BLUE MISSION BANOS

Supporting the Mission Ocean Lighthouse in the Baltic and North Sea Basins

Deliverable 1.5
Draft Knowledge and Data Management Plan

Grant Agreement number	**101093845**
Project title	BlueMissionBANOS – Supporting the Mission Ocean Lighthouse in the Baltic and North Sea Basin
Deliverable title	Draft Knowledge and Data Management Plan
Deliverable number	1.5
Deliverable version	1
Contractual date of delivery	31.05.2023
Actual date of delivery	31.05.2023
Document status	Final
Document version	Version 1
Online access	Yes
Diffusion	Public
Nature of deliverable	Public
Work Package	Nr.1
Partner responsible	University of Southern Denmark (SDU)
Contributing Partners	All
Author(s)	Maéva Vignes, SDU
Project Officer	Daniele De Bernardi
Editor	Alberto Terenzi, SUB
Approved by	Alberto Terenzi, SUB
Keywords	BlueMissionBANOS, Mission Ocean, Data Management Plan, FAIR, Ethics

HISTORY OF CHANGES		
Version	**Publication date**	**Changes**
1.0	31.05.2023	Final draft version

TABLE OF CONTENT

ACRONYMS

BMB	BlueMissionBANOS - Supporting the Mission Ocean Lighthouse in the Baltic and North Sea Basin
PP	Project Partners
WP	Work Package
KDMP	Knowledge and Data management Plan
FAIR	Findable, accessible, interoperable, re-usable
SDU	University of Southern Denmark
PREP4BLUE	Overarching CSA in the Mission Ocean, project Grant No. 101056957
MIP Ocean	Mission Implementation Platform of the Mission Ocean
SUB	SUBMARINER Network for Blue Growth EEIG, coordinator of BlueMissionBANOS

EXECUTIVE SUMMARY

The objective of D1.5 is to develop the first draft of a Knowledge and Data Management Plan (KDMP) to share with all project partners. This draft has been developed using a free Data Management Template from DMPonline (a tool for writing data management plans, https://dmponline.deic.dk/) and with support from the Data Management team at the University of Southern Denmark.

Following EU policies and best practices, this KDMP includes:

a. information on handling research data during/after the project

b. data types that will be stored and shared, and

c. explicit statement whether data will be shared or made open access.

During the BMB project, partners will strategically map and collect data from EU, national, regional and local open-access public databases or share their internal data. The data to be collected covers a broad spectrum: project information, funding opportunities, policies, stakeholders, citizen science activities, and relevant private and public services and products. Each data point collected is integral to our project objectives, and we ensure these diverse datasets will be appropriately integrated, managed and protected in a centralised database.

Our project is firmly grounded in **FAIR data principles** (Findability, Accessibility, Interoperability, and Reusability) and strongly emphasises appropriate intellectual property rights (IPR) management.

   Findability: Each dataset will be systematically indexed with unique and persistent identifiers, making it easily discoverable to relevant parties. Rich metadata will be created and associated with each data record to improve its findability.

   Accessibility: We are committed to open data principles and will strive to provide access to data and metadata, conditional on appropriate security and privacy regulations. Our data will be accessible via a well-documented and standardised API, ensuring users can easily access and understand the data. Where restrictions apply, a comprehensive justification will be provided.

   Interoperability: To promote integration and use of our data with other datasets, we will adhere to recognised data standards, formats, and vocabularies where possible. This approach will ensure our data can be combined with other datasets, thereby increasing its value and potential for synergy.

   Reusability: Our data will be accompanied by clear usage licenses (by default Creative Commons CC BY-NC), thus providing clarity on the conditions of reuse. Metadata will include sufficient information to allow future users to understand the original study and replicate the methods if needed.

In terms of **IPR Management**, our project acknowledges the necessity of appropriate protections to safeguard the interests of all stakeholders. We will ensure compliance with all relevant IPR regulations, taking into account ownership, licensing agreements, patents, copyrights, privacy concerns and security. We will identify, assess, and manage IPR issues throughout the project's life

cycle, ensuring all involved parties know their rights and obligations and follow the policies reflected in Annex 5 of the Grant Agreement Model (AGA).

This version of the KDMP will be shared for feedback/input by partners and managed by WP6 task leaders, with regular reviews by the BMB WP leaders in line with new data collection or changes in regulations and best practices. It is further expected to include multiple versions as the project grows (minimum during periodic assessments of the project) to ensure adaptability where warranted (https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm).

1. KNOWLEDGE AND DATA MANAGEMENT PRINCIPLES

The BMB knowledge and data management principles follow the data management principles and guidelines developed in the context of the Mission Ocean CSA PREP4BLUE (Grant No. 101056957) by ERINN Innovation. This document refers to the guidelines, policies and best practices presented in Deliverable 1.3 of BMB. The most important principles for data collection and management are reported below.

The BMB Data Management Plan (DMP) aims to provide a strategy for managing background/ data generated and collected during the project and optimise access to and re-use of research data. It is intended to be a 'living' document that will outline how the BMB research data will be handled during and after the project, and so it will be reviewed and updated at regular intervals.

The DMP describes the Research Output Management (ROM) life cycle for datasets based on results from BMB. It covers:

•   Guiding principles around data management following EC Horizon Europe requirements.

•   How to make BMB data FAIR: Findable, Accessible, Interoperable and Re-usable.

•   Data management cost and allocation of resources.

•   Data security and ethics, and confidentiality.

BMB will generate diverse outputs, including databases, innovation outputs, new methodologies, data, protocols, experimental approaches, and strategies. A structured and efficient Research Output Management is crucial to integrate the work performed within the different work packages and themes and align the scientific tasks with the overarching strategy of the BMB project.

The DMP also includes the following annexes, which will also evolve over the course of the project:

• Annex 1 consists of valuable resources relevant to data management to support partners in making their research and data openly accessible in the context of Horizon Europe.

• Annex 2 will be added to the DMP as the project is implemented, will consist of the BMB Inventory of Datasets, and will be generated as and when data is generated and managed.

1.1. KNOWLEDGE AND DATA MANAGEMENT GUIDING PRINCIPLES

The Data Management Plan (DMP) of BMB is coordinated under WP1 and WP6 jointly and is articulated around the following key points:

a) This DMP follows the definition of the **obligations/mandatory practices in the AGA Article 17 – Annex 5**)**.** The elaboration of the DMP based on these definitions will allow BMB partners to address all IP protection and data issues in line with the obligations/mandatory practices. The DMP will evolve with the project and be updated regularly as and when significant changes arise.
b) The Consortium will **comply with the requirements of Regulation (EU) 2016/679 and of the Council of 27 April 2016, as well as GA Articles 14 and 15** on the protection of natural persons

with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

Type of data, storage, confidentiality, ownership, management of intellectual property and access: Procedures surrounding data collection, storage, access, sharing policies, protection, retention and destruction are in line with EU standards as described in the BMB Grant Agreement (GA) and the Consortium Agreement (CA), particularly GA Article 20.1 Keeping records and supporting documents; GA Article 16 Intellectual Property Rights; GA Article 16.1 Background and access rights to the background, especially GA Article 16.2 Ownership of results and GA Article 16.4 Specific rules on IP, results and background; CA Article 8 Results; CA Article 9 Access Rights; CA Article 10 on Confidentiality; and GA Annex I – "Description of the Action".

1.2. BMB RESEARCH OUTPUT MANAGEMENT POLICY

As outlined in the executive summary, the collection, organisation, and formatting of data will be the responsibility of the relevant data owner. Each partner is responsible for their records and documentation in relation to the data generated, which must be in line with the accepted standards in the respective field, overseen by Task Leaders. To avoid losses, partners must take measures to ensure that data are backed up using reliable methods.

The right of the data owner (members of the research team) to use research data is reserved when providing open access. Right to use, here, refers to the right of the data owner to execute the original project plan before opening the data for further use.

All results produced during the project will be assessed for the need for IPR protection by the coordinating team in consultation with the partners concerned. The coordinating team is also responsible for strategic issues, ethics, and the exploitation of results, to ensure mutual understanding and facilitate creative collaboration.

All data underlying peer-reviewed scientific publications must be made openly accessible within and beyond the Consortium and uploaded to an open-access repository complying with the FAIR principles within one month of publication unless it has been evaluated as confidential. The collection and upload of data and underlying publications to the repository is the responsibility of the data owners.

If data owners wish to restrict access to data that underlies publications, they must provide to the coordinator a 'justifiable' reason for doing so as soon as possible and provide evidence of such reason upon prior notice **one month before publication** of the associated peer-reviewed scientific publication. For a list of 'justifiable' reasons under EC rules, please see section 2.1.3.2 below.

For each dataset collected, processed and/or generated in the project, the following elements will be documented and reported in the next DMP):

- **Data owner** – Description of the data owner, including the project partner name, originating work package, task and activity, the responsible researcher(s)' name and the primary contact details for enquiries regarding the data.
- **Data Summary**– A description of the data and an overview of how it is being captured and stored, including the name and description of the dataset, how it is being created and captured, the type and format of the data, the expected overall storage size of the data and whether an IP evaluation is needed.

- **Findability** – Description of domain-relevant repositories, whether the data will be made identifiable by a standard identification mechanism and the type of metadata that will be provided.
- **Accessibility** – Confirmation of whether all data is accessible and the methods or software tools needed to access the data. In cases where it is not accessible, a rationale for keeping it restricted must be provided.
- **Interoperability** – Description of whether the data is interoperable and the standard data/metadata vocabularies/ontologies relevant to it.
- **Reusability** – Description of the data licensing, any limits to the re-use of the data and the date (where applicable) the data will be made available for re-use.
- **Allocated resources** – Description of the estimated costs required to make the data FAIR and how these costs will be covered (e.g., covered by the WP budget).
- **Security** – Brief description of the data security measures in place, including confirmation of a plan for recovery, secure storage and protection over the transfer of sensitive data.
- **Ethics** – Any potential ethical issues must be noted.

A detailed description of the type and format of BMB data that will be generated and collected in the BMB Data Plan Inventory will be shared on the BMB project Microsoft TEAMS group. The inventory will be shared in the next update of the DMP in M18. Additional datasets may be identified and added to future versions of the DMP as necessary.

1.3. KEY PRINCIPLES AND PROTOCOLS (ALIGNED WITH PREP4BLUE PRINCIPLES AND PROTOCOLS)

The PREP4BLUE DMP is to be used as a reference document by partners to ensure compliance with both internal project and external EU rules and regulations. This section will first provide a clear and easily accessible summary of the **Key Principles** and **Protocols** all partners must follow. The following sections provide further context and detailed explanations outlining the necessity of the established protocols.

Every partner is responsible for ensuring they understand and follow the protocols established in this document. Protocols have been summarised under the following headings:

- Data Management
- Ensuring Data Findability
- Ensuring Data Accessibility and Reusability

The DMP operates as a functional manual and will be updated over the course of the project whenever significant changes arise, such as:

- New types of data
- Changes in consortium policies
- Changes in consortium composition and external factors

As new data are generated, they will be logged in the format of the Data Inventory Table in the BMB TEAMS by the data owner.

KEY PRINCIPLES & **PROTOCOL - Data Management**

- The **organisation and formatting of data collected will be the responsibility of the relevant data owner**, which must be in line with the accepted standards in the respective field, overseen by Task Leaders. This includes formatting and metadata. BMB aims to automate data ingestion and FAIR (see below), which should be the focus for all data owners.
- **Each partner is responsible for ensuring the security of their datasets**. To avoid losses, partners must take measures to ensure that data are backed up using reliable methods. Partners are advised to consult with their organisation's IT professionals to set up and manage data security and ensure the right safeguards are in place.
- From the start of each relevant task, project partners should consider identifying the most suitable file types and structures to be used to support external interoperability once ready and published where appropriate.

Best practices indicated hereafter should be followed by individuals involved in the project's research outputs and generating data.

1. When relevant, register at ORCID: http://orcid.org, which provides a consistent identity for individuals.
2. Follow prior notice procedures as outlined in sections 8 and 9 in the BMB CA.
3. Ensure that EC funding is acknowledged, including the project name and GA number (see section 4.1 and the PEDC for official text to use or contact SUB).
4. Ensure that peer-reviewed scientific publications based on BMB results are published in Open Access (e.g., Gold or Green).
5. Submit data underlying the published peer-reviewed scientific publications to the Coordinating Team (SUB) for evaluation.
6. When submitting these data, if a partner already intends to protect the data (see section 2.1.3.2), they should communicate this to the Coordinating Team to ensure an optimum level of confidentiality is upheld from the earliest stage.
7. Once cleared by the Coordinating Team, or if no objection is raised within 10 working days of receipt, ensure that data underlying published peer-reviewed scientific publications are deposited in an appropriate open-access data repository no later than 1 month after publication.

PROTOCOL – Findability

For each dataset collected or generated through BMB, partners should consider the following **Open-Access (OA) repository requirements** (for *all* data approved for open-access):

1. Researchers have the authority to select their repository of choice. Where possible, it should be domain-specific. To facilitate linkages, the suggested repository is Zenodo, but if that is not applicable, please consider the following:
   a. The BMB Database, once running, may be used as the suggested OA repository for BMB, if appropriate. The data will be made available through the cloud services of the eScience Centre at SDU (UCloud, DEiC EOSC-Nordic initiatives).
   b. Use an external data archive or repository already established for your research domain to preserve the data according to recognised standards in your discipline. You may search for data repositories (see the non-exhaustive list of suggested repositories in Annex I.
   c. Or, if available, use an institutional research data repository or your research group's established data management facilities.

    d. Or use a cost-free data repository.

2. Ensure the chosen online repository facilitates the identification of data and refers to standard identification mechanisms (ideally persistent and unique identifiers such as Digital Object Identifiers (DOIs)).
3. The organisation and formatting of the data collected will be the responsibility of the relevant data owner.
4. Each data owner will be responsible for depositing relevant data in the appropriate repository.
5. Ensure that research outputs and datasets are cross-referencing each other (e.g. scientific publications and the data behind them).
6. Outline the discoverability of the data (give metadata provision). Each dataset owner will ensure that their chosen repositories disseminate (meta)data to OpenAIRE to maximise data sharing, finding and re-using research outputs from BMB.

PROTOCOL – <u>Accessibility and Reusability</u>
Prior notice and confidentiality review

- Although not obligatory, in the interest of FAIR data, BMB partners are encouraged to make any other project datasets that do not require protection – regardless of whether they are connected to a publication – available open-access.
- Information on all datasets must be entered into the Data Plan Inventory, which will be made available on the BMB TEAMS.
- Whenever information on a new dataset is uploaded to the BMB Microsoft TEAMS or a previously created dataset is changed in such a way as might alter its protection level, the partner primarily responsible for the dataset must notify the Project Coordination team.
  - If a partner intends to protect any data, they should communicate this to the Project Coordination team to ensure an optimum level of confidentiality is upheld from the earliest stage.
  - Any evidence of applications for protection and/or associated legal processes relating to said dataset should be sent to the Project Coordination team within six months of this notification.
- If no evidence of protection is provided, the Project Coordination team may request that data be made accessible.
- The Project Coordination team will provide feedback on the type of protection mechanism, if any, they believe should be applied to the data.
- If a dataset is deemed suitable for open access, the Project Coordination team will inform the data owner and request they submit information on their data to the Coordinating Team for evaluation. If cleared by the Coordinating Team, or if no objection is raised within 10 working days of receipt, the data owner is responsible for uploading it to an appropriate open-access repository within 30 days of Coordinating Team approval (see below for repository requirements).
- If a dataset is deemed to be unsuitable for open access, the data owner may still be required to place relevant metadata in a suitable repository, providing said metadata does not itself constitute a protection breach.

When considering the potential to make data open access, **partners are requested to review the BMB Consortium Agreement (CA).** This defines the main approach regarding the ownership, protection and access to key knowledge like Intellectual Property (IP) and data. This approach will allow the BMB partners, collectively and

individually, to pursue opportunities arising from the project's results. Some of the major aspects covered in the CA are briefly summarised below:

- **Confidentiality:** Each partner will treat information from other partners as confidential unless otherwise stated and not disclose it to third parties unless the information is publicly available.
- **Results-/ Background-/ Data- owners** will notify the partnership of their planned intent to upload datasets to open-access repositories following the same prior notice procedure as is set up for the publication of results (45 days).
- **Pre-existing Know-How**: Each partner is, and remains, the sole owner of its IPR over its Pre-existing Know-How. The partners have identified and listed in the CA the Pre-Existing Know-How over which they may grant access rights for the project. The partners agree that the Access Rights to the Pre-existing Know-How needed for carrying out their own work under the project shall be granted on a royalty-free basis.
- **Ownership and Protection of Results:** The ownership of results will belong to the partner(s) generating the results. Protection will be implemented appropriately. When the result is the outcome of work carried out by two or more partners, and their respective share of the work cannot be ascertained, joint ownership will be agreed between the partners as established in CA Article 8.2. If a partner wishes to assign any knowledge to a third party, they should do so while observing the conditions set out in the BMB GA, especially articles 8, 9, 10, and should inform the other partners and request their consent, which should not unreasonably be withheld.
- **Access Rights:** Partners grant each other royalty-free access rights to knowledge generated in the project and to the background knowledge they bring to the project to the extent needed to successfully perform the project tasks allocated to them (see CA article 9).
- **Patents:** Under Article 16.4 and Annex 5 of the GA, partners with knowledge suitable for patents are obliged to make applications for patents or similar forms of protection and shall supply details of such applications to the other partners. Information relating to patents that have been registered must be submitted under the 'IPR' section of the EU Funding and Tender Opportunities Portal.
- **Use and Dissemination:** If dissemination of knowledge does not adversely affect its protection or use and is subject to legitimate interests, the partners shall ensure further dissemination of their own knowledge as provided under the GA (see Article 17), which has been signed by all partners.

2. INITIAL DATA SUMMARY

The primary data we are collecting and structuring or anticipating to collect, enrich or generate are listed below:

- WP2 Information on governance structures, responsibilities and roles at sea-basin, national and regional levels.
- WP3 Information on projects, initiatives and actors with respect to Citizen Engagement and the Mission Ocean.
- WP4 Transformative solutions and their barriers and R&I needs to be scaled—successful methodologies to be transferred on local, national and EU levels.
- WP5 Data and Information related to Sustainable, carbon-neutral and circular Blue Economy and the progress towards the Mission Objectives and sub-objectives.
- WP6 Projects and Technical Services related to the Mission Ocean in the BANOS area, together with information on the participants of the projects and services providers (at the institutional and personal level).

2.1. WILL YOU RE-USE ANY EXISTING DATA, AND WHAT WILL IT BE RE-USED FOR?

SDU will reuse data from EU, national, regional or local open-access public repositories such as CORDIS Open, Horizon Dashboard & Results Platform (CINEA Project Portfolio), Kohesio, OpenAIRE, UN Decade of the Ocean, LIFE Project database and Interreg project databases (keep.eu), Maritime Data Hub, data from the Mission Ocean Charter (European Maritime Forum).

We will also reuse data from our partners when they are willing to share their internal databases. So far, the internal databases we are integrating are:

- The BMB coordinator SUB will share their internal database of institutional stakeholders (excluding personal contacts).
- VLIZ (PP) will share a dataset of citizen science activities linked to the Mission Ocean objectives collected in PREP4BLUE.
- The Blue Bio COFUND project partners participating in PREP4BLUE and BMB share their database of projects and stakeholders.
- The Mission Ocean secretariat shares updates of the Mission Ocean Charter data in a table format.

The list of data that will be re-used will increase following the evolution of the needs during the whole BMB project duration.

2.2. WHAT TYPES AND FORMATS OF DATA WILL THE PROJECT GENERATE OR REUSE?

Given the scope of the project and its objectives, we can expect to collect a diverse range of data types and formats. These can be broadly categorised as follows:

Project Data: This includes structured data related to various research projects, funding opportunities, and initiatives. These data may come in CSV, XLSX (Excel), JSON, or XML formats.

Policy Data: Information about policies related to ocean health, pollution control, and carbon emissions in the blue economy. These could be in text format (TXT, DOCX, PDF) and would need extensive metadata to ensure they are findable and usable.

Stakeholder Data: This could include data on individuals, organisations, and institutions involved in or impacted by the project. These may be stored as structured data formats (CSV, XLSX, JSON, or XML) but may also include unstructured data such as text or email correspondence (TXT, EML).

Citizen Science Activities: This could be a broad range of data types, including geospatial data (e.g., GIS files for mapping pollution sources), sensor data (CSV, XLSX), and potentially images or videos (JPEG, PNG, MP4) documenting relevant activities.

Services and Product Data: This could include structured data about various public and private services and products related to Mission Ocean's objectives. Formats might include CSV, XLSX, or JSON files.

Metadata: Crucially, each data type listed above will have accompanying metadata, providing information about the data's context, quality, condition, and characteristics. Metadata can be stored in various formats, such as XML or RDF, depending on the standard followed.

It's important to note that the data types and formats collected will depend on the data sources and the project's specific needs. A more detailed view of the data types and formats collected will be added in Annex II of this KDMP before the next periodic review (M18).

2.3. WHAT IS THE PURPOSE OF THE DATA GENERATION OR RE-USE AND ITS RELATION TO THE PROJECT OBJECTIVES?

The data collected will help achieve the project's aims and, ultimately, the Mission Ocean objectives and sub-objectives.

More precisely, for WP6, the rationale for building a digital catalogue of projects and technical services is: "The geographic and sectoral fragmentation of existing technical services and projects catalogues is a major hurdle to achieving the Missions' objectives. Information is available but scattered among regional/national innovation and knowledge hubs, described in specific languages, often using non-standard metadata systems which are difficult to access and unknown to end-users". Hence BMB will strive to integrate this fragmented information and enrich it to establish (i) a catalogue of existing services needed to deploy innovative solutions for decarbonisation and circularity of the blue economy sectors in the BANOS area; (ii) establish and curate a catalogue of recent (2014 –), ongoing and future (2022-2025) projects on national, regional, sea basin and European levels relevant to BANOS lighthouses – considering both public and private funders and including the results of the projects.

Building a digital catalogue of projects and technical services in the form of a database will enable direct online matchmaking between actors developing and deploying innovative solutions and service providers that could support their efforts through testing, validating, upscaling, reproducing

and customising. A dedicated and customised BMB web portal will be developed to make the catalogue accessible to end-users.

Data collection and integration is a continuous process during the whole BMB project duration. We will collect and list relevant projects and initiatives across the BANOS area in a semi-automated manner (i) through consultation with relevant stakeholders (ii) through automated data collection from existing databases like CORDIS Open, Horizon Dashboard & Results Platform, Kohesio, SUB's Actors Map & Catalogue, and the future MIP Ocean Platform. We will model data and use dedicated terminology (ontologies) to combine the content of different types, such as surveys, databases, web content, free text descriptions or interviews. Each data point/entry will be associated with metadata relevant to diverse end-users and sector perspectives and for generating sustainable and interoperable digital catalogues.

It is important to note that the BMB Database is developed jointly with the Mission Ocean Ecosystem database of PREP4BLUE[1] and the BlueMissionAA[2] Database. The web interfaces of these three projects will also be developed and deployed jointly. Once the database and web interface are released in their beta version (planned in November 2023 for the 1st BMB Mission Arena in Gothenburg), EcoDALLI and BlueMissionMed[3] will be able to access the platform and database as users. If they have databases/datasets they would like to integrate into our database, SDU will offer to integrate these databases.

2.4. WHAT IS THE EXPECTED SIZE OF THE DATA THAT YOU INTEND TO GENERATE OR REUSE?

The expected size of the BMB database will be in the range of several hundred Gigabytes to a few dozen of Terabytes. Structured tabular data, including descriptions of projects, products, or stakeholders' expertise, is typically in the range of Gigabytes when including several thousands of entries. We anticipate including pictures or videos to describe services or products that would then lead to a database reaching several Terabytes in size.

2.6. TO WHOM MIGHT YOUR DATA BE USEFUL ('DATA UTILITY') OUTSIDE THE PROJECT?

The data collected will be useful for BMB partners and all stakeholders interested or involved in the Mission Ocean. The type of stakeholders that we target we the BMB database and web interface are:

- EU level Mission supporters.
- Partners in relevant projects funded under the Mission and running parallel to BMB.

---

[1] https://prep4blue.eu/

[2] https://bluemissionaa.eu/

[3] https://bluemissionmed.eu/

- Transnational and national organisations involved in the governance of the Mission Ocean.
- Public and private funding programmes with synergies with Mission Ocean.
- Sub-regional policymakers interested in the Mission Ocean.
- Blue economy industry and supporting organisations or associations.
- Citizen engagement organisations or NGOs such as Aquaria/Science Museums or BlueSchools.
- Public and private research institutes and service providers contributing to the Mission Ocean objectives.
- Local, regional and national Media.

Different stakeholders will likely be interested in different ways to access the data. For example, we can anticipate that: the EU Mission ocean supporting services (including the MIP Ocean) will get access to raw data through APIs; R&I actors from academia or industry will be more interested in searching through the projects or services/products data to find out who is working on similar topics as they are working on; journalists from national or local media might be interested in taking snapshots of overviews of the projects or stakeholders involved in the Mission Ocean.

3. MAKING BMB DATA FAIR

3.1. MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

Identifiers

Persistent identifiers (PIDs) are important because they unambiguously identify data and facilitate data citation. BMB partners will select data repositories that assign a persistent identifier, e.g., a Digital Object Identifier (DOI). Partners will also be encouraged to register with ORCID to create a personal, professional ID.

Metadata

Metadata is data that provides information about data. The primary metadata categories for BMB research data are descriptive (title, abstract, author, keywords, technical characteristics), structural (pages, chapters, tables, pictograms, diagrams) and administrative (file type, permissions, and when and how the file was created). Metadata is as important as the data itself as it allows applications to categorise and archive datasets, facilitating users to search and find relevant datasets easily. For this reason, metadata produced in BMB will adhere to the OBPS recommendations on the use of common vocabularies.

OpenAIRE is a platform funded and supported by European Commission with the Mission to shift scholarly communication towards openness and transparency and facilitate innovative ways to communicate and monitor research. **Each dataset owner will ensure that their chosen repositories disseminate (meta)data to OpenAIRE to maximise data sharing, finding and re-using research outputs from BMB.** Any repository that is listed in this DMP is interoperable with OpenAIRE. Partners who wish to use a different repository may make an inquiry to the Coordinating

Team to check whether it disseminates to OpenAIRE. If it does not, the Coordinating Team may contact OpenAIRE to see whether it can be added.

As the project progresses and datasets are identified and collected, further information on standards specification for metadata creation will be outlined in subsequent versions of the KDMP. Information on naming conventions used, approach towards search keywords, approach for clear versioning, and making data findable and crosslinked will also be provided.

3.2. MAKING DATA ACCESSIBLE - REPOSITORY

BMB will develop its own database, the BMB Database, which will be made publicly available and promoted through the project website and a dedicated web platform. As well as this, BMB partners will be asked to select the most appropriate data repositories that facilitate the finding, accessing, re-using and inter-operating of data sets, which are the basic principles with which Horizon Europe projects must comply. Among the field-relevant repositories to which it is anticipated that BMB could contribute are INSPIRE, EMODNET and Copernicus. A full list of repositories is presented in Annex I of this document.

Partners are encouraged to consider the Registry of Research Data Repositories (re3data) and Directory of Open Access Repositories (OpenDOAR) for useful listings of repositories that might be suitable for BMB outputs. A BMB community page within Zenodo will also be evaluated as a suitable repository for all project-related Open Access publications and datasets.

In view of the synchronisation of the BMB Database with the MIP Ocean Platform, we have dedicated resources in BMB WP6 to develop APIs that will increase the data accessibility but also reusability and interoperability.

3.3. MAKING DATA INTEROPERABLE

Labelling of data: To make our data interoperable, reusable and cross-compatible, we will stratify and label data about services and projects using a controlled Mission Ocean Ontology (MOOnt). The MOOnt will ensure the integration of BMB data with the wider community of the sustainable and circular blue economy. The MOOnt will build on other EU projects (PREP4BLUE, iMarine, BlueBRIDGE EMODNet, EU Digital Twin Ocean) and use a standard 'linked open vocabularies' (e.g. MarineTLO, e-Government Core Vocabularies, EURIO, EuroSciVoc, other vocabularies from LOV and W3C Semantic Web resources). Each data entry of our database will be labelled using the MOOnt, with categories describing the user/person (researcher, lawyer…), organisation (academic, industry, governmental…), sector of activity (technology, economy, environmental, societal, policy, legal…) and impact (decarbonisation, climate, biodiversity, circularity, innovation, digital transformation…).

Inter-operability, synergies and sustainability: The use of this MOOnt will ensure interoperability with the existing Ocean and Water knowledge systems and, more broadly, with the EC digital knowledge systems (e.g. EMODNet, iMarine & BlueBRIDGE, Digital Twin Ocean, Horizon Result Platform and Kohesio). In addition to making data accessible, we will build robust and sustainable APIs for end-users to retrieve all our open data automatically and use it for their own purpose. We

will also do a regular dump of open data (aligned FAIR principles) available on Zotero and OpenAIRE. Through constant dialogue with the MIP, PREP4BLUE and other Mission Ocean services, we will align expectations, avoid duplication of work, exchange best practices and synchronise data whenever relevant or required. The web interfaces will be developed using open-source libraries. The code will be thoroughly annotated to ensure its re-use during and after the end of BMB, with SDU/SUB committed to maintaining the database and interfaces post-project subject to resources.

3.4. INCREASE DATA RE-USE

We have several strategies in place to ensure that data generated and collected is reusable not just by the project team but by other researchers, policymakers, and stakeholders who may benefit from our work.

Comprehensive Documentation: We will create comprehensive data documentation to help users understand, validate, and reproduce our analysis. This will include the following:

- README files, which will describe the data's purpose, structure, and contents.
- Methodological descriptions outlining how the data was collected, processed, and analysed.
- Details on variable names, definitions, possible values, and units of measurement.
- Details on data cleaning processes to assist with the replication of our methodology.
- Analysis scripts or code, if applicable, to provide complete transparency about how the data was manipulated and analysed.

All these elements will be stored together with the data. They will be accessible whenever data is accessed to ensure potential users have all the information they need to interpret and utilise the data effectively.

Data Availability: Our aim is to make our data as openly available as possible, facilitating its reuse and potential for impact. However, data will only be made publicly available after thoroughly assessing ethical, legal, and commercial constraints. Where full open access is not possible, we will provide as much access as possible while respecting these constraints.

Data Licensing: We plan to use Creative Commons licenses to clarify the data reuse terms. Whenever possible, our default choice of license will be CC BY-NC in line with the obligations set out in the Grant Agreement and reflecting our commitment to open science.

Data Usability by Third Parties: Our commitment to the FAIR principles ensures that the data produced during the project will be usable by third parties, even after the project's end. This will be achieved by ensuring the data is effectively stored, documented, and accessible. The use of standard formats, clear and comprehensive metadata, and careful documentation of methodologies will ensure that third parties can effectively understand and reuse the data. Wherever possible, data will be stored in an accessible data repository beyond the project's end to ensure its continued availability.

Data Provenance: The provenance of the data will be thoroughly documented to ensure transparency and traceability. This will include information about where the data originated (source),

who collected it when it was collected, any transformations it has undergone, and how it has been used throughout the project. Provenance will be documented using recognised standards to promote interoperability. Documenting provenance will provide important context, help establish the data's credibility, and support data reuse.

Data Quality Assurance: Ensuring high data quality is central to our project. We will implement rigorous data quality assurance processes, including:

- *Data Verification*: After data collection, we will carry out data verification to check for any inconsistencies or errors in the data.
- *Data Validation*: We will ensure that the data accurately represents the reality it is supposed to depict. This may include cross-checking with other sources or with subject-matter experts.
- *Data Cleaning*: Any errors detected will be addressed in a data cleaning process, with the original raw data preserved for transparency.
- *Routine Quality Checks*: Regular audits and reviews will be conducted to maintain data quality throughout the project.

4. OTHER PROJECT OUTPUTS

Beyond the data itself, which will be shared through an interactive web interface and dedicated APIs, the project will likely generate a range of other research outputs that can provide valuable insights and resources to stakeholders. These will include:

Analytical Reports: Detailed reports summarising findings, trends, and insights derived from the data. These can be shared as downloadable PDFs through the web interface or as interactive web pages allowing users to explore findings in more detail.

Data Visualisations: Graphs, charts, maps, and other visual representations can be embedded within the web interface to facilitate understanding of the data. Interactive data visualisations could also be provided, allowing users to explore the data more dynamically and engagingly.

Methodological Guides: Clear, step-by-step guides explaining how the data was collected, cleaned, analysed, and interpreted. These could also include directions on how to use the web interface effectively. These can be shared as downloadable documents or as part of the web interface's help section.

Training Materials: If the project involves novel methods or tools, training materials (such as tutorials, webinars, or workshops) could be produced to help stakeholders understand and utilise the database and user interface. These could be shared through the BMB TEAMS or on external platforms like YouTube.

Publications: If the project leads to academic papers or conference presentations, these can be shared through the BMB Website, CORDIS and OpenAIRE. BMB project partners are fully committed to Open Science. We will ensure that all publications, reports and public deliverables (when approved by the EC) are published in Gold and Green OA. All WP deliverables will be open access on the project website, incl. the interactive Mission BANOS web platform (WP6).

Software and Code: If the project develops new software, algorithms, or code, these will be shared on the GitHub platform after considering the IPR protection of involved partners. Clear documentation will accompany these outputs to facilitate their reuse.

Datasets: Beyond the database itself, curated datasets (i.e., subsets of the data that have been carefully cleaned, annotated, and organised for specific analyses) can be valuable outputs. These can be made available for download through the web platform and BMB website.

In terms of how these outputs should be shared, best practices suggest that they should be made as openly available as possible, following the FAIR principles and ensuring appropriate licensing and IPR management. Stakeholders should be informed about new outputs through updates on BMB TEAMS and project consortium meetings.

5. ALLOCATION OF RESOURCES

Costs related to open access of research data in Horizon Europe are eligible under the conditions defined in the BMB Grant Agreement (GA), in particular Article 6 – Eligible and Ineligible Costs, such as Article 6.2.C.3 – Other goods, works and services, but also other articles relevant for the cost category chosen. These include the costs of data deposit, long-term storage and cost in time and effort needed to prepare the data for sharing and preservation. Costs cannot be claimed retrospectively. Project partners will be responsible for including any relevant costs in their financial statements. SDU is anticipated to bear most of the data management, processing and storage costs and has a dedicated budget allocated for this purpose.

5. DATA SECURITY

Data associated with the project may be stored and shared via the BMB TEAMS platform between all Consortium members. Each partner may manage levels of access. If a more secure platform is needed, or if data is expected to remain within the private domain of a partner, this should be clearly communicated and explained. Transferring sensitive data is possible via several routes through SDU and should be explored, as needed, by the Consortium.

To store and process the data included in the BMB database, SDU uses an internal cloud service provided by the eScience Center at SDU. This cloud service is linked to the EOSC and DEIC Nordic initiatives and has been certified ISO 27001 since February 2020, following a formal evaluation by an accredited external auditor, DNV GL.

Partners must use multiple methods to back up and copy research data and protect data from unauthorised access or use, or disclosure. All partners should consult with their organisation's IT professionals to set up and manage data security and ensure the right safeguards are in place. Partners should also consider that data files may need encrypting before storage or transfer and take the appropriate steps to ensure their security.

Data uploaded to Zenodo.org will remain available for long-term preservation. Long-term data management primarily addresses security in relation to where the data will be kept and how it will be

Funded by the European Union under Grant Agreement ID 101093845. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

20

accessible after the research project is complete. Partners must ensure that data are stored in certified repositories for long-term (a minimum of five years after the end of the project) preservation and curation.

6. ETHICS

The coordinator and WP leaders ensure compliance with fundamental ethical principles. Ethics aspects will be monitored throughout the project implementation, incl. the "do no significant harm" principle and Intellectual Property (IP) rights protection.

Data from questionnaires, briefing materials, and the project deliverables will be reviewed to ensure adherence to ethical principles in line with the Nuremberg Code, European Textbook on Ethics in Research, and EC Ethics Appraisal Procedure. BMB develops a secured Stakeholder Repository with individual access rights and assigns Stakeholder Interlocutors to ensure clear responsibilities for handling and protecting personal data and adherence to GDPR and other relevant laws. An Informed Consent form will be developed and used to inform and obtain consent from each person participating in the project stakeholder activities and each person whose personal data is to be collected in the project (i.e., in case of interviews, interactive events, recorded trainings and webinars, surveys).

BMB uses the Microsoft TEAMS secure management system as an online repository for internal documents with individual access rights. In general, the coordinator and WP leads will consider any ethical questions arising to ensure compliance with ethical principles and applicable international, EU and national law in the implementation of the research activities. Should any project activities raise ethical concerns, these will be handled following rigorously the recommendations provided in the European Commission Ethics Self-Assessment Guidelines. Should ethical issues arise unexpectedly during the project, the coordinator will provide the information as requested via the Funding and Tenders Portal on how the Consortium intends to handle it.

Oceanographic specific best practices
- https://www.oceanbestpractices.org/
- https://www.bodc.ac.uk/resources/vocabularies/

Open Science in Horizon Europe
- Open Science: https://openscience.eu
- Open Research Europe: https://open-research-europe.ec.europa.eu/

FAIR Findable (repositories)
- Directory of Open-Access Repositories: http://www.opendoar.org/
- Registry of Research Data Repositories: https://www.re3data.org/
- ZENODO Open-Access Data Repository: https://zenodo.org/
- INSPIRE: https://inspire.ec.europa.eu/
- EMODNET: https://emodnet.eu/en
- Copernicus: https://www.copernicus.eu/en/library
- Horizon Results Platform (HRP) Horizon Results Platform (europa.eu)
- The official portal for European data: https://data.europa.eu/en
- European Open science Cloud: https://eosc-portal.eu/

FAIR: Findable (Identifiers)
- Making Data 'Findable' using Persistent Identifiers: https://www.openaire.eu/how-to-make-your-data-fair
- Using Identifiers for Open Access- for Authors and Research Materials: http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?action=display&doc_id=4607

FAIR: Findable (Metadata)
- Oceanographic specific vocabularies for metadata: https://www.bodc.ac.uk/resources/vocabularies/
- Explanations of Scientific Metadata: http://www.dcc.ac.uk/resources/curation-reference-manual/chapters-production/scientific-metadata
- Metadata Standards Directory Working Group: http://rd-alliance.github.io/metadata-directory/
- Open Data and Metadata Standards: https://joinup.ec.europa.eu/sites/default/files/document/2015-05/d2.1.2_training_module_2.2_open_data_quality_v1.00_en.pdf
- Use of DataCite for metadata provisions: https://guidelines.openaire.eu/en/latest/data/use_of_datacite.html

FAIR: Interoperable
- OpenAIRE Guidelines for Literature Repositories, Data Archives, and CRIS Managers based on CERIF-XML: https://guidelines.openaire.eu/en/latest/

FAIR: Reusable (licensing)
- Creative Commons licensing : https://creativecommons.org/licenses/
- Advice on commercialisation of research data: https://eudat.eu/data-access-and-re-use
- Licensing Wizard: https://b2share.eudat.eu/
- IPR Helpdesk Advice on seeking IP Professionals: https://www.iprhelpdesk.eu/sites/default/files/documents/Guide-IP-professionals.pdf

- IPR Helpdesk Factsheet on IPR Valuation:
https://www.iprhelpdesk.eu/sites/default/files/newsdocuments/Fact-Sheet-IP-Valuation.pdf