ABNJ GEF Deep Seas Project Component 4:

A review of data storage and sharing options for the South-East Pacific to support area-based planning in Areas Beyond National Jurisdiction

UN environment

COMMON OCEANS

gef

A review of data storage and sharing options for the South East Pacific to support area-based planning in Areas Beyond National Jurisdiction

Authors

Juliette Martin[a], Miles Macmillan-Lawler[b], Ruth Fletcher[a], Osgur McDermott Long[a], Naomi Kingston[a], Hannah Thomas[a]

[a] UN Environment World Conservation Monitoring Centre (UNEP-WCMC)
[b] GRID-Arendal

Citation

Martin, J.C.G., Macmillan-Lawler, M., Fletcher, R., McDermott Long, O., Kingston, N., Thomas, H. (2018). A review of data storage and sharing options for the South-East Pacific to support area-based planning in Areas Beyond National Jurisdiction. Cambridge (UK): UN Environment World Conservation Monitoring Centre. 16 pp.

Graphics

Cover illustration: Graphics used with permission from the FAO/UN Environment GEF ABNJ Deep Seas Project.

219 Huntington Road, Cambridge, CB3 0DL, UK

Telephone: +44 (0)1223 277314

Email: marine@unep-wcmc.org

Teaterplassen 3, Arendal, N 4836, Norway

Telephone: +47 47 64 45 55

Email: grid@grida.no

Partners involved in the GEF ABNJ Deep Seas Project are:

1. Outline and purpose of the report

Human activities occurring in Areas Beyond National Jurisdiction (ABNJ) are mainly managed through sectoral measures. Currently, there are limited mechanisms to bring the data and information required by these different sectoral management bodies together. Additionally, there are both legal and geographic gaps in the management of human activities in ABNJ. This has potential for management decisions made by one sector to adversely impact the management decisions of another sector. Therefore, cumulative impacts of human activities on vulnerable deep sea biodiversity may not be adequately captured under the current situation. One simple mechanism to help mitigate these impacts is to ensure that data used to support the management of human activities and on key environmental features are accessible to various sectoral management bodies.

Aim of the Report

In this report, we explore potential options for storing and sharing pre-existing data collated for the South-East Pacific as part of the GEF ABNJ Deep Seas Project ('the ABNJ Deep Seas Project') Component 4 on area-based planning in ABNJ. This Project is jointly implemented by FAO and UN Environment, the South-East Pacific being one of Component 4's two pilot study areas. However, the management of data is relevant across all the Components of the ABNJ Deep Seas Project.

We aim to identify both short-term (duration of the ABNJ Deep Seas Project) and long-term (post ABNJ Deep Seas Project) options for ensuring data collated as part of the pilot project in the South-East Pacific are stored centrally and shareable. It is important that there is sustainability of information collected as part of the project, ensuring it is available in the long term if needed. Any proposed data sharing option needs to align with the Permanent Commission for the South Pacific (CPPS) Regional Seas Programme strategy on data management and where possible build on other regional data management initiatives.

Methodology of Review

In this report, we highlight the needs of the ABNJ Deep Seas Project for the storage and sharing of pre-existing data collated in the South-East Pacific through a data portal. For the purpose of this report, the most basic definition of a *data portal* was used, namely an online list of datasets with associated guidance on how these datasets can be accessed, which may or may not include further functions such as data visualisation tools or detailed metadata. As a result, the portals that were included in this report vary in their nature and purpose, from geospatial data portals which are entirely customisable (e.g. UN Biodiversity Lab), to portals focused on only one type of data (e.g. the EBSA data portal) and bespoke portals that do not currently exist. The highlighted requirements aim to support the needs of the CPPS Secretariat and its member countries for data storage and sharing.

In this report, we created a set of criteria (detailed in Section 2) against which existing and new data management options were evaluated. Some of these options include currently available data portals in the region and worldwide. Finally, we provide recommendations for the centralised sharing of data collated under the ABNJ Deep Seas Project, both during the duration of the ABNJ Deep Seas Project and after its completion (long-term sustainability).

2.   Data management requirements

The ABNJ Deep Seas Project, through its pilot project in the South-East Pacific, is collating information on data (i.e. metadata) relevant to area-based planning in ABNJ, and may create new synthesis data products. There are existing data layers at a global scale with features within the South-East Pacific area. A metadata inventory collating information on these layers was created in order to increase transparency of the biodiversity within the region (Weatherdon et al., 2016). Follow up work is being done to identify what sectoral management data and biodiversity information are available at a regional and national scale in the South-East Pacific. The aim is to create a tailored inventory of this metadata information to support increased transparency.

It is important that data discovered or created as part of the project are available to the project partners in the long term. A number of data management requirements need to be fulfilled in order to make these data available to, and shareable by, sectoral organisations, including the CPPS Secretariat and its member countries.

This report is focused on options for storing and sharing data in the South-East Pacific, as well as metadata where they are available. Metadata are therefore described in more detail below to explain the distinction between data and metadata.

Metadata and data discoverability
Metadata are frequently defined as "data describing the context, content and structure of other data and their management through time". In simple terms, metadata are information about data. One approach to facilitate data discoverability is the application of a comprehensive and well-designed metadata standard to data falling under the remit of the South-East Pacific pilot region. Without the context provided by metadata, such as the date of creation, data are merely raw facts with limited usability.

Organisations may be hesitant to share what they view as sensitive and important data, which are often costly to collect. Sharing of metadata can overcome some of the issues associated with the sharing of the raw data, as sharing metadata makes it clear what data exist, what they can be used for and what the conditions of their use are. More information on metadata and their benefits can be found in Annex 1.

3. Criteria for assessing the data storage, management and sharing characteristics of available data portals

There are currently a number of existing and possible data storage and sharing solutions available to the ABNJ Deep Seas Project in the form of existing data portals. In order to assess the best option for ABNJ-relevant data storage and sharing, a comprehensive set of criteria was developed, each composed of several sub-criteria (Table 1). These criteria are based on the ABNJ Deep Seas Project and pilot region's needs, as well as best practices for user interface and data structure.

Table 1: List and description of the criteria used to assess different data management solutions

Criterion	Sub-criterion	Description
Mandate	Regional relevance	Supports the Regional Sea Programme's data management objectives.
	Geographic scope	Coverage of existing data and information, or geographic mandate relevant to the pilot region and/or ABNJ.
	Multi-sectoral	Stores data from or relevant to multiple sectors.
	Authoritative	Portal is seen as containing authoritative data or can harvest directly from the authoritative source. Additional documentation such as reports, related studies and other supporting non-spatial data are available with the data.
Sustainability	Sustainable	Meets the Global Environment Facility (GEF) requirements for long term data management. For example, there are sufficient funds and technical capacity to maintain the data management solution beyond the life of the project.
Availability	Available	The data management solution is available to the ABNJ Deep Seas Project for immediate use (i.e. it currently exists).

Criterion	Sub-criterion	Description
	Accessible	Is accessible to users from the region, including options for multiple languages.
Functionality	Ease of Use	Easy to use and requiring little technical capacity, i.e. easily navigable, intuitive and structured in such a way that users' technical knowledge or experience levels do not constrain their usage of the portal and its resources.
	Versatile	A range of data and metadata types can be stored.
	Discoverable	Metadata and/or factsheets are available alongside data in order to enhance the data discoverability and to guide the appropriate use of the datasets. Metadata should include agreed thematic keywords to aid in discovering the data.
Visualisation	Web mapping	Spatial data can be visualised interactively through online maps and, where terms and conditions for use allow, can be downloaded. Viewing spatial data online can be valuable and compelling. It can provide quick understanding of what the data cover and their potential to be used in a planning process. Therefore this functionality of a portal can be very useful, in particular for users who do not have spatial data software available to them.

4. Assessment of data portals of relevance to the South-East Pacific

The results of the assessment are presented in table 2, which shows whether the assessed data management option met a given criterion (denoted by "✓") or not (denoted by "x"). Based on these results, recommendations for the most appropriate options are made in section 5. Detailed descriptions of the listed portals are provided in Table 1 in Appendix 2.

Table 2: How well the assessed portals meet the selected set of criteria in the South-East Pacific

Criterion	Mandate				Sustainability	Availability		Functionality			Visualisation
Portal name and URL	Regional relevance	Geographic scope	Multi-sectoral	Authoritative	Sustainable	Available	Accessible	Ease of Use	Versatile	Discoverable	Web mapping
Bespoke ABNJ Project portal *	✓	✓	✓	✓	✓	x	✓	✓	✓	✓	✓
SPINCAM Atlas http://www.spincamnet.org/	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
UN Biodiversity Lab https://www.unbiodiversitylab.org/	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IW:LEARN http://geonode.iwlearn.org	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
PacGeo http://pacgeo.org/	x	✓	x	x	x	✓	✓	x	✓	✓	✓

Criterion	Mandate				Sustainability	Availability		Functionality			Visualisation
Portal name and URL	Regional relevance	Geographic scope	Multi-sectoral	Authoritative	Sustainable	Available	Accessible	Ease of Use	Versatile	Discoverable	Web mapping
EBSA data portal https://www.cbd.int/ebsa/	✓	✓	x	✓	✓	✓	✓	✓	x	✓	✓
SeaDataNet https://www.seadatanet.org/	x	✓	x	✓	✓	✓	✓	✓	✓	✓	✓
The Knowledge Network for Biocomplexity https://knb.ecoinformatics.org/	x	✓	✓	✓	✓	✓	✓	✓	✓	✓	x
Bespoke online search tool *	✓	✓	✓	✓	✓	x	✓	✓	x	x	x
Ocean+ Data http://wcmc.io/oceandata	x	✓	✓	✓	✓	✓	✓	✓	✓	✓	x

Denotes data portals which do not currently exist. As a result, the assessment was carried out based on the potential of such a portal to fulfil the included criteria.

5.   Recommendations

The global biodiversity data landscape is becoming increasingly complex, giving rise to a plethora of different data portals worldwide (Bingham *et al*., 2017). A number of these existing or potential data portal options have been assessed against a range of criteria to determine the best option for fulfilling the ABNJ Deep Seas Project aims of supporting conservation and sustainable fisheries management. The results provided in table 2 can be used to identify the most appropriate data storage and sharing options for data collated in the South-East Pacific and for long-term data management for the region.

A long term goal of any data storage and sharing option chosen for this project is to facilitate cross-sectoral planning by providing access to data from various marine sectors and industries. Recognising the difficulty in balancing short- and long-terms benefits the assessed options, the following suggestions are made, all of which are subject to funding:

Key Recommendations:

1. As a short term option, for the remaining duration of the ABNJ Deep Seas Project, the **IW:Learn** Platform was considered most appropriate for hosting ABNJ data due to its global accessibility, relevance to the GEF requirements, and immediate readiness.
2. An interim and complementary option is the development of a **bespoke online search tool** listing all relevant datasets and their metadata for the South-East Pacific ABNJ. This is in line with the metadata inventory work package under Component 4, undertaken by the CPPS with support from UNEP-WCMC.
3. As an immediately available alternative, the **Ocean+ Data platform** would allow regional metadata to be stored, filtered and downloaded. Yet, it does not provide map services to interactively view spatial data.
4. As a long-term option, post completion of the ABNJ Deep Seas Project, the **SPINCAM Atlas** would allow for the highest regional and ABNJ Project relevance. This Atlas is hosted by CPPS with technical support provided by IOC-UNESCO. Since 2017, CPPS is also an OBIS node, which demonstrates the commitment of members countries in their long-term vision regarding the storage and sharing of marine data generated in the region.

In the intricate global biodiversity informatics landscape, new data portals need to occupy a unique niche to avoid being redundant and duplicating effort. The approach of creating a new data portal was therefore not considered beneficial due to the multitude of existing portals that are suitable for hosting data under the ABNJ Deep Seas Project, and the resources needed to create a new portal.

6.   Next steps

Providing tools to empower people with access to relevant information is a complex task, which requires considerations that reach beyond the scope of this report. Such considerations include the political contexts of various data portals, their current data hosts and users, their user interface design, data organisation and their general navigability. Furthermore, various data management issues, such as appropriate data sharing, data licensing and potential use constrictions will also need to be considered in more detail for any data shared in the future. Some of these elements, in addition to more specific user needs, could be identified through an expert workshop in the pilot regions could provide the best impact for the Project. Nonetheless, this report offers initial recommendations and a first step towards the most appropriate data storage and sharing approaches for the ABNJ Deep Seas data in the South-East Pacific.

The identified data portal options will also inform and complement a capacity assessment undertaken in the South-East Pacific under the ABNJ Deep Seas Project (Macmillan-Lawler et al., in press), which identified access to marine biodiversity information as a key capacity element for the region.

To achieve effective area-based planning in ABNJ and be able to assess impacts on deep-sea ecosystems, stakeholders from different sectors need to be able to access and share data on the different human activities taking place in ABNJ. An important next step of this report is therefore an assessment of existing or potential data portals that would support the sharing of cross-sectoral data on a global scale, rather than only one region of the world. This would involve a similar scoping exercise of currently available portal options for hosting and sharing these types of data, as well as considerations of the type of data portal functionalities required to achieve this.

References

Bingham H, Doudin M, Weatherdon L, Despot-Belmonte K, Wetzel F, Groom Q, Lewis E, Regan E, Appeltans W, Güntsch A, Mergen P, Agosti D, Penev L, Hoffmann A, Saarenmaa H, Geller G, Kim K, Kim H, Archambeau A, Häuser C, Schmeller D, Geijzendorffer I, García Camacho A, Guerra C, Robertson T, Runnel V, Valland N, Martin C (2017) The Biodiversity Informatics Landscape: Elements, Connections and Opportunities. *Research Ideas and Outcomes* 3: e14059. https://doi.org/10.3897/rio.3.e14059

Macmillan-Lawler, M., Thomas, H., Fletcher, R., and Martin, J. (in press) Capacity assessment for area based planning in areas beyond national jurisdiction for the CPPS. Final Report 16pp.

UNEP-WCMC (In Prep) Guidance for the formation of a marine and coastal metadata standard: An application within Myanmar. Cambridge: UNEP-WCMC.

Weatherdon LV, Martin JCG, Fletcher R, Martin CS, Blyth S, Fletcher S (2016). Introduction to marine datasets of biodiversity importance in the South-East Pacific. Cambridge (UK): UN Environment World Conservation Monitoring Centre. 17 pp. (+ 3 annexes)

Annex 1: Benefits of metadata

Without the addition of metadata, "raw" data often lack the contextual information required to understand data origins and what types of analyses they can support. Well-constructed and maintained metadata can therefore improve understanding of datasets. It should be provided in a format that allows a dataset to be used, re-used, and integrated with other datasets and analyses, thus ensuring they appropriately inform decisions. All efforts to implement, organise and manage metadata can help to improve knowledge of an organisation's data holdings, and support operations and decision-making. Some common benefits of well-structured and informative metadata are:

- **Streamlining data discovery and data heritage**. Most data users and providers are increasingly having to manage systems in multiple locations and across various platforms. Through the effective management of metadata, an organisation can create an inventory of its data and maintain a lineage of changes made to different versions of the data throughout the lifecycle, while also providing information on the format and location of these data.
- **Reducing loss of institutional knowledge**. Often, institutional knowledge of data is held only by certain employees or actors. If these individuals leave, this knowledge can be lost. Metadata helps to preserve knowledge about the data, reducing the risk of losing important contextual information that can guide use.
- **Increasing the continuity, lifecycle and usefulness of data**. Missing or unavailable metadata may render data more difficult or impossible to use for a particular purpose(s) beyond its initial purpose.
- **Increasing confidence in data and data-driven analyses for end users**. Tracking data lineage gives context to users, which can help with identifying and correcting data errors and thereby increase the accuracy and reliability of outputs.
- **Increasing the value of data by facilitating reuse and sharing.** Streamlining data discovery can increase awareness of what data exist and how they may be used for multiple purposes. Through comprehensive metadata, data created by an organisation will be more accessible to multiple users. However, access to metadata does not always mean that the data described are accessible: the data manager still has control over data and their use.

Annex 2: Descriptions of data portals

Portal name	Description
Bespoke ABNJ Deep Seas Project portal	A new portal would be created especially for the ABNJ Deep Seas Project. It would therefore be entirely customised to the Project needs.
SPINCAM Atlas	The SPINCAM project ("Southeast Pacific data and information network in support to integrated coastal area management") was funded by the Flemish Government (Belgium). It has been developed under the coordination of IOC-UNESCO and CPPS. SPINCAM was designed to establish an integrated coastal area management (ICAM) indicator framework at national and regional level in the South-East Pacific, focusing on the state of the coastal and marine environment and socio-economic conditions. The SPINCAM Atlas provides stakeholders with information on the sustainability of existing and future coastal management practices and development.
UN Biodiversity Lab	The UN Biodiversity Lab is an online platform that allows policymakers and other partners to access global data layers, upload and manipulate their own datasets, and query multiple datasets to provide key information on the Aichi Biodiversity Targets and nature-based Sustainable Development Goals. Its partners include the GEF, CBD, UNDP, UN Environment, GRID-Geneva and UNEP-WCMC. The UN Biodiversity Lab data is powered by MapX, a UN-backed geospatial mapping software which collects the best available scientific data and acts as a trusted data broker to help people find spatial solutions to environmental and natural resource challenges.
IW:LEARN	The IW:LEARN Spatial Lab aims to provide a customizable mapping of the GEF International Waters (IW) portfolio and provide a visualization of studies, reports and intervention results on the various transboundary water basins pulled together with the help of the GEF IW projects, partner organizations and authoritative sources and other contributors. The IW:LEARN Spatial Lab aims to provide a central service for sharing and discovering geodata. The platform lets users upload, manage, and browse and visualize data, and export public data in reusable formats. IW:LEARN is executed by GRID-Arendal, and its implementing agencies are UN Environment and UNDP.
PacGeo	PacGeo is an open access geospatial data repository for the Pacific Region providing premier geophysical, geodetic, and marine spatial datasets. It was developed through collaboration between the GeoScience, Energy and Maritime Division of the Pacific Community (SPC), University of Sydney, Geoscience Australia (GA), and GRID-Arendal.
EBSA data portal	Displays EBSAs which are special areas in the ocean that serve important purposes, in one way or another, to support the healthy functioning of oceans and the many services that it provides.

SeaDataNet	SeaDataNet is a distributed Marine Data Infrastructure for the management of large and diverse sets of data deriving from in situ of the seas and oceans. Professional data centres, active in data collection, constitute a Pan-European network providing on-line integrated databases of standardized quality. The online access to in-situ data, metadata and products is provided through a unique portal interconnecting the interoperable node platforms constituted by the SeaDataNet data centres.
The Knowledge Network for Biocomplexity	The Knowledge Network for Biocomplexity (KNB) is an international repository intended to facilitate ecological and environmental research. It is an efficient way to share, discover, access and interpret complex ecological data. The data originate from a highly-distributed set of field stations, laboratories, research sites, and individual researchers.
Bespoke online search tool	This would be simple website on which datasets and metadata are listed and searchable by generating a query-based inventory of datasets in report format. Complementary to this, listed datasets could be integrated into one of the proposed data portals.
Ocean+ Data	Ocean+ Data provides an overview of global marine and coastal datasets of biodiversity importance, including some datasets of regional interest. This online library lists over 180 resources and directs users to their sources using up-to-date and standardised metadata sheets.

ABNJ DEEP SEAS PROJECT

The Sustainable Fisheries Management and Biodiversity Conservation of Deep Sea Living Resources in Areas Beyond National Jurisdiction Project (ABNJ Deep Seas Project for short) is a five year project supported by the Global Environment Facility, and implemented jointly by the Food and Agriculture Organization of the United Nations, and the United Nations Environment Programme. The UNEP project component is executed though the UNEP World Conservation and Monitoring Centre.

The Project is designed to enhance sustainability in the use of deep-sea living resources and biodiversity conservation in the ABNJ through the systematic application of an ecosystem approach. It brings together over 20 partners who work on deep-sea fisheries and conservation issues in the ABNJ globally. The partnership includes regional organizations responsible for the management of deep-sea fisheries, Regional Seas Programmes, the fishing industry and international organizations.

The Project aims to:

Strengthen policy and legal frameworks for sustainable fisheries and biodiversity conservation in the ABNJ deep seas;

Reduce adverse impacts on VMEs and enhanced conservation and management of components of EBSAs;

Improve planning and adaptive management for deep sea fisheries in ABNJ; and develop and test methods for area-based planning.

The ABNJ Deep Seas Project started in September 2015 and is one of four projects under the **GEF Common Oceans Programme**.

More information is available from **www.commonoceans.org**