation processes in selected Demo Sites	Literature review, media analysis
	T3.3.2: Data from interviews with Demo Site leaders and stakeholders	Semi-structured interviews
	T3.3.2: Photographs for the Story Map	Photographs of Demos Sites etc.
	T3.4: Evaluation of the governance toolbox	Demo Sites, literature review

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

WP	Data to be collected	Sources
WP4	**T4.1**: Inventory of data sources from Demo Sites	DS and data repositories and platforms
	T4.2: Data and contact details of LSWG members	Demo Site leaders, stakeholders
	T4.3: Data on MPA funding plans and financial models	Demo Site leaders
	T4.4: Compilation of best practices and lessons learnt from Demo Sites	Demo Sites
WP5	**T5.1**: Data on best practices	Literature review
	T5.2: In-situ ecological and environmental DS monitoring data	In-situ monitoring using innovative techniques
	T5.4: Factors contributing positively or negatively to the ecological performance of the DS	Demo Sites, Literature review
WP7	**T7.2**: Local stakeholders database (closed)	LSWGs, Demo Site leaders
	T7.3: Workshop participant database (closed)	Workshop participants, Demo Site leaders
	T7.4: Newsletter subscription list (closed)	Newsletter subscription form on the website
	T7.5: Mission Ocean projects and contacts database (closed)	Mission Ocean communication group

3.6. Data utility

Data collected by BLUE CONNECT will be used to create and test the Blueprint, which will improve the designation and management of MPAs, including shifting their status from “protected” to “strictly protected” and quantifying their conservation effectiveness. It will do this by providing a holistic modelling framework, incorporating tools for the establishment of science-based conservation objectives, the co-definition of measures with stakeholders, the design of conservation measures focusing on passive restoration and co-management, and for effectiveness measurement and monitoring. This will ultimately contribute to the implementation of the European Green Deal and its EU Biodiversity Strategy for 2030, the European Restoration Law, and the EU Mission "Restore our Ocean and Waters". The Blueprint will be openly available but specifically aimed at MPA managers, regulators/authorities, MSP Planners, researchers/research organisations, local NGOs and communities, and intergovernmental organisations. The project’s target groups will be further specified in D7.1. WP7 is responsible for communication and dissemination directed towards Demo Site-specific groups such as local stakeholders, MPA managers, and citizens (T7.2, T7.3), and marine conservation- and management-focused policymakers and researchers (T7.4, T7.5).

By following Horizon Europe's Open Science policy, BLUE CONNECT will ensure open access to its data wherever possible. Relevant tools and data produced by the project, particularly in-situ data from WP5, will be integrated into the European Digital Twin of the

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

Ocean (DTO) and data platforms such as EMODnet. Data that meet the requirements of ingestion into EMODnet or OBIS will first flow to those data platforms, from where there will be an automated flow to the DTO data lake. Alternative flows will be implemented for data which does not meet EMODnet or OBIS ingestion requirements. Some data may also be of interest to other Mission Ocean projects and related EU projects and initiatives such as Protect Baltic, MPA Europe, and the Mission Lighthouses. T7.5 will focus on building cooperation and partnerships between these projects, including sharing data.

4. Data organisation / data flow

The Data Management Officer (VLIZ) will coordinate the process of data management during the project. During the data collection phase, each partner will be responsible for managing the datasets they collect and must follow the DMP. During initial data generation and quality control, the data will be stored at the data provider's organisation and uploaded to the BLUE CONNECT SharePoint folder after taking any relevant steps to remove personal data (see section 9). It is the data provider's responsibility to ensure that the data is stored securely at this stage. All internal sharing of data, outputs, and drafts will happen via the SharePoint folder, hosted by SUBMARINER. No personal or sensitive data should be uploaded to the SharePoint folder. Once the material has been uploaded, other project members should be informed appropriately (e.g., at a meeting and/or via email or similar).

As soon as possible, i.e., after initial quality control/pre-processing, and no later than 6 months before the end of the project, the data manager will upload the data to the Marine Data Archive (MDA), hosted by VLIZ's Data Centre, where it will be archived. In this archive, the data can initially be kept private and shared only between BLUE CONNECT partners. When the data is published, or at the latest 2 years after the end of project, datasets will be made openly accessible on MDA and receive a DOI. The metadata will be integrated into the Integrated Marine Information System (IMIS), VLIZ's marine science information system. The submission workflow (Figure 1) is as follows:

Metadata submission

1. Data providers submit the metadata via SharePoint using the metadata template (Table 4).
2. Data managers check the metadata is correctly filled in.
3. Data managers integrate the metadata in IMIS.
4. Data managers assess the most appropriate data format and the data's compliance with the standards for submission to MDA and inform the data submitter.

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

Data submission

5. If necessary, data providers format the dataset with support from the data manager.
6. Data providers submit the datasets via SharePoint if the files do not exceed 3GB, or via their online file transfer platform of choice (e.g., ftp, Eudat B2Drop, wetransfer, web services) if the files exceed 3GB. Additional metadata fields may be requested by data managers depending on the repository targeted.
7. Data managers check that the data are provided in the appropriate format and provide an overall quality check.
8. Data managers integrate the dataset in the MDA.
9. If relevant, data managers inform other data repositories (e.g., EMODnet, OBIS, GitHub) of the intention to submit the datasets and data providers assist the repositories with submitting their datasets and passing the necessary quality control steps.

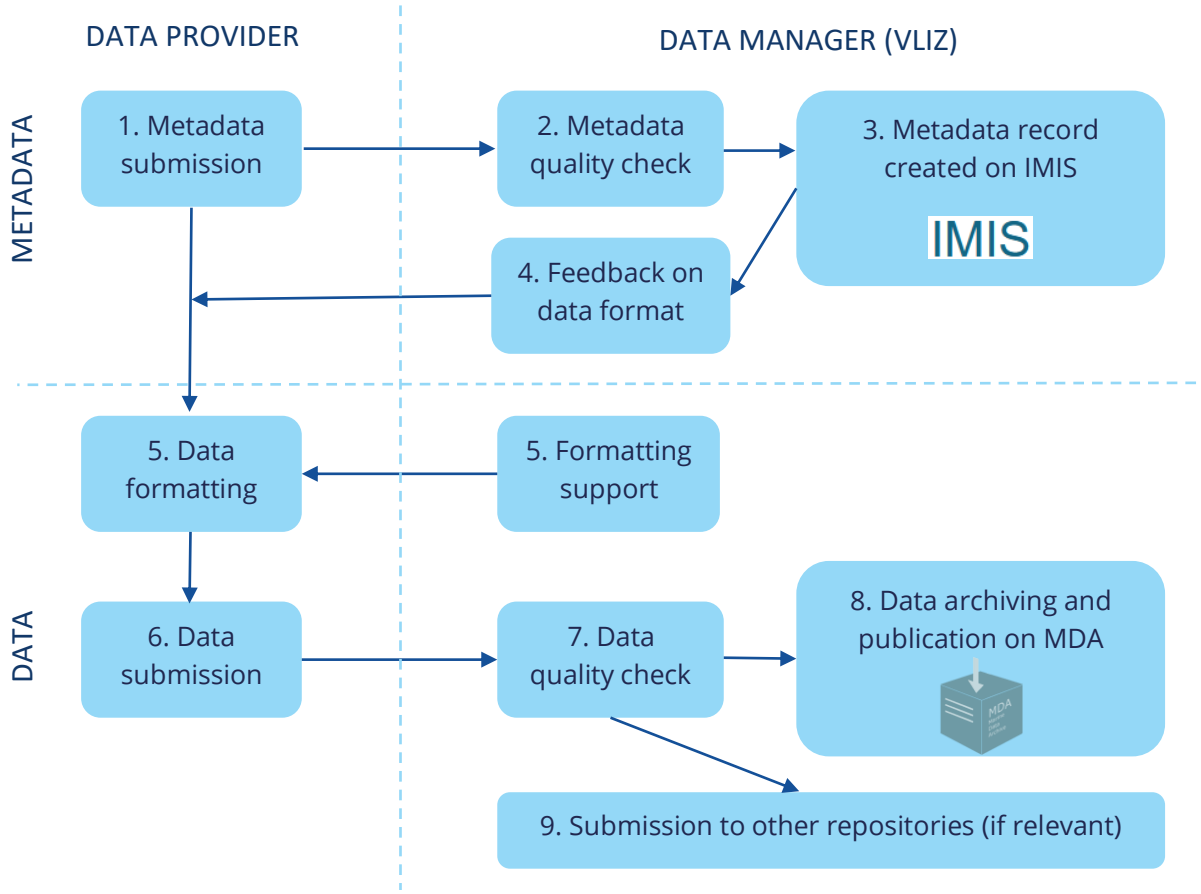

Figure 1: Data flow including metadata and data submission processes

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

The data provider will be responsible for quality control during data collection and for submitting the data in a recommended or accepted file format. Before the data are stored on the MDA or in another repository, the data provider will also be responsible for data security and back-up and recovery of the data. This could involve, for example, storing copies of the data in the BLUE CONNECT SharePoint folder, another secure cloud service, and/or an external hard drive.

The main purpose of MDA is long-term preservation, but to increase findability and reuse, data and derived products may also be published in different international repositories such as EMODnet, OBIS, GBIF, and GitHub depending on the relevance and type of the data. Relevant data will also be integrated into the DTO. Although most of the project data will be uploaded to MDA, it may not be suitable for some data types such as raw data which take up large amounts of storage space (audio, images, etc.). These will be archived in other, data type-specific, repositories. Before publication, these data types can be shared upon request between partners. However, the data derived from these (e.g., ecosystem valuation data derived from audio interviews) will be archived on MDA. For all datasets and data products generated by the project, an IMIS metadata record will be created, linking to MDA. In cases where there is both a 'raw' dataset and a 'derived' dataset, these will be linked to each other.

5. FAIR data

5.1. Making data findable, including provisions for metadata

Data findability will be maximised by assigning a DOI or another globally unique and persistent identifier (assigned by MDA) to each dataset/data product and publishing its associated metadata record. Metadata of all datasets/data products will be stored and published on the IMIS, which is hosted by the Data Centre in the Flanders Marine Institute (VLIZ). Some data may also be submitted to repositories such as EMODnet or OBIS and/or be integrated into the DTO, where relevant, to increase findability. The metadata will adhere to international standards such as ISO 19115 and will also be available in EML (Ecological Metadata Language) 2.2.0, INSPIRE and Dublin Core formats where relevant. Table 4 provides a template for IMIS- and INSPIRE-compatible metadata.

Table 4: IMIS- and INSPIRE-compatible metadata template.

Metadata field	Metadata sub-field	Description	Importance
Creator	• Creator name • Name identifier (e.g., ORCID iD)	The main researchers involved in producing the	Mandatory

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

Metadata field	Metadata sub-field	Description	Importance
	• Affiliation (organisational or institutional)	data, or the authors of the publication, in priority order.	
Title	-	A name or title by which a resource is known.	Mandatory
Publisher	-	The name of the entity that holds, archives, publishes, prints, distributes, releases, issues, or produces the resource.	Mandatory
Publication year	-	Year of publication	Mandatory
Subject	-	Subject, keyword, classification code, or key phrase describing the resource, e.g., biology, economics.	Recommended, not mandatory
Contributor	• Contributor type (e.g., contact person, data collector) • Contributor name • Name identifier (e.g., ORCID iD) • Contributor affiliation	The institution or person responsible for collecting, managing, distributing, or otherwise contributing to the development of the resource.	Mandatory when applicable
Date	• Date type (e.g., accepted, available)	Different dates relevant to the work in YYYY-MM-DD format (e.g., date of acceptance, submission...).	Mandatory
Resource type	-	A description of the resource, e.g., dataset, text, model.	Recommended, not mandatory
Size	-	Size of the resource, e.g., number of MBs or pages.	Optional
Format	-	Format of the resource, i.e., file extension.	Optional
Rights	-	Any rights information for this resource, e.g., licence, embargo.	Mandatory when applicable
Description	• Description type (e.g., abstract, methods)	All additional information that does not fit in any of the other categories, e.g., an abstract providing a short summary of the dataset and its content.	Mandatory when applicable

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

Metadata field	Metadata sub-field	Description	Importance
Geolocation	• Geolocation point (WGS 84 coordinates) • Geolocation box (WGS 84 coordinates) • Geolocation place (text)	Spatial region or named place where the data was gathered or about which the data is focused.	Optional
INSPIRE Theme	-	Relevant INSPIRE theme.	Mandatory when applicable

5.2. Making data accessible

In accordance with Horizon Europe's Open Science policy, all data collected or produced in BLUE CONNECT should be open access and unrestricted by the end of the project if possible. All data, except for personal data, will therefore be made freely accessible at no charge to third parties and available for long term reuse by depositing it in MDA. The data and metadata will be available for as long as those portals are available and, as per the policy in MDA and IMIS, metadata will be available in perpetuity, together with the data, unless the data owner requires it to be removed.

Exceptionally, some data will not be openly accessible, namely datasets containing personal data such as raw data from interviews and questionnaires with Demo Site stakeholders, LSWGs, and local communities (WP3, WP4, WP7). In these cases, and if possible, datasets containing personal data will be pseudonymised or anonymised before being made openly accessible in accordance with GDPR (see section 9 – Ethics). If subjects of the research can still be identified even after anonymising the data, which may happen if the data relates to small communities or social groups, the data will be made available in aggregated or summarised form. In this case, as much information as possible will be retained while removing any references to research subjects.

5.3. Making data interoperable

To maximise interoperability, internationally recognised standards and controlled vocabularies will be applied to metadata and data. The use of common controlled vocabularies in the metadata and data, and the use of common data formats, are important prerequisites for consistency and interoperability, and enable records to be interpreted by computers. BLUE CONNECT data generators should use the recommended file formats indicated in Table 1 for their datasets. The metadata must be submitted in English, and we recommend that the data are translated into English before

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

submission if possible. Biological data should be structured to comply with the Darwin Core standard and environmental data should adhere to the Global Ocean Observing System's guides, best practices and standard operating procedures.

The following controlled vocabularies for standardisation of names should be followed:

- BODC Vocabularies (British Oceanographic Data Centre), for parameters, units, and instruments.
- World Register of Marine Species (WoRMS), for crucial quality control of taxonomic data, for biological data or any other type of data that includes species names. Taxonomy can be standardised by referencing the matching aphia_id from WoRMS. In this way, differences in spelling, or use of species names will be corrected.
- Marine Regions, for standardised marine georeferenced place names and areas.

When unique or uncommon ontologies or vocabularies are used by data providers, they must be mapped to more commonly used ontologies.

5.4. Increase data reuse

Most of the data generated by BLUE CONNECT will be open-access and licensed under Creative Commons CC-BY (Attribution) or CC-0 (Public Domain Dedication) so that it can be reused, in line with Open Science principles. The licences will be specified in the metadata records. Data will become open-access and available for reuse as soon as research involving that data is published. The project's reports and results will become available immediately on BLUE CONNECT's website and in an open-access repository. The Blueprint, the project's final output, will be made available as a web platform. No embargoes or moratoriums are currently foreseen to be applied to any data collected in BLUE CONNECT. If a moratorium or embargo is applied to any data, it will be detailed and justified in the Final DMP (D1.4).

Data not (yet) included in published research will become open-access at the latest 2 years after the official end-date of the project. Access rights after disclosure will depend on the Creative Common licence those data or products hold, as described above. For products created with restricted data, any necessary discussions with the data owners for the reuse of those datasets and or/products need to be arranged by the partner using those restricted datasets.

Initial quality control of datasets produced by BLUE CONNECT is the responsibility of the data provider. Additional quality control will be performed by the data manager (VLIZ) before publishing and being made available through MDA. Any quality control procedures applied will be described in the metadata.

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

6. Other research outputs

In addition to raw data and processed data products, BLUE CONNECT will produce tools, models, reports, and scientific publications. These will be shared as widely as possible with other projects and initiatives (WP1, WP7) and with Demo Site stakeholders and communities (WP4, WP7) in line with Open Science and FAIR principles.

6.1. Physical samples

Samples will be preserved at least 5 years after publication of data and termination of the project.

6.2. Reports and publications

Reports and publications will be made available to the public through the project's website (for non-peer-reviewed literature) or open access facilities (for peer-reviewed literature). BLUE CONNECT conforms with Horizon Europe requirements to make all peer-reviewed scientific publications generated by the project open access immediately after publication. Publications will either (1, preferably) be published in a Gold Open Access journal, or (2) be deposited as a preprint or postprint in a trusted repository for scientific publications (e.g., ResearchGate, Zenodo, institutional repository) in machine-readable format and be made open access (i.e., Green Open Access). Only publication fees in full open access venues for peer-reviewed scientific publications are eligible for reimbursement.

6.3. Code

BLUE CONNECT aims to share all open-source scripts and code on a dedicated GitHub page. It is the responsibility of the creator of the code to document the scripts and make sure they are easy to understand and follow. The MIT software licence is recommended because it is a simple permissive licence with conditions only requiring preservation of copyright and licence notices.

6.4. Models and tools

WP2 will develop new ecological models and tools tailored to address the needs of the Demo Sites. These will be integrated into a holistic framework with existing models and tools and made available through the Blueprint and the DTO.

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

7. Allocation of resources

The short- and long-term costs and maintenance of MDA and IMIS will be covered by VLIZ, the provider of these services. Data management personnel costs are included in the BLUE CONNECT budget under T1.2. Publication costs for Gold Open Access journals are also included in the budget. The Data Centre at VLIZ, supported by the IT Department, will coordinate and lead the data management activities, providing the necessary expertise and capacity to ensure proper data management. The project team has a designated Data Management Officer (VLIZ), whose main tasks are to support the partners engaged in data acquisition with tools for data transformation, quality control and upload. The Data Management Officer is responsible for the implementation of the DMP, the coordination of data evaluation, and the overview of data management practices within the data life cycle frame. The Data Management Officer will also update the DMP throughout the project and contribute to the final DMP (D1.4) at the end of the project.

8. Data security

During the project, raw data (pseudonymised or anonymised where appropriate), data summaries/analyses, and internal outputs and drafts will be shared between partners via a dedicated SharePoint folder with subfolders for each WP and task. Access to this folder is granted only to project partners by the coordination team led by SUBMARINER. No personal or sensitive data should be uploaded to the SharePoint folder; datasets containing such data should remain on the institutional server of the data collector and be pseudonymised before being uploaded to the SharePoint or shared with other project partners.

For long-term archiving, data will only be stored in trusted repositories which have provisions in place for data security. The metadata and data will be archived on IMIS and the MDA respectively, hosted by VLIZ, where there are authentication and authorisation procedures for giving editing rights to certain datasets. This procedure ensures long-term preservation of the data; if VLIZ ceases to exist or terminates its data-archiving activities, VLIZ will transfer the data files to a similar organisation that will continue the agreement with the data provider under similar conditions if possible. If VLIZ fails to find a similar organisation, VLIZ will transfer all files to each intellectual property right holder or data provider. Some BLUE CONNECT data may also be replicated to other repositories (e.g., OBIS, EMODnet) and relevant data and models will ultimately be integrated into the DTO, further assuring long-term data accessibility and resilience to possible fault occurrences affecting one of the data systems where data are stored.

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

9. Ethics

BLUE CONNECT will follow existing national, EU, and international guidelines and frameworks on ethics. Since these are explored in detail in D1.1 (Strategy for Addressing the Ethics and IP in the Project), this section focuses on ethical issues related to data collection and management.

Data providers are required to be compliant with the fundamental ethical principles required of all research activities carried out under the Horizon Europe Framework Programme, detailed in Regulation 2021/695, Articles 18 and 19. When collecting, (re)using, and sharing data, BLUE CONNECT will therefore comply with the principle of proportionality, the right to privacy, the right to protection of personal data, the right to physical and mental integrity of a person, the right to non-discrimination, and the need to ensure high levels of human health protection. Most relevant to this DMP is the protection of personal data; data from interviews, surveys and workshops will therefore be protected and stored securely. The research carried out in BLUE CONNECT will not involve the use of genetic resources or traditional knowledge associated with genetic resources, so the project will not fall within the scope of the EU Access and Benefit Sharing (ABS) Regulation.

9.1. Personal Data

BLUE CONNECT will comply with the European General Data Protection Regulation (GDPR) regarding the protection of personal data. BLUE CONNECT will collect personal data from those interacting with the project through questionnaires, workshops, interviews, demo sessions, observations, or the newsletter, always in compliance with the GDPR. Consent will always be obtained before collecting personal data using an Informed Consent Form (D1.1, Annex 1), in which data subjects will be informed in writing about the purpose of the research, their role in the project and how collected data will be used. Data subjects will also be informed that participation in BLUE CONNECT activities will always be entirely voluntary and they can withdraw their consent to participate at any time via a written statement. Informed consent will be sought from each individual outside of the consortium participating in the project activities and/or each individual whose personal data is to be collected in the project (i.e. for the purpose of the newsletter, public surveys, interviews, questionnaires, training sessions and webinars). The current version of the Informed Consent Form can be adapted to meet the requirements of particular activities, as long as it allows the data collector to gather informed consent.

Personal data will also be collected related to certain groups participating or interacting with BLUE CONNECT, namely project partners and AB members (WP1), local stakeholders

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

(LSWGs) at DS (WP4 & WP7), and newsletter subscribers and contact points for other Mission Ocean projects (WP7). These data will be maintained by the partners leading the associated tasks (i.e. SUBMARINER for WP1 and WP7, Demo Site lead partners for LSWGs), and will not be shared within the consortium. BLUE CONNECT will use a secure management system server (SharePoint) with restricted access for internal documents with access rights, for instance LSWG Informed Consent Forms.

Certain data protection measures must be followed when collecting personal data:

- **Data minimisation**: the collection of personal information should be limited to what is directly relevant and necessary for the project's activities. The data should be retained only for as long as is necessary to fulfil that purpose.
- **Anonymisation**: this technique strips all personal data out of a dataset to the point that the subjects are no longer identifiable, so the data are no longer recognised as personal data under GDPR. All BLUE CONNECT datasets containing personal data should be anonymised before making them publicly available. Audio and video data (e.g., interviews) should be transcribed and then fully anonymised. In some cases, techniques such as aggregation may be appropriate, for example if the subjects remain identifiable even after all personal data has been removed. Aggregated data is displayed in ranges or categories (e.g., age 20-30 instead of a specific age), omitting individual responses. Since some information is lost from aggregated data, this technique should only be used when necessary.
- **Pseudonymisation**: this technique removes all directly identifiable information from the data (e.g., name, address, contact details) while leaving indirect identifiers (e.g., age, gender, personal opinions). A unique, random code is then assigned to each data subject, and the links between codes and data subjects are recorded in a separate encrypted file accessible only by the data collector. Since pseudonymised data still qualify as personal data under GDPR, this technique should be used when sharing data between project partners, but not when making the data publicly available.

9.2. Special categories of personal data

Under GDPR (Art. 9), there are special regulations for several categories of personal data: personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health, and data concerning a natural person's sex life or sexual orientation. Such data can only be processed if the data subject has given explicit consent to the processing of those personal data. They must

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

therefore only be collected if necessary, and the data subject must give their informed consent.

In BLUE CONNECT, data on stakeholders' political opinions may be collected by WP3 to analyse barriers and enablers to effective MPAs, and power structures, distributional issues, and agency in the Demo Sites. Project partners collecting such data should store them securely using their own institution's facilities and only share them with other project partners via the BLUE CONNECT SharePoint folder in pseudonymised form. These data should be anonymised (and aggregated/summarised if necessary, e.g., as interview summaries) before being shared externally. Data belonging to the other special categories listed in the previous paragraph are unlikely to be collected in BLUE CONNECT. Data collectors are responsible for assessing whether their data belong to any of the special categories outlined in this section and for following the appropriate procedures.

9.3. Internal review of ethical issues

The ethical dimensions of the project's actions and outputs will be reviewed by the Ethical Advisory Board (EAB) as detailed in D1.1. The EAB will ensure adherence to the ethical principles in line with the Nuremberg Code, European Textbook on Ethics in Research, and EC Ethics Appraisal Procedure, and will monitor ethical features throughout the project such as the 'do no significant harm' principle and Intellectual Property Rights (IPR) protection. Considerations will be made on how to balance IPR and Open Science regarding the exploitation and dissemination of results and the application of IPR in relation to the FAIR principles. The EAB will also provide advice in case any ethical issues related to data management arise. The implementation of the strategy for addressing ethics in the project (D1.1) will be monitored and D1.1 will be updated accordingly, if necessary, in M12, M24 and M34 of the project.

10. Other issues

National legislation will be followed regarding sampling procedures in protected areas. National and European data protection regulations will be followed. Questionnaires and interviews will be carried out with full ethical approval of the partner and all details will be recorded. Some activities will be carried out in non-EU countries (namely Cabo Verde, Norway, and the United Kingdom); no local resources from these countries will be used, and no non-data materials will be transported from the EU to these countries or vice-versa. However, personal data may be collected during the project's activities in these non-EU countries, namely interviews, surveys and questionnaires. If this is the case, all data collection and protection procedures must be in accordance with (1.) the laws of the country in which the data was collected, and (2.) EU law. The GDPR applies to all data-processing procedures carried out by BLUE CONNECT, both within and outside of the EU.

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759

Any partner(s) who collect and/or process data in non-EU countries should determine what legal obligations apply based on local and EU laws and take whatever action is necessary to comply with them. Partners must also be able to demonstrate compliance upon request.

When personal data collected from non-EU countries is exported to the EU, research participants must understand and provide their consent by signing the Informed Consent Form (see D1.1). In this case, the data protection measures outlined in D1.1 and section 9 of this deliverable will be implemented, including pseudonymisation and anonymisation where appropriate. To ensure that personal data are transferred securely, appropriate organisational and technical measures should be taken.

11. References

- Abramic, A., Norton, D., Sarretta, A., Menegon, S., Katsika, M., Gekas, V., Rybka, K., Fernández-Palacios, Y. (2023). Maritime Spatial Planning Data Framework (MSPdF). How to structure input data for MSP process, monitoring & evaluation. Produced by Technical Expert Group (TEG) on Data for MSP. Supported by CINEA and DG MARE (EC). 45p. https://doi.org/10.2926/440667.
- General Data Protection Regulation (Regulation (EU) 2016/679). https://gdpr-info.eu
- Horizon Europe Data Management Plan Template (2022). V1.1 https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm#A1-template
- Leonelli S. (2015). What Counts as Scientific Data? A Relational Framework. *Philosophy of science*, *82*(5), 810–821. https://doi.org/10.1086/684083.

This project has received funding from the European Union's HORIZON Innovation Actions under Grant Agreement No 1011656759