ANNEX F - OPEN ACCESS AND FAIR DATA

TABLE OF CONTENTS

1. Context

The Sustainable Blue Economy Partnership supports a sustainable transformation of the blue economy which relies on wide and immediate access to quality data, across a range of marine and maritime disciplines and human activities. Better access to ocean data is essential in supporting innovation, advancing our understanding of marine ecosystems or also reducing costs and risks for maritime operators. More broadly speaking, it helps in delivering impact of the research activities, to feed-in decision-making processing, in enabling good governance, and also for Ocean Literacy.

In accordance with its ambitions for the blue economy, including the digital transition, the SBEP strongly supports Open Science principles, including the open sharing of research data to stimulate novel approaches to the collection, reuse, analysis, validation, and management of data and information.

Data generated by funded projects, including metadata and raw data, should be made publicly available, possibly after a short period of exclusivity. Sharing of data should be done:

- *in accordance with the FAIR principles for data management: Findable, Accessible, Interoperable, Reusable.*
- *with a balanced approach towards openness, according to the principle "as open as possible, as closed as necessary". Making data available prevails as long as there are no legitimate reasons to constrain access, such as risks to the privacy of individuals that may arise from personally identifiable data or the need to protect commercial potential.*

To support the funded projects in this process, a Data Management Plan (DMP) will be requested by SBEP from the development phase (pre and full-proposal) and to be updated throughout the lifecycle of the research project. Regional and national PFOs might also have specific requirements in terms of open access to data. Applicants are encouraged to plan resources in their proposal for the management of their research data and to comply with the requirements of their PFOs (if any).

In accordance with the principles of Open Science, funded projects must also comply with full and immediate open access for scientific publications. This compliance implies to publish in full and immediate open access journals, or full and immediate open access via a trusted open access repository (such as Open Research Europe). When using the repository option, partners should ensure that electronic copies of published articles are machine-readable and openly licensed,

including bibliographic metadata. Information should be provided about all the other scholarly objects, tools and instruments that are needed to validate the conclusions of the publication.

2. Data Management Plan (DMP)

The following section provides the core requirements in terms of DMP respectively at the stages of pre-proposals, full proposals, and awarded projects of SBEP. The set of questions in the DMP is adapted from the Practical Guide to the International Alignment of Research Data Management - Extended Edition of Science Europe. Applicants are invited to consult this practical guide which provides additional guidance and explanations to respond to the following core requirements.

a. Pre-proposals - Preliminary Data Management Information

In the data management section of the pre-proposal template, please address the following questions:

- *Who on your team will be responsible for developing, implementing, overseeing, and updating the data management plan?*
- *What data sets of long-term value do you expect that the project will produce? "Long-term" means those data sets that, over time, will or may be of value to others within your research community and/or the wider research and innovation community. Data of long-term value should meet the FAIR principles, i.e., they should be findable, accessible, interoperable, and reusable.*
- *How will you account for the costs required to manage the data and other materials to ensure long-term availability?*

b. Full proposals - Data Management Information

In the data management section of the full-proposal template (to be included in your single pdf to be uploaded on the EPSS), please address the following questions. Those that are repeated from the earlier stage should be elaborated on as appropriate:

- *What types of datasets of long-term value do you expect the project will produce or reuse? "Long-term" means those data that will or may be of value to others within your research community and/or the wider research, innovation, and stakeholder communities.*

- *How do you intend to ensure that the data from your project complies with the FAIR principles (for instance, in terms of financial and time resources)?*
- *Who will be responsible for developing, implementing, overseeing, and updating the DMP (role, position, and institution)? For collaborative projects, explain the coordination of data management responsibilities across partners.*
- *How do you intend to manage the data during the project and to ensure their long-term protection? For example, where will the data be held during the project, who will have access, and will a specialised data manager be part of the project team?*
- *How and by whom will the data be managed after the project ends to ensure their long-term accessibility? For example, will the outputs be published with a Persistent Unique and Resolvable Identifier (such as a Digital Object Identifier (DOI), Accession Number, Handle, etc.), and/or be placed in a recognised, trustworthy long-term domain or other repository or data centre. When will this occur? (Further information about repositories include, but is not limited to, the Re3data.org registry of research data repositories, CoreTrustSeal list of certified data repositories, etc.)*
- *What restrictions, if any, do you anticipate could be placed on how the data can be accessed, mined, or reused? Are there possible restrictions to data sharing, and embargo retention? Please explain the reason in case of restrictions (ex. intellectual property protection).*
- *What supporting documentation and other information (e.g., metadata) do you plan to make publicly accessible to support the longer-term re-use of the data and digital outputs?*
- *How do you account for the costs required to manage the data and digital outputs to ensure long-term accessibility?*

c. Awarded Projects - Full Data Management Plan

Awarded projects will be requested to provide a DMP at the beginning of their project. Updates will also be requested in their mid-term and final reports.

A full DMP for an awarded project is a living, actively updated document that describes the data management life cycle for the data and other digital outputs to be collected, reused, processed and/or generated. As part of making research data findable, accessible, interoperable, and reusable (FAIR), the DMP for a funded project should elaborate on the information provided at the Full proposal stage.

Applicants are advised to include the full costs of implementing the data management plan in their proposed project budget.

A session will be organised at the beginning of the funded projects during the kick-off meeting to exchange on best practices related to data management. For assistance in developing their DMPs, project leaders are encouraged to consult with relevant domain repositories, librarians, and information specialists at their respective institutions.

d. Research data and data infrastructures

Research data and digital outputs include, but are not limited to:

- *Data generated by research activities such as experiments, analysis, surveys, measurements, instrumentation and observations, video, audio, and computer simulations.*
- *All metadata describing the data and digital outputs, their acquisition (including model description and related metadata for simulations and workflows), and other details for the use and reuse of the data.*
- *Secondary data resulting from automated or manual data reduction, transformation, analyses, and results, together with the associated code, software, workflows, and provenance information.*
- *Stakeholder-oriented digital outputs such as maps (including GIS layers), decision support tools, tutorials, videos, local language resources, lesson plans, curricula, policy memos, and whitepapers.*
- *Descriptions of (and metadata relating to) physical samples connected with the call - but not the actual physical samples.*

Shared data must be discoverable through machine-readable catalogues, information systems and search engines. Sufficient metadata must be provided and made openly accessible to enable data (including models, workflows, software, methods, etc.) to be discoverable, accessible, understandable, interoperable, and effectively reused by others, including those outside the discipline of origin and the context of acquisition. Data must be curated, including maintaining integrity, quality, and veracity, using internationally or community-agreed standards and protocols. Data must be preserved, protected from loss, and remain accessible and usable for future research in sustainable and trustworthy repositories.

Resulting publications must list where or how to locate the underlying supporting data and other research materials, including agreed persistent identifiers, processing details and any workflows, software, and code. Academic journals may also set specific requirements for Data Accessibility

Statements to be included within published research results (primary research articles). Researchers should ensure that metadata and documentation created to support research datasets and other digital outputs retained for the long term is sufficient to allow other researchers a reasonable understanding and trust of those materials, thereby minimising unintentional misuse, misinterpretation, or confusion.

In the development of data infrastructures, it is important to leverage existing resources, platforms, standards, and recognised practices together with a clear sustainability plan. Projects that propose to develop data infrastructures are asked to work closely with and support relevant international networks, infrastructures, and standards organisations. They should make as much use as possible of existing certified domains, and national or international data repositories (for further information, possible resources include, but are not limited to re3data.org, Core Trust Seal, Group on Earth Observations (GEO), FAIRsharing.org, etc). Projects should also coordinate with, and make use of, the products and practices developed by recognised research and operational data policy and sharing organisations such as the Committee on Data for Science and Technology (CODATA), the Research Data Alliance (RDA), the ICSU-World Data System (WDS), and the European marine data networks and services: European Marine Observation and Data Network (EMODnet, please see below), Pan-European Infrastructure for Ocean & Marine Data Management (SeaDataNet) and Copernicus Marine Environment Monitoring Service (CMEMS). More information on these important components of Marine data management, along with other initiatives, are described in the European Open Science platform Blue-Cloud (a component of the European Open Science Cloud for the marine domain).

3. FAIR data and knowledge for the OceanEye and European Ocean Pact

The European Ocean Data Directive sets out that environmental observations are high value datasets and that Member States and other stakeholders have a responsibility to make these data accessible (either fully open or at least open source through metadata). In addition, the EU INSPIRE Directive provides a framework for geospatial data and metadata. In the marine domain, these are implemented by the EU marine data service EMODnet which provides a public service to Member States and the wider in situ ocean observation community to assemble and integrate all relevant in situ ocean observations into its data collections. EMODnet offers open access to more than 68 parameters of the marine environment and human activities at sea, harmonised at pan-European level and beyond, and with EMODnet experts producing data products e.g., the

EMODnet Bathymetry Digital Terrain Model and EUSeaMap. EMODnet offers guidelines and support to projects in preparing their data, and metadata describing the data collection e.g., owner/ data producer, location, parameter(s) and ocean observation device/platform, etc. Going forward, all SBEP projects are encouraged to contact the EMODnet Helpdesk (helpdesk@ec.emodnet.europa.eu) at the start/during the project to express their planned data collection (parameters, volume, geographical coverage), prepare their (meta)data according to established standards (EU INSPIRE, International ISO) to ensure they are interoperable with - and ready for integration into - EMODnet.

4. Links for further information and documentation

- Open Research Europe platform

https://open-research-europe.ec.europa.eu/

- Practical Guide to the International Alignment of Research Data Management - Extended Edition from Science Europe

https://scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management/

- Research Data Alliance (RDA)

https://www.rd-alliance.org/

- Blue Cloud platform

https://blue-cloud.org/

- European Marine Observation and Data Network

https://emodnet.ec.europa.eu/en

- SeaDataNet

https://www.seadatanet.org/

- Copernicus Marine Service

https://marine.copernicus.eu/