Ocean Data and Information System - Concept Paper

By Tobias Spears, with contributions by Ms Pauline Simpson, Ms Cyndy Chandler, Mr Yutaka Michida and Mr Peter Pissierssens

Executive Summary
This paper describes a recommended strategy to move towards the implementation of a universal marine data and information system in response to the 2016 external audit of the IOC and its activities. After considering the observations presented in the audit, identifying the root causes which have contributed to the current state of the marine data and information systems landscape, it is recommended that the IOC work with existing stakeholders, linked and not linked to the IOC, to improve the accessibility and interoperability of existing data and information, and to contribute to the development of a global ocean data and information system, to be referred to as the Ocean Data and Information System in this document, levering established solutions.
# Table of contents

1. Introduction ............................................................................................................. 1
2. Scope of audit .......................................................................................................... 2
3. Key root causes underlying issues described in the IOC audit ................................. 7
   3.1 Engaging data and information providers and the role of the IOC ......................... 8
   3.2 Disparities between regions which impede participation ........................................ 8
   3.3 Competition (or perceived competition) between efforts which may or may not be linked to IOC ................................................................................................................. 9
   3.4 Diversity in content standards ................................................................................ 10
   3.5 Diversity in technology solutions ............................................................................ 12
4. Options to consider in response to the IOC audit ..................................................... 14
   4.1 Option 1: Deem the vision presented by IOC as unattainable, instead focusing on harmonizing access to resources under IOC/IODE control using IODE resources .................. 14
   4.2 Option 2: Develop the Ocean Data and Information System independently using IODE and/or external resources ............................................................................................................. 15
   4.3 Option 3: Work with one or more established systems to extend their scope to support the global community and seek support from global, regional, and national bodies on areas where their investment and support are required ............................................................................................................. 16
   4.4 Recommending a path forward ............................................................................. 17
5. Proposed solution ..................................................................................................... 19
   5.1 Portal implementation ............................................................................................ 20
   5.2 Metadata brokering/mediation tier ........................................................................ 22
   5.3 Data brokering/access .......................................................................................... 24
   5.4 Data assembly, submission, and provider support .................................................. 25
6. Implementation plan ................................................................................................ 27
7. Conclusion ................................................................................................................. 30

ANNEX 1: Ocean Data and Information System - Conceptual Architecture’.
ANNEX 2: Existing operating models and relevant activities
ANNEX 3: Acronyms
ANNEX 4: Reference Documents
1. Introduction

In April 2016, a team of external auditors was engaged by UNESCO to undertake an audit of the following aspects of Intergovernmental Oceanographic Commission (IOC) operations:

A. Management framework
B. Budget management
C. Human resources
D. IOC activities and programs

The audit of IOC activities and programs includes "Section 5: Information exchange system and marine data management.", which outlines observed barriers and issues related to marine data and information discovery, access, and exchange, along with other organizational issues likely contributing to the current state of the marine data and information system landscape.

The observations presented in the audit included two critical points of particular relevance to IODE and JCOMM:

157. There is no common database for all marine sciences, which is certainly unattainable when one considers the amount, complexity and heterogeneity of the information to be assembled, but there is also no common portal for all marine sciences that connects all websites and relevant disciplines through web links.

158. This array of data and information reflects the diversity of the stakeholders, the complexity of the subjects, the variety of material and the obsolescence of several systems. For its users, it appears to be part of the avowed landscape of marine science. For the outside observer, it seems to be the product of historical, technical, organizational, and disciplinary build-up, rather than the result of rational construction. This is even more evident as programmes like the International Oceanographic Data and Information Exchange (IODE) have existed for more than 50 years, with the express goal of encouraging the exchange of data, at least of metadata, between the different counterparts of oceanographic disciplines.

These observations reflect the complexity and lack of integration within the marine data and information landscape, while simultaneously highlighting the fact that even with longstanding organizations such as the IODE, there is still no global scale system even for simple search and discovery purposes.

The audit also resulted in a number of recommendations including:

**Recommendation No. 15.** The External Auditor recommends that a draft resolution be submitted to the IOC Assembly calling for Member States to work together, with the support of IOC, to construct a universal information system and ocean data portal, along with a cost-benefit analysis prepared in advance by the IODE project.

The objectives of this working document are to:

1. Elaborate on the scope of the audit and root causes for some gaps and inefficiencies reported in the audit
2. Expand upon the results of the audit and the connections to efforts related to global
marine data and information management
3. Describe the support required to address the issues and recommendations from the audit
4. Present existing working models and project/program activities and their potential relevance in addressing the recommendations in the audit
5. Propose a working model to address the recommendations in the audit
6. Present a high-level DRAFT workplan for consideration by IOC, JCOMM, and the IODE

2. Scope of audit
The IOC audit specifically refers to data and information in a number of forms:

- **Websites** - Web accessible locations containing unstructured information related organizations, projects, themes, etc. Websites may also contain data stored as unstructured information (e.g. simple HTML table) or as structured information (e.g. HTML microdata) hosted outside a managed repository.
- **Databases** - Structured repositories of data and information, regardless of technology or access protocol).
- **Portals** - A web site or other entry point to collection(s) of data and information, typically supported by search and reporting/access facilities.

It is important to stress that websites, databases, and portals are not mutually exclusive, but in the context of the audit, they represent distinct classes of resources having unique characteristics which become clear on publication (e.g. data and information encoding standards impact data and information providers) and discovery and access (e.g. presentation tier, web service/API endpoints impact data and information users).

**Websites**
There have been many marine data and information inventories produced in support of centralized discovery of data and information relevant to a particular topic (project, theme, issue). Such inventories tend not to be maintained as they are out of date before they are completed and they require a significant amount of manual effort in order to assemble and sustain them.

In the internet age, search engines such as Google or Bing attempt to index web accessible content with the goal of providing centralized search and discovery facility. However, the success of these systems is heavily dependent on how information is physically encoded and made available (impacts ability to index content), semantic richness (ability for search engines to retrieve resource in response to search requests), and volume of content (ability to retrieve most appropriate resources from the full set of potential matches). Even with these factors addressed, there remains the fundamental issue that the presence of outdated and even irrelevant content confounds the search and discovery process.

An early leader in global marine science inventories was IODE OceanPortal which was a web-based high-level directory database of ocean related web sites and databases. It was a highly regarded international source-of-choice for marine science information, and was one of the most visited of all IOC websites. OceanPortal was closed in 2009 because it was decided that there were sufficient other web-based search engines to seek ocean related

---

web sites. In addition, the work of keeping the inventory current by discovering new ocean related websites and the maintenance of web links became labour intensive.

More recently, the Atlantic Ocean Research Alliance (AORA) had undertaken the development of a marine data and information system inventory (summary report available at http://ices.dk/explore-us/projects/Documents/AORA-CSA-WP11_European-Catalogue_ver2017.PDF), and although this work was constrained in terms of its geographic coverage, many global systems and web sites were identified through the cataloguing activity which led to the published report. As with the OceanPortal effort, the current inventory was developed manually and focused on specific aspects of these sites’ capabilities.

Databases/Repositories

Although there are many successful large-scale data/information systems in operation today, it is recognized that the diversity of themes and sheer number of content producers renders the vision of a single source database infeasible. A more realistic alternative is to provide federated and/or aggregated search across individual databases and repositories as appropriate.

Beyond the logistics associated with handling such complexity at the data level, there are also the technological factors. New approaches and technology solutions continue to evolve as the volumes of data and information produced continue to grow. However, these advances have only further fragmented the landscape as the most basic solutions (file-based archives) are still in broad use today.

A response to the IOC audit must consider the ongoing efforts focused on improved coordination and integration of data and information within the marine community and to leverage these efforts to the extent possible.

The JCOMM Cross-cutting Task Team for Integrated Marine Meteorological and Oceanographic Services within WIS (TT-MOWIS) (http://www.jcomm.info/index.php?option=com_oe&task=viewGroupRecord&groupID=318) was established to provide a mechanism for international coordination of oceanographic and marine meteorological observing, data management and services.

The work of this team represents a close collaboration with the WMO Information System (WIS), leveraging the success of the WIS operating model and resulting in greater harmonization of operational marine meteorological data and services. Although this effort is focused primarily on operational centres, the implementation of a WIS-like model under JCOMM requires a robust and highly available technical environment, technically and semantically interoperable data and information, both supported through personnel and processes.

To manage the disparities in capabilities of the contributing institutes/systems, and to implement redundancy/failover and high availability of data and information, the type of hierarchical data network structure employed by WIS is relevant to the Ocean Information System to be developed under MOWIS. These areas of focus are all relevant to the Ocean Data and Information System envisioned through the IOC audit, and there are opportunities for the global effort to draw upon the expertise from TT-MOWIS and related efforts.
At the regional scale, the AtlantOS project (https://www.atlantos-h2020.eu/) seeks to implement an integrated Atlantic Ocean Observing System and has completed assessments in several work areas requiring coordination in order to facilitate integration of data across the contributing systems.

The AtlantOS data harmonization report (https://www.atlantos-h2020.eu/download/deliverables/7.1%20Data%20Harmonization%20Report.pdf) outlines many of the integration challenges raised in the IOC audit and although there are recommendations related to standardization on certain core elements, the report acknowledges that setting a single set of standards for all data and supporting information interchange is not realistic. As a result, leveraging the role of integrators to enable federated access to data and information across contributing systems, while simultaneously providing advice back to those systems in order to promote standardization where possible.
The concept of data or information portals has evolved significantly since the rise of the internet. At one time, the goal was to build a single, all-encompassing system to serve needs of all potential stakeholders. As approaches to data and information exchange have evolved, along with the ability to link resources (as opposed to mirroring them), the vision has changed. Although there is still a role for generic portals, the demand is now for theme or issues-based views built upon managed repositories of data and information. Rather than attempting to implement one central repository to contain everything, the current practice is to federate access across disparate data and information stores, harvesting where required for performance and other efficiencies. This adds additional complexity in terms of portal implementation and management, but these thematic resources are shown to provide greater value and relevance to their respective communities. In cases where central portals are provided, they should consider various classes of users and content providers in order to provide appropriate user experience for the various stakeholder groups.

**IODE Information Products**

The International Oceanographic Data and Information Exchange currently supports a number of marine data and information products and repositories and the state of these systems is mirrors the types of variability that can be found within the broader global community.

**IODE OceanKnowledge Platform**

(http://www.iode.org/index.php?option=com_content&view=article&id=367&Itemid=100082) is a new IODE activity intended to offer the user a single access point to various linked IODE information and data products such as researcher profiles, publications, data, learning objects, etc. and will furthermore facilitate social networking between specialized research communities.
The strategic goal of the IODE Ocean Knowledge Platform Pilot Project:

(i) Combining multiple inter-linked IODE (and partner) Data and Information knowledge sources and making them discoverable and accessible through a single access point with social networking functionality;
(ii) Linking IODE (and partner) Data and Information Products through semantic web technologies;
(iii) Promoting IODE (and partner) Data and Information Products within the ocean research and observation community.

**OID Events** ([http://www.iode.org/index.php?option=com_oe&task=eventoCalendar&Itemid=45](http://www.iode.org/index.php?option=com_oe&task=eventoCalendar&Itemid=45)) provides information about meetings, workshops, conferences and IOC/IODE assemblies. This information includes participant list, documents and other metadata related to a particular event. The display template is embedded with microdata from the schema.org Event schema ([http://schema.org/Event](http://schema.org/Event)). This data is available through a SPARQL endpoint.

**Ocean Biogeographic Information System (OBIS)** ([http://www.iobis.org/](http://www.iobis.org/)) builds and maintains global alliances that collaborate with scientific communities to facilitate free and open access to, and application of, biodiversity and biogeographic data and information on marine life. It uses an extended version of the Darwin Core metadata schema ([http://rs.gbif.org/core/dwc_occurrence_2015-07-02.xml](http://rs.gbif.org/core/dwc_occurrence_2015-07-02.xml)), for sampling events and species occurrences. At present, OBIS provides species occurrence data from some 2000 datasets. The metadata regarding these datasets is available for inclusion into a triple store repository.

**OBIS-ENV-DATA** ([http://www.iode.org/index.php?option=com_content&view=article&id=463&Itemid=10020](http://www.iode.org/index.php?option=com_content&view=article&id=463&Itemid=10020)) is an in extension to OBIS in order to cover coincident environmental data. The goals of the project are (i) to ensure mixed data sets stay together, (ii) to demonstrate that OBIS nodes, NODCs and ADUs, that hold mixed data sets, can collaborate on the joint management and exchange of biodiversity observation data including environmental measurements, and (iii) to show what the benefit of this approach is for marine sciences, biological analysis and modelling.

This project will also investigate how these mixed datasets can additionally flow to regional and global data repositories such as OceanDataPortal, World Ocean Database, SeaDataNet, etc.

**Ocean Data Portal** ([http://www.oceandataportal.org/](http://www.oceandataportal.org/)) is a system that collects, integrates and manages physico-chemical data from participating data centres and projects. It is based on a model where data is received from individual nodes (mostly based at IODE NODCs and ADUs) and is presented to the users at a single location. It uses a metadata schema based on ISO 19115. It also has transformation procedures 19139. It supports OGC, CSW, SRU and OAI-PMH.

**OceanDataPractices** ([http://www.oceandatapractices.net/](http://www.oceandatapractices.net/)) is a repository of best practices, manuals and guides. It is a project supported by a group of organizations such as World Meteorological Institute (WMO), Scientific Committee on Oceanic Research (SCOR), Joint Technical Committee for Oceanography and Marine Meteorology (JCOMM) and International Council for the Exploration of the Sea (ICES). IOC/IODE plays an important role in supporting the initiative by hosting the repository and providing content related to its best practices, manuals and guides. Technologically it is similar to OceanDocs and uses DSpace as the repository software and supports Linked Open Data (LOD).

**OceanDocs** ([http://www.oceandocs.org/](http://www.oceandocs.org/)) is a repository of marine science related publications but also includes other media types. It does not exclusively, but mainly, focuses on works
from the IOC Member States and IOC Partners. It offers a free hosting service to any marine related organization, programme or project who needs to expose and permanently archive its research output e.g. LMEs. It employs extended Dublin core metadata schema (http://dublincore.org/documents/dcmi-terms/). Open source software called DSpace is used for the repository. It has native support for Linked Open Data (LOD). Metadata is converted to RDF and is loaded into a triple store. Persistent URI’s are obtained through an external service known as Handle.Net (http://handle.net/) and resolved locally using a handle server.

OceanExpert (http://www.oceanexpert.net) is a database of individuals and institutions. It also includes information on event and job postings. Since its creation in 1997, it has been using a self-enrolment method. All information for an individual of an institution, except some contact information is available publicly online. Output templates use schema.org schemas for presenting content for Individuals (https://schema.org/Person) and Institutions (https://schema.org/Organization), in a structured format. Microdata is managed using a triple store and the system has a SPARQL endpoint.

OceanTeacher Global Academy Courses (http://classroom.oceanteacher.org/) has built a comprehensive Learning Management System (OceanTeacher) that, in combination with classroom training, has trained nearly 2000 students from 120 countries since 2005. Its databases include data on: individuals, locations, courses, study resources such as PowerPoint, reading lists, and course videos. It uses Moodle a free and open-source software learning management system written in PHP.

See ANNEX 2 for existing operating models and relevant activities including some other examples of national, regional, and global databases/information products

Relevance to the IOC and responding to the recommendation from the audit

In the context of the IOC audit, the changing landscape in terms of technology, the ability to link resources across systems and domains, and fundamental changes in how data and information are published and communicated to their intended audiences will all contribute to any solutions proposed through this concept paper. Although the audit points to the variety of standards in use as an issue, the any recommendations put forward must respect that these established standards exist and that efforts should be placed on federation rather than attempting to gain support for single standards (though convergence should always be assessed and adopted where possible).

3. Key root causes underlying issues described in the IOC audit

Many of the issues described in the audit are not new developments, but have been evolving over time. In order to fully appreciate the relevance of the points from the audit, and to propose options to move forward, it is important to be aware of key root causes as these will be raised throughout the response to the audit:

- Engaging data and information providers and the role of the IOC
- Disparities which impede participation
- Competition (or perceived competition) between efforts which may or may not be linked to IOC
- Diversity in content standards
- Diversity in technology solutions
3.1 Engaging data and information providers and the role of the IOC

In order for IOC to achieve its goal of a universal data and information system, the IOC must find a way to engage communities ranging from those already linked to IOC to those who are operating independently. Stakeholders will not always share IOC/IODE perspective on sharing of marine data and information. The issues of competition, geostrategic, and legal barriers mentioned in the audit are potential factors which may impact the IOC’s ability to foster collaboration.

Failure to engage the community of stakeholders not currently linked to the IOC will prevent (or significantly inhibit) the ability to deliver on the IOC vision regardless of whether or not all other supporting elements are in place.

3.2 Disparities between regions which impede participation

Although the IOC audit refers to inter-regional disparities, inequalities exist at different levels and can manifest themselves within regions as well. In order to provide feasible recommendations in response to the audit, the following forms of disparity must be considered:

- **Disparities in representation** - In keeping with the engagement barrier mentioned previously, any solution put forward must be open in terms of participation and also consider stakeholder groups that may not be directly linked to IOC today.

- **Disparities in technical capabilities and knowledge** - In order to succeed in the implementation of the Ocean Data and Information System, it is critical to consider the capabilities of data and information providers and users and the impact on data and information submission as well as discovery, access, and use. In essence, the different levels of technical capability required from both the provider and user end would include (i) generalist (non-technical, limited access to technical infrastructure, requires highly synthesized level of support and user experience), (ii) domain specialist (more technical expert, may have access to greater technical infrastructure, requires more flexibility in terms of domain standards support and ability to interact with any global system), and (iii) system to system (technical) connection (highly technical and non-interactive access to data and information, high demand for technical infrastructure, highly standardized in terms of interoperability protocols).

- **Disparities in technical environment/capacity** - The operational stewardship of data from acquisition through to dissemination involves an inherent level of IT capacity. However, it is not uncommon to find data providers managing data and information on infrastructure well below the levels required in an operational setting. The limitations in these environments raise themselves as data providers commit to broader sharing of data, resulting in increasing numbers of requests against the infrastructure, increasing demands in hours of service, and demands for new services to be implemented for accessing data and information.

- **Disparities in human capacity** - Coupled with expectations for in-kind contributions of IT capacity, it is also common for enabling human capacity to be considered an in-kind responsibility. In cases where individual researchers or programs take on this enabling support internally, they may be limited in how they can respond to requests to share data and information. From low level systems administration (server resource provisioning, operating system updates and maintenance, backups and monitoring), to application service implementation and monitoring, to data management support, committing to providing high performance, reliable, and secure access to data and information may present cost and skillset barriers that cannot be easily borne by the data provider.
Because the disparities referred to in the audit can occur at many levels, from data providers
to established data or information systems, recommendations in response to the audit must
take these realities into account and strive to solicit support as close to the source as
possible. For example, in order for a frontline data provider to contribute to the Ocean Data
and Information System, they must have a system to which they can submit data for long
term stewardship and downstream publication. This level of support would start with the
NODCs and ADUs in the context of IOC.

3.3 Competition (or perceived competition) between efforts which may or may
not be linked to IOC

The audit mentions three established systems which different in their technological
implementation, but which overlap in terms of thematic coverage and mandate:

- The IODE Ocean Data Portal was established to “provide seamless access to
collections and inventories of marine data from the NODCs in the IODE network and
will allow for the discovery, evaluation (through visualisation and metadata review)
and access to data via web services.”

- The goal of SeaDataNet is to provide “a standardized system for managing
the large and diverse data sets collected by the oceanographic fleets and the
automatic observation systems. The SeaDataNet infrastructure network and enhance
the currently existing infrastructures, which are the national oceanographic data
centres of 35 countries, active in data collection. The networking of these
professional data centres, in a unique virtual data management system provide
integrated data sets of standardized quality on-line. As a research infrastructure,
SeaDataNet contributes to build research excellence in Europe.”

- WOD is the “world’s largest collection of vertical profile data of ocean
characteristics available internationally without restriction.”

In terms of their respective levels of success:

- The IODE Ocean Data Portal has not achieved full coverage of the IODE
network since its establishment in 2007. There has been a level of success in terms
of metadata exchange under the Ocean Data Interoperability Platform brokering
prototype activity (where SeaDataNet collection level records can be searched from
the ODP), but from the perspective of a data user, these system to system
relationships can result in more confusion if the purpose for the connection is not
clear.

- SeaDataNet has been very successful in in establishing a centralized system
for providing access to data from many European countries, though its level of
openness is limited to shopping basket access to data as a registered user.

- The World Ocean Database has been very successful as a data compilation
product, but as with SeaDataNet data access is limited to request/download and
there is overlap with data present in both SeaDataNet and IODE Ocean Data Portal.

Beyond this example, the situation described in the audit is not new. Even within the scope
of the IOC, the data and information systems most closely linked to the IOC are not
accessible through a single integrated portal and underlying repositories are also
separate.
It is important to consider that data may be duplicated or harvested into multiple systems for many reasons:

- To expose and facilitate the use of data in analysis and product generation.
- To integrate data with that from other sources in order to provide a more comprehensive view of a region or domain of interest.
- To incorporate data into a long-term archive where this cannot be resourced by the data provider’s organization.
- To enable transformation of data into a standard not supported by the originating organization.
- To provide greater performance, failover, and other capabilities.

Although standards, IT capacity, and open data culture have advanced, there still exists a degree of competition and duplication between systems. This lack of coordination and direct and indirect competition between systems creates challenges for both data and information providers and users:

- Competition between efforts results in duplication of back-office systems and front-line portals used in the discovery and access to data and information.
- Data and information published to multiple systems results in duplication of content when federating searches across systems, confounding search results and leaving the task of de-duplication to be resolved.
- The sheer number of data and information systems that could potentially serve as input into the Ocean Data and Information System creates technical complexity - how to be engaging while being cognisant of the interoperability, availability, and reliability issues that exist.

Although the longer-term goal must include improving collaboration between the various marine data and information systems, this is not likely to be fully achieved in the short term. Therefore, recommendations in response to the audit must include clear statements on what can be achieved in the short term versus longer term, and the approach for federating discovery and access to data and information must include layers of aggregation in order to manage the complexity which exists in the current ecosystem.

### 3.4 Diversity in content standards

Content standards and best practices are critical to sustainable operations and facilitating the discovery, access, and integration of data across individual systems. However, the information management and technology landscapes have exploded with new, and sometimes competing, options for accomplishing the same tasks. As a result, data and information providers are left with a great deal of flexibility in how they publish, leaving data and information integrators and systems providers faced with increased complexity when trying to provide some level of harmonized discovery and access.

Some common examples include:

- **Metadata standards** - Metadata is critical to capturing structural and contextual information about a published resource, be it a dataset, a web page, or document, or a value-added product. The technologies used to make data and information available are increasingly metadata driven, if for no other reason than to enable the user to filter the wealth of resources available. The richness and complexity of the standards vary, but a feature they have in common is that the “content standard” (what information is collected) is generally accompanied by an “implementation (encoding) standard”, which tends to be machine readable.

- **Data and information standards** - Similar to metadata interchange, many standards exist for the encoding of data and information for exchange and these standards may be more open (e.g. XML, CSV, NetCDF) or constrained based on a more rigorous specification (e.g. HTML 5 or ePub for internet content, Climate Forecasting (CF) implementation of NetCDF). Regardless of the standard employed, the use of open standards (documented, easily accessible through the use of freely available software and tools) is preferred over proprietary standards where use of commercially licensed tools is a barrier to use.

- **Resource themes** - Resource themes are increasingly used to categorize data and information in order to provide a more efficient user experience. Themes may be derived from many sources such as the ISO topic categories included in ISO 19115 metadata content standard (http://standards.iso.org/iso/19115/resources/Codelist/cat/codelists.xml), legislated sources such as the INSPIRE Theme Register (http://inspire.ec.europa.eu/theme/), domain themes such as the Global Change Master Directory (GCMD) Science Keywords list (http://gcmdservices.gsfc.nasa.gov/static/kms/sciencekeywords/sciencekeywords.csv?ed_wiki_keywords_page), organizational themes such as the Government of Canada Core Subject Thesaurus (CST) (http://www.thesaurus.gc.ca/recherche-search/thes-eng.html), organizational mandate/structure/priorities, and so on.

- **Vocabularies** - Although internet search engines, machine learning techniques, and other tools continue to develop in order to improve the quality of searches without adding undue effort on the part of resource providers, there continues to be a need for specificity in how the attributes of data and information resource are defined. Most catalogues which provide faceted search capabilities are employing fixed vocabularies or ontologies. Examples of commonly used vocabularies include the BODC Parameter Usage Dictionary (P01) (http://vocab.nerc.ac.uk/collection/P01/current), GCMD Science Keywords (https://wiki.earthdata.nasa.gov/display/CMR/GCMD+Keyword+Access), and World Registry of Marine Species (WoRMS) (http://marinespecies.org/).

- **Persistent identifiers** - The unambiguous identification of an entity, such as a dataset, an individual, or a device, is a critical step in the documentation of its provenance. In the context of a dataset published on the internet, provenance is key to identifying the unique version provided by the originator versus mirrors or deprecated versions which may appear on other data portals or web sites versus a version of the data which may have been subjected to ancillary processing and quality control and can be seen as a related, but distinct entity in its own right. Systems such as DOI and Handle are used to identify various classes of information entities. Going beyond the provenance of a specific resource, there are other examples of where persistent identification systems are providing critical foundation to the semantic web. The persistent identification of individual people and organizations is being used in linking these entities to their published works. More importantly, these connections are being populated by many organizations and are not reliant solely on self-identification of publications/resources by individuals. As
these linked data systems continue to evolve and connections are made, it will be possible to automate more of the community engagement, similar to how systems such as Facebook and Linked-In operate now. Examples of such efforts include IODE OceanExpert (http://www.oceanexpert.net/), regional efforts such as the European Directory of Marine Organizations (EDMO) (http://seadatanet.maris2.nl/v_edmo/welcome.asp), and community efforts such as ORCiD (https://orcid.org/) and VIVO (https://researchlink.rdc-drc.ca/vivo/). In addition, proprietary and nonproprietary personal identification/authentication systems such as ResearcherID (http://www.researcherid.com), International Standard Name Identifier (ISNI) http://www.isni.org/ and OpenID (http://openid.net/) contribute to the PID offer.

These examples represent a small subset of the rich landscape of content, technical, and semantic related standards and tooling which exists. Therefore, recommendations in response to the audit must include efforts to foster the support and use of mature standards where they exist, educating the data and information provider communities in order for them to adopt appropriate standards, and to ensure that the Ocean Data and Information System leverages advanced capabilities such as those in the semantic and linked data spaces in order to gain the benefits of work already being undertaken at no cost.

3.5 Diversity in technology solutions

The technology solutions that can be employed at any level is variable to some degree and the choices made can be driven by many factors such as:

- **Legacy support requirements** - The length of time an organization has been in operation can have a significant impact on its forward-looking technology strategy. Organizations which employ light technical solutions may provide a more limited user experience (e.g. data download services via FTP as opposed to GIS aware web services), but are also less impacted by change in IT industry.

- **Standards in use within the community** - For operational environments and other specialized communities, the technology solutions and interoperability requirements may be set by the community itself.

- **Size or capacity of the organization** - Organizations having limited technical capacity or expertise may choose simple or more stable technology solutions where organizations having greater technical capacity or expertise may choose a more agile approach (i.e. accepting more frequent change with the benefit of providing a richer user experience).

- **Needs of the communities supported by the organization** - Organizations may choose to employ technology solutions based solely on the requirements of communities they support.

Although the importance of well-managed data cannot be overstated, the implementation of a cross cutting data and information system does not necessarily equate to a complete tear-down exercise on the part of data and information providers. Rather than attempting to enforce a single technical solution from the ground up, it is important to focus on where existing data and information systems end and where the Ocean Data and Information System envisioned by IOC begins.

At a conceptual level, the critical aspects of any data or information system include:
- **Ingestion tools** – There is need for data and information providers to have access to supported submission processes and ingestion tools to facilitate contribution the systems that will feed the Ocean Data and Information System. Although these technologies are generally outside the scope the higher-level discovery/access portals, these are critical resources for organizations delivering data and information management functions.

- **Storage and management infrastructure and tools** - Technology options here range from traditional file and database driven infrastructures to large scale unstructured archives, to linked data systems. From the perspective of the Ocean Data and Information System, these technologies are generally out of scope. However, the performance and scalability of the hardware infrastructure, along with the interoperability protocols and data and information exchange standards are of critical importance. As these systems become more connected and visible to the user community, this increased presence has the potential to increase demands on IT infrastructure.

- **Publication tools** - Publication tools may exist at multiple levels and from the perspective of the Ocean Data and Information System, these may be relevant a) if the Ocean Data and Information System allows for direct contribution by individuals or systems (as opposed to harvesting or federating access to existing regional, national, or thematic systems), b) if the Ocean Data and Information System seeks to provide guidance or standardization at the global scale (e.g. in order to obtain buy-in for individuals to adopt a new standard, it must feasible for them to do so), or c) if the existing tools do not fully address the large network of data and information providers (resulting in barriers to participation)

- **Discovery tools** - Because the Ocean Data and Information System will provide a level of federation across existing systems, there may be opportunities to adopt established technologies, though the specific decisions will be based to a large extent on partners involved in the implementation along with the functionality and interoperability and content standards which need to be supported. This ability to provide multiple levels of search capability and to search across content standards, will be critical features.

- **Visualization tools** - Although there may be some opportunity to leverage globally accessible data and information to produce new value-added products, much of this work will likely be the results of federated access to products originating from domain experts. In this scenario, it is not so much the visualization tools that will matter, but how access to those tools and products is achieved.

- **Access tools** - As is the case with discovery tools, the Ocean Data and Information System will likely be employing federation across systems. The specific technologies employed to enable access to data and information are less important than the specific types of access provided. Interoperability protocols and content standards, coupled with the capacity and performance of the physical IT infrastructure, will be critical.
4. Options to consider in response to the IOC audit

Given the scope of the problem presented by IOC, the numerous past and ongoing efforts by many organizations to deliver data and information to the marine community at various levels (national, regional, global, domain-specific), and the realities of the highly variable state of the marine data and information communities, IODE has a number of avenues to explore:

- **Option 1:** Deem the vision presented by IOC as unattainable, instead focusing on harmonizing access to resources under IOC/IODE control.
- **Option 2:** Develop the Ocean Data and Information System independently using IODE and/or external resources.
- **Option 3:** Work with one or more established systems to extend their scope to support the global community, and seek support from global, regional, and national bodies on areas where their investment and support are required.

4.1 **Option 1: Deem the vision presented by IOC as unattainable, instead focusing on harmonizing access to resources under IOC/IODE control using IODE resources**

At the technical level, this option follows on the past and ongoing efforts of the IODE Ocean Data Portal and OceanKnowledge Platform, and involves re-aligning work activities and resourcing infrastructure and related resources to meet the demands (technical and non-technical) and the IOC deadlines.

There are many benefits of this approach:

1. By focusing on systems under IOC direct control, there is a higher probability of partial success in terms of delivering on a more modest scope. There may be a need to provide support at the regional level, and the IOC has indicated that a level of support would be provided.

2. This option provides an opportunity to improve resourcing and support to the IODE OceanKnowledge platform, which is intended to provide this harmonized access on an IOC scale with the broader concept as a vision.

3. This option provides an opportunity to realign existing contributions to IODE technical activities, which is linked to the efforts to ongoing proposals to restructure the IODE in a broader sense.

4. This option provides an opportunity to make fundamental improvements in how data and information from IOC programs are published, providing a real-world use case for educating and supporting IOC member states on more open and interoperable data and information publication.

5. This option provides an opportunity to scale the solution, if successful, to include data and information from IOC member states.

There are a number of risks and issues associated with this approach as well:

1. Focusing solely on the IOC linked systems and content does not address the
need for the Ocean Data and Information System, which is a failure in terms of delivering on the recommendation from the audit. The resulting system will not reflect the full marine community and existing “siloes” such as those described earlier in this document, will remain.

2. IODE resourcing has been, and continues to be, a challenge. Delivery on IT related projects in the past has resulted in mixed success and it is not clear that continuing with the current model will result in success even within the IOC space.

3. Although there may be some efficiencies arising from this work, it is also possible that the costs to IODE in terms of computing infrastructure and people support, will increase. These ongoing costs are currently unclear and represent a risk should support from IOC and other sources not materialize.

4. Until the IODE restructuring has been presented and agreed, it may be premature to assume realignment of existing efforts or impose increased demands on teams that are providing in-kind support.

4.2 Option 2: Develop the Ocean Data and Information System independently using IODE and/or external resources

In essence, existing projects and systems will continue as they have been, but new development will be from the bottom up using current standard/technologies to deliver an optimized solution.

Taking a ‘clean slate’ approach provides a number of benefits:

1. This option provides the ability to leverage advances in terms of computing infrastructure, software technologies, content standards, without the potential burdens associated with legacy support/issues associated with a pre-existing system.

2. This option provides an opportunity to re-engage data and information providers as part of a new collaboration, possibly resulting in some improved working relationships by putting aside previous perceived and real competition, and other non-technical issues from the past.

3. This option provides more control in the implementation and resourcing of the final solution. By not having associated pre-existing commitments and support issues, determining the implementation and operating costs for the Ocean Data and Information System will be easier to achieve and the results of this exercise will be not confounded by other arrangements that must be accommodated (either politically or due to pre-existing commitment).

Although this option provides an opportunity to establish a more robust starting point for delivering on the recommendation from the audit, there are many risks associated with this approach:

1. Taking a ‘clean slate’ approach may be seen as more competition and a repetition of past failed efforts to implement the Ocean Data and Information System.
2. This option has the potential to result in a much longer development cycle because decisions already made with existing systems will need to be re-visited. The fundamental approach taken to implement the Ocean Data and Information System, its design and architecture, standards, support and delivery options, are all on the table for discussion.

3. Building a new system from the bottom up initially requires more testing and support at all levels, and there is inherent risk that data and information providers could walk away from the system before it stabilizes. In any distributed or multi-tiered system, one deals with foundation elements (e.g. hardware infrastructure, data access layers, security layers, etc.), along with content delivery and related tiers (e.g. data and information delivery, analytics, user experience). It is critical that foundation tiers be stable and robust in order to focus on needs and support for data and information providers and users (the higher-level tiers). Well established systems have a stable foundation whereas new systems will not have this initially.

4. Given existing capacity issues, IODE must carefully consider its role in the execution of this option. With efforts to restructure IODE already under development, undertaking a new project such as this should align with how IODE sees its role going forward. Initiating this work following the ‘old’ mindset has the potential to maintain the status quo, which is already not fully successful.

The audit included a point related to resourcing the development effort, suggesting that firms involved in compiling and selling relevant data to specific industries (e.g. oil and gas) may be willing to provide support to this effort. Enabling data and information providers to publish these resources in discoverable and interoperable forms, supports the goals of both IOC and data brokering organizations, but may introduce a commercial complication.

4.3 **Option 3: Work with one or more established systems to extend their scope to support the global community and seek support from global, regional, and national bodies on areas where their investment and support are required**

There are many data and information systems already in place and are seen to be effectively serving their specific communities. Rather than building another system in isolation, it may be possible to enhance and leverage the strengths and connections between mature systems to deliver the Ocean Data and Information System envisioned by the IOC.

There are many benefits of this approach:

1. This option is intended to fully delivery on the recommendation from arising from the audit, thus increasing the presence of IOC within the marine community.

2. By building on the success of established systems, there is reduced potential for the Ocean Data and Information System to fail at the foundation level. Other projects such as AtlantOS and MOWIS are considering fundamental issues ranging from infrastructure, standards, barriers/strategies for
harmonization, etc. in their implementation strategies, and the Ocean Data and Information System can benefit from these complementary efforts.

3. This option provides an opportunity to help re-shape the role of IODE in a landscape where its’ resourcing is very low compared to other organizations working in the same space. Rather than building from the bottom up, IODE plays a key role in evaluating and recommending working solutions where they exist, and providing input into these solutions in order for them to scale to the broader marine community.

4. This option provides an opportunity to realign existing contributions to IODE technical and support activities, which is linked to the efforts to ongoing proposals to restructure the IODE in a broader sense.

5. This option provides an opportunity to make fundamental advances to how data and information from IOC and non-IOC programs are published.

There are a number of risks and issues associated with this approach as well:

1. By focusing on both the IOC and non-IOC systems, greater levels of engagement are required and the associated risk of failure is higher.

2. IODE resourcing has been, and continues to be, a challenge. There is a potential that IODE work itself into irrelevance by relying on much larger programs for the delivery on this effort.

3. Although there may be some efficiencies arising from this work, it is also possible that the costs to IODE in terms of people support, will increase. These ongoing costs are currently unclear and represent a risk should support from IOC and other sources not materialize.

4. Until the IODE restructuring has been presented and agreed to, it may be premature to assume realignment of existing efforts or impose increased demands on teams that are providing in-kind support.

4.4 Recommending a path forward

In order to recommend a path forward, it is important to take stock of existing capabilities and where they may be better leveraged at the global level to achieve the goals of IOC.

Annex 2: Current and past activities relevant to the Ocean Data and Information System, lists a number of regional and global efforts and systems whose capabilities and supporting operating models are relevant when considering options on to address the recommendation from the audit.

To support this task, a light capability assessment was developed and is currently being applied across a number of established systems in order to better understand their respective operating models, their approach to serving their respective communities, their use of technology, standards and best practices, client engagement, and their business models in terms of operational support and forward planning. Using this information, potential collaborations between IODE and existing players will develop, along with recommendations in terms of supporting activity where IODE may be a participant or lead, and also where operational support at the regional and national levels is seen as necessary.
The accompanying document, ‘Analysis of marine data system capabilities’, describes the assessment approach being undertaken to review existing, well-established marine data systems, in order to support recommendations back to IODE and IOC. Attempts are being made to evaluate the capabilities of the selected systems in a number of areas:

- Core technical capacity
- Content coverage and standards
- User experience
- Data provider experience
- Onboarding costs and sustainability
- Quality of service

Systems that were evaluated fit within one or more of the following classes:

- Portal provider
- Metadata broker
- Data aggregator
- Secondary (product) archive
- Primary archive

Furthermore, these systems classes were further sub-classified by the types of provider and user relationships they currently support.

The systems were also evaluated to the extent possible on their support for three classes of data and information providers and users:

- **Simple provider/user** - Limited technical abilities/infrastructure; More ad-hoc in terms of content and structure; Less formalized in terms of standards use; Fewer requirements in terms of system functionality (e.g. simple submission processes, simple query capabilities); Limited organizational support in terms of IT and data management

- **Advanced provider/user** - May be limited in terms of technical abilities/infrastructure; More likely to employ standards in terms of content and structure; More formalized in terms of standards use; Greater requirements in terms of system functionality (e.g. advanced submission processes, complex query capabilities); May be limited organizational support in terms of IT and data management

- **System-based provider/user** - Advanced in terms of technical abilities/infrastructure; Greater use of standards in terms of content and structure; More formalized in terms of standards use; Greater requirements in terms of system functionality (e.g. system to system interoperability arrangements, advanced submission processes, complex query capabilities); May be limited organizational support in terms of IT and data management; More likely to require extended hours of service

The result of this effort is a **conceptual architecture model** including the various classes of infrastructure components, potential collaborators to work with, and a breakdown of key enabling functions/activities required to successfully deliver and sustain the environment.

Where possible, practices already employed by existing systems/organizations, have been identified as examples of where broader implementation is recommended.

Please refer to ANNEX 1: Ocean Data and Information System - Conceptual Architecture.
5. Proposed solution

Addressing the observations in the audit will require effort on multiple levels. Aspects of this work may be undertaken by IODE and its partners directly, while other aspects will require higher level support and contributions by the IOC member states. The degree of success will depend heavily on collaboration, providing adequate support to stakeholders, and employing an evolutionary approach in order to improve the efficiency and robustness of the proposed system over time.

Although all three options presented previously were considered, it is recommended that efforts be made to undertake Option 3 and to work with major stakeholders in the marine community to collaborate on implementing the environment envisioned by the IOC. The goal will be to use this new working relationship to evolve the Ocean Data and Information System over time in order to improve its capabilities and richness in terms of content, interoperability, resource provider and user support.

Although mobilizing disparate teams with varying levels of capacity will have its challenges, this strategy is seen to be more realistic than expecting existing players to invest in a completely new system. Not only does this approach ensure short term results leveraging existing data and information resources, but by seeking opportunities to increase participation in the existing systems as inputs into the Ocean Data and Information System, both the IOC and the contributing systems will benefit. Ideally, the level of new development directly under the responsibility of IODE would be very limited.

In essence, by working with data and information providers to make their resources more accessible, interoperable, and semantically rich, the solution will be the implementation of an ocean knowledge platform where with working level responsibility for infrastructure, data and information repositories, provider support would remain with the programs and organizations equipped to manage these responsibilities.

In order to fully realize the vision of the Ocean Data and Information System, it is critical to consider not only the technical scope of the system, but also the roles of IOC, IODE, and JCOMM in the delivery. The full audit includes statements related to limitations in resourcing, varying levels of involvement of IOC in existing programs and systems, and other barriers to IOC’s leadership and role in the delivery of the Ocean Data and Information System. Rather than fighting these realities, a more useful approach would be to embrace the barriers and related challenges which exist. Rather than attempting to lead and implement the full technical solution with its partners, a more practical approach would be for IOC to act more as a client and product owner of the Ocean Data and Information System, using IODE as a vehicle for translating high level business requirements into technical requirements at a level which can be responded to by technology/system partners and which can be measured against for success.

Although a small number of potential collaborators are mentioned by name in this chapter, it is important to stress that the full landscape of contributing organizations and data and information systems is too large to list in full. The end goal is to establish the Ocean Data and Information System such that it is inclusive and is architected to support stakeholders at all levels.

The conceptual model envisioned for the ocean data and information system would be as follows:
5.1 Portal implementation

IODE would solicit partners to work with in terms of implementing the front-end portal that would act as the front end to the data and information systems and resources across the marine landscape.

Recommended partners would include:

1. **IODE OceanKnowledge Platform and Galway Knowledge Platform work packages** - These efforts would assist in developing requirements the Ocean Data and Information System in terms of user experience, categorizing published resources, and defining core versus optional functionality. In both cases the end goal of the project is a unified knowledge platform spanning data and information. Achieving this will require technology solutions and supporting semantic models, and in both cases user experience is important. These user driven projects are more likely to result in the type of system that IOC, leveraging technical solutions (e.g. brokering, etc.) implemented and supported by technical teams.

   It is expected that this product owner function will include analysis of existing portals delivering data and information at varying levels of complexity (supporting different classes of providers and users) and that there will be opportunities to showcase innovative capabilities from the contributing systems.

2. **WMO/WIS, JCOMM/OIS, EMODNet, OBIS, GEOSS** – The technical implementation of the portal component of the Ocean Data and Information System should take into account the experiences of mature global scale technology partners and consider both the technical aspects of delivery as well as user experience. Systems such as WMO/WIS and the planned implementation of JCOMM/OIS are focused on operational marine meteorological and oceanographic systems, the challenges associated with delivery of a high performance, highly available, and fault tolerant data systems also applies to the implementation of a global portal. Furthermore, this collaboration may provide an opportunity to leverage the WMO telecommunications network and to add a level of capability that may not be readily available through commercial offerings. Systems such as EMODNet, OBIS, and GEOSS provide insight into user experience and demonstrate approaches for managing the ingestion/access of data through development of synthesized products, and providing access through theme-based systems.

3. **Internet search providers, translation engines** – Existing marine data and information systems tend to be limited in terms of linguistic support. The implementation of the Ocean Data and Information System will require a much broader linguistic profile, and although it is unreasonable to expect all content from all regions to be available in all languages, it is expected that components of the Ocean Data and Information System will require much broader linguistic support than what is currently available in existing systems. Internet search providers and dedicated translation engines exist which may be leveraged in the delivery of a more globally accessible user experience.

4. **Implementation team** – Between the IOC and external collaborators, a lead implementer will need to be established who will work with the other
collaborating teams to develop and support the Ocean Data and Information System. This role may involve multiple partners implementing a distributed system having sufficient fail-over and performance optimization to deliver an effective user experience.

**Implementation strategy:** In terms of implementation strategy, the approach that will be taken for the implementation of the portal tier will follow a more typical agile development lifecycle, developing through iteration, with key milestone releases being implemented in production.

**Requirements definition:**
The successful delivery of the Ocean Data and Information System is contingent on identifying the target audience and ensuring that the user experience is tailored to support the intended classes of users. With growing demands to improve engagement with the public, to support non-technical decision makers, while continuing to support more traditional technical and research communities, it is anticipated that a multi-faceted user experience will be required. The knowledge platform initiatives are already engaged and tasked with supporting the delivery of data and information to specific communities. By taking the experience in these efforts into account, the goal will be to leverage their experience to define requirements such as:

- classes of user to be served by the Ocean Data and Information System
- identify and prioritize expected outputs from the system
- identify types of interactions to be supported between users and the Ocean Data and Information System
- establish the core (mandatory) and secondary (optional) capabilities to be delivered by the system by use class, along with core and secondary capabilities that will apply across user classes
- establish the general look and feel of the portal
- identify elements requiring semantic development (e.g. data and information categories, themes to be covered by the portal, etc.)
- identify requirements for accessibility and device support

These requirements will help form the technical design of the system and related processes, other dependent requirements that will need to be addressed by the system.

**Solution design:**
As a global level operational system, the WMO/WIS (and JCOMM/OIS as this work progresses) team is an excellent collaborator in terms of planning and implementing a scalable, highly available system having entry points across many continents. In terms of user experience, design, and implementation, looking to how systems such the EMODNet thematic portals, OBIS, GEOSS delineate simple versus advanced access, employ products versus providing source data, and manage discovery level versus full metadata will optimize the experience for support user classes.

At the portal level, technical elements include:
- define interfaces with users, the metadata and information brokering and mediation tier, and external services
- development of technical and information architectures
- selection of technology solutions and level of adoption (adopt as is, adopt with customization, new development)
- develop infrastructure provisioning, management, and support requirements
- develop change management and other enabling processes
Implementation and testing:
The implementation of the portal tier for the Ocean Data and Information System is expected to occur in phases, scaling out as the support for resources feeding the metadata brokering/mediation tiers increases. Functional, performance, and load testing will be performed by technical teams, and acceptance testing will be undertaken by knowledge platform participants in conjunction with clients from communities they serve.

Acceptance and close-off:
Once the portal tier is accepted, it will be promoted to production through the change management strategy developed by the technical team.

5.2 Metadata brokering/mediation tier

The efforts of standards bodies such as ISO, W3C, OGC, coupled with the technology development efforts under programs such as GEOSS, ODP, ODIP, data.gov.uk, and the commercial and open source communities provide options for describing (metadata, semantics) and enabling access (interoperability protocols) to data and information across domains and systems. However, the level of horizontal integration between standards tends to be clustered. In cases where content encoding standards are enforced, transformations are more readily available compared to standard having more flexibility in terms of encoding. For example, integration of metadata from ISO 19139, Dublin Core, and OGC CSW Records tend to be supported together whereas less restricted standards (or more flexible, depending upon one’s perspective) such as SPARQL and OpenSearch may not be readily integrated across systems.

By defining a core (minimum) set of search capabilities and metadata content, it would be possible to further integrate search across the catalogue and metadata standards employed within the marine community and to identify gaps in metadata content required to facilitate organization of resources within the Ocean Data and Information System.

Beyond the standards already employed within the marine science community, with the evolution of the internet and search provider tools (e.g. APIs) and services it is the right time to engage internet search providers in the use of their resources to supplement the discovery of data and information that are already known through ‘self-registration’ facilities. Working directly with search engine providers also provides an opportunity to improve search results from search provider home pages by focusing efforts to improve discovery of information served through the Ocean Data and Information System.

Recommended partners would include:

1. **IODE OceanKnowledge Platform, Ocean Sites, and Galway Knowledge Platform work package** - Between the two active knowledge platform efforts and program leads from within the IODE and JCOMMOPS, it should be possible to develop a levels of minimum metadata content requirements for data and information resources (minimum view, advanced view), and these requirements would be supplemented through IODE participation in other semantic web working groups.

2. **GEOSS, MOWIS, IODE Ocean Data Portal, SeaDataNet/SeaDataCloud, PANGAEA/WDS** - There are many established metadata brokering solutions in the data-space, and it is recommended that one or more existing technologies be evaluated as a candidate to support this capability within the Ocean Data and Information System. Taking a technology such as GI-cat, the underlying technology used in the delivery of the GEOSS common broker
which is also the technology employed within the metadata brokering prototype under the ODIP project, would be one logical option to assess. Evaluating the scalability of GI-cat to the global scale (the existing GEOSS common broker is already supporting brokering from numerous systems) and extending its capabilities to metadata content standards such as that employed by PANGAEA is one of the options which should be considered for handling the high volume of metadata arising from the catalogues that would feed the Ocean Data and Information System. Furthermore, the approach employed by systems such as WMO/WIS or the Ocean Data Portal to implement layers of aggregating and linked catalogues, as well as the efforts of SeaDataCloud to implement a centralized European catalogue system in the cloud, both represent other relevant options to be considered at the global scale.

3. **Microsoft and/or Google and/or other search providers** - Search providers now offer many options for leveraging working with and transforming web content indexed by the provider, including machine learning and other advanced capabilities which can employed by customers from within their own systems. By engaging search providers, it may be possible to identify and connect data and information sources not captured through other inventory exercises.

**Implementation strategy:**

It is expected that the content broker employed by the Ocean Data and Information System will be based upon, or developed through an extension to existing data brokering technology, as mature options exist which deliver this capability. It is also expected that the implementation supporting the Ocean Data and Information System may be separate from existing implementations to ensure existing regional or national systems are not impacted by additional workload originating from the Ocean Data and Information System. In terms of implementation strategy, the approach that will be taken for the implementation of the metadata brokering/mediation tier will follow a more typical agile development lifecycle, developing through iteration, with key milestone releases being implemented in production.

The metadata brokering/mediation tier seeks to integrate support for structured and unstructured content. Because we are dealing with a highly variable landscape where there is no uniform level of content standards, accessibility, or technical interoperability, the goal will be to implement a ‘best experience’ based on the state of the underlying content, and it is expected that the capabilities in terms of brokering unstructured content will evolve over time.

The existing knowledge platform will work with teams already delivering metadata and other content brokering systems to develop the functional requirements for the enhanced brokering tier required by the Ocean Data and Information System in areas such as:

- request (query) and response capabilities
- content requirements
- integration of semantic resources (ontologies) into query process
- integration of multi-lingual support through interoperability with translation APIs
- support for querying unstructured resources through search provider APIs
5.3 Data brokering/access

In terms of data access, connecting with data and information providers should be implemented at highest (aggregate) levels possible, leaving the mobilization of individual data and information providers, implementation and management of IT and related solutions, as close to source as possible and supported operationally and not through special project funds. This approach acknowledges the operational nature of this function/infrastructure, and ideally, will not involve the development of this tier within this project. Instead, the feeding of the Ocean Data and Information System is achieved through negotiation and establishment of interoperability relationships with existing systems.

It is important to stress that although the Ocean Data and Information System would ideally federate across higher level aggregating systems, it is recognized that not all data and information providers are sufficiently connected to such services. Therefore, it is critical to engage providers at the national level in order to verify their state of connectedness and to be prepared to implement connections to national systems in order to provide an inclusive environment. As with the Ocean Data and Information System itself, it is expected that the connections between national, regional, and global systems will evolve over time and that these efforts will also provide opportunities to implement levels of harmonization and possibly standardization.

In the longer term, attempting to foster connections at the aggregator level will simplify the data architecture for the Ocean Data and Information System, and it is much more likely that the data networks supported working with the data aggregators with ongoing operational funding will be in a better position to implement any optimization and failover required to provide the high performance, reliable, and scalable environment needed to feed the Ocean Data and Information System.

Recommended lead/partners would include:

1. **MOWIS, SeaDataNet/SeaDataCloud, IODE Ocean Data Portal** – In terms of data brokering and access, WMO/WIS and SeaDataCloud are examples of large-scale infrastructure projects serving data providers at regional and global scales. Working through a multi-tiered network of centres supported through a combination of dedicated networking infrastructure (varying capacities by class of WIS centre) software infrastructure, and content specifications, processes, and dedicated staff. The technical capacity and expertise underlying the implementation of WIS may be relevant to both the Ocean Data and Information System envisioned by IOC, as well as the implementation of data brokering and access solutions for regions lacking capacity to participate using current systems. SeaDataCloud focus on current technologies and cloud based architecture and service delivery provides expertise in areas of IT now becoming the norm. Within the IODE, the Ocean Data Portal also supports multiple layers of capability ranging from simple data provider up to full node implementation providing the same distributed network capabilities available from the ODP global node. The packaging of the ODP technology components as a set of loosely coupled Virtual Machine mages also supports simple and flexible deployment of the technology stack.

2. **IODE and JCOMMOPS information and data systems, EMODNet** - Theme-base systems such as the OBIS, the EMODNet thematic portals, and JCOMMOPS data products (from GOOS, Argo, and other networks) are examples of data and product sources that do not represent full data centres, but which have a standing history in terms of client focus and operations, and which represent end-user focused contributions to the Ocean Data and Information System.
3. **PANGAEA/WDS** – Beyond traditional data centres, other speciality data publishers such as PANGAEA exist which support the publication of data associated with research literature. PANGAEA is one example of such a system and there are other organizations providing similar services. Unlike more traditional data systems, these services focus on securing and stewarding data, and in some cases, associated literature.

4. **Regional and national systems** – IODE NODCs, ADUs, and other regional and national systems providing data aggregation services are key stakeholders in the implementation of the Ocean Data and Information System. Ideally, connections are made with data aggregators, but that is not always possible as there may be systems in existence which are not connected with any higher-level data aggregator.

**Implementation strategy:**

Data harmonization studies have been completed at the program level in the past, and even at this level the effort required to enable harmonized data access can require significant investment. Looking beyond individual data and information networks, the complexity and effort only increase. The strategy proposed to be employed by the Ocean Data and Information System is to focus primarily on promoting access to data and information in open and interoperable forms, leaving the communities of practice to develop community profiles on these open standards.

5.4 **Data assembly, submission, and provider support**

The front-line assembly and submission of data and information into systems which feed the Ocean Data and Information System envisioned by IOC is a critical and enormous task which cannot be undertaken by IOC alone. These enabling functions require investment of national and regional organizations in the IOC and non-IOC communities. Although not funded by this project, this operational support is the first step to participating in the Ocean Data and Information System envisioned by IOC.

The levels of service required by the different classes of providers and user described earlier in this document must be taken into account, along with the disparities which act as barriers to participation. Otherwise, if participation is difficult or impossible, it is likely that resource providers and users will develop siloes elsewhere or not participate at all.

In terms of core capabilities, data and information providers require support in:
- Securing raw data and information to ensure there is no risk of loss
- Incorporating data and information into the IT systems which ultimately make them available, including applying standards/transformations that may be beyond the capabilities of the provider
- Enabling high performance and reliable access to data and information and managing appropriate technical infrastructure to deliver this capacity (or working with enablers who perform these functions)

Although data assembly and submission is not directly linked to the Ocean Data and Information System, the regional disparities raised in the audit directly impact to completeness of the global view. Challenges such as limited access to long term data archives and support can occur even in well-resourced regions. There is a need to ensure that data and information provider are providers have access to long term archives, and that submission and other processes and standards themselves do not create new barriers.
Beyond the need to ensure data and information providers have a mechanism to participate, the greater challenge is the proliferation of systems and data, and the downstream impacts this has on data and information users. As was described earlier in this document, it is not realistic to build a single data and information system to hold all content related to the marine sciences. However, being deliberate when choosing to build a new system over enhancing an existing one, and taking care when ingesting data and information from other sources (by managing provenance) will facilitate the identification of the individual versions of these resources.

In the longer term, employing linked-data strategies to access data and information from source, rather than duplication of content. However, this level of de-duplication and ensuring a high performance and reliable service comes at the costs of computing and network infrastructure.

Recommended lead/partners would include:

1. **Data and information aggregators and publishers** – The centres actively aggregating and publishing data represent key nodes facing the Ocean Data and Information System. These entities may be assembling data coming from IOC and non-IOC centres and their capabilities (e.g. infrastructure, service availability, people support, etc.) will directly impact the upstream content brokering and performance of the Ocean Data and Information System.

2. **Data assembly centres (IOC and non-IOC)** – The centres providing frontline assembly and archival/management support are a critical to enabling data and information providers to participate in the Ocean Data and Information System. It will be important to establish strong connections to these centres, either directly or through higher level data networks (e.g. AtlantOS, WIS) and to proactively identify and close gaps in terms of regional, thematic, and other coverage. Within the IODE context, this is an area where extending the roles of existing NODCs and ADUs, implementation of new ADUs or facilitating the flow of data into established non-IOC systems.

3. **IT Infrastructure partners** – Dedicated IT service providers exist, both as independent entities and as embedded enablers within larger organizations. These service providers may provide support across sectors or they may be focused on specific commercial, academic and research, or government clients. Furthermore, the available service offerings and associated costs provided by these organizations may vary significantly by client sector. With the desire to span these sectors, opportunities must be sought to foster the ‘right’ agreements with infrastructure service providers in order to gain access to required capacity while managing costs. Systems such as WIS have been built upon high performance, fault tolerant, highly available infrastructure and opportunities must be sought to extend these practices into domains that have been traditionally thought of as pre-operational, thus not requiring the same level of performance and availability. The JCOMM Ocean Information System (OIS) to be implemented under MOWIS will follow a very similar model to WIS and will also be architected to support operational activities. With the ever-increasing volumes of data and information being collected and generated, coupled with the increasing demands for global (e.g. 24/7) access to data and information, requires the pre-operational domains to think more operationally. These operational environments will be hosted on high capacity and connected infrastructures supported by dedicated IT service providers at various level.

4. **Data and information producers** – Data and information producers must be engaged in order to encourage their participation in the global community and to establish connections with the assembly centres, aggregators, and brokers which provide support from long term stewardship of data and information through access and dissemination.
Coupled with assessing the state of data assembly centres, it will be important to identify and address gaps and barriers which impact the ability to participate. Otherwise, it will be impossible to establish the Ocean Data and Information System with a truly global view.

**Implementation strategy:**

The broad range of stakeholders anticipated to contribute to the Ocean Data and Information System will be engaged and supported through a variety of mechanisms. It is also expected that changes required to address identified gaps will be implemented through phases, and likely through parallel activities undertaken by different teams. The lack of an existing unified coordination function will prove to be challenging, and is an area where the IOC and non-IOC communities must come together to ensure the data and information provider community is fully engaged.

Within the IOC context, it would be useful to assess the current state of the member states in terms of data assembly centres, identifying known marine data and information systems, and to have members self-identify any known capacity and related challenges. If the IOC is not able to resolve issues within the member states, the developing a complete global view is impossible. Furthermore, as connections are made between national and regional systems and the Ocean Data and Information System, new capacity challenges. The IODE, in co-operation with the NODCs and ADUs, will need to seek opportunities to address these challenges within the IOC community.

**6. Implementation plan**

It is proposed that the implementation plan for the Ocean Data and Information System be executed in phases, with preliminary investigations to determine the feasibility and cost-benefit analysis of proceeding.

**Foundation activities:**

These foundation activities form the basis for any well governed project:

- **Project management** – The project management function will encompass activities such as work plan development and monitoring, reporting and communications back to IOC and non-IOC stakeholders, management of risks and issues, and development of other project management artefacts.

- **Stakeholder engagement** – The delivery of the project, from feasibility and cost-benefit assessment though implementation of the Ocean Data and Information System, will rely heavily on engagement with IOC and non-IOC stakeholders (data and information system providers, users of the Ocean Data and Information System, collaborators supporting the implementation of the Ocean Data and Information System). As the requirements for the Ocean Data and Information System, supporting data and information access, semantics, and infrastructure capacity evolve, there will be a need to enhance the underlying systems contributing to the Ocean Data and Information System, implementing change through a combination of operational and project level investment.

**Establish implementation team:**

In order to implement a reliable and sustainable global system, the structure of the implementation team will be critical. As described through Section 5 in this document, there are established marine data and information systems, and other relevant activities which will contribute to the delivery of the Ocean Data and Information System. The team will include leads in the following areas:

- **Infrastructure implementation** - It is recommended that attempts be made
to secure support from a partner already supporting a regional or global scale data and information system. A strong infrastructure partner would leverage existing infrastructure capacity and support regimes, reducing costs and improving sustainability of the Ocean Data and Information System. The WMO/WIS team is an example of a global scale partner, though it is possible that another IOC partner may take a leadership role.

- **Metadata/content brokering** - Metadata and content brokering partners will span data and information, and the choice of partners depending in part on the results of feasibility and cost-benefit analyses. It is expected that the implementation lead will be the team supporting an existing brokering technology such as GI-cat, working with service/API endpoint into an internet search provider’s search engine.

- **Portal implementation** – There is some flexibility in terms of the portal implementation team as the majority of complexity is handled at the brokering level. That being said, it is expected that the IODE OceanKnowledge Platform and Galway Knowledge Platform work package will be instrumental in defining the scope of the portal in terms of functionality, etc. and that these activities will also participate in the evaluation and acceptance of the portal delivered through the project.

**Feasibility study and cost-benefit analyses:**
It is recommended that a preliminary assessment be completed evaluate capabilities, scalability, and other features of existing brokering solutions, as well as to evaluate options for integrating discovery and access for unstructured information. The results of this assessment will allow the team to refine the scope for the Ocean Data and Information System.

In terms of the technical evaluation in support of metadata brokering, the following areas would be assessed:
- **Ability to connect to catalogues/metadata brokers built upon accepted metadata content standards**
- **Ability to execute federated simple and advanced searches across catalogues and brokering systems via metadata brokering technologies considered for the Ocean Data and Information System**
- **Ability to implement brokering technology in a fault tolerant mode**

In terms of the technical evaluation in support of brokering of unstructured content, the following areas would be assessed:
- **Availability of API and/or service endpoints that would facilitate system to system access from the broker for the Ocean Data and Information System**
- **Access to internet search provider data focused on the marine community**
- **Ability to extract discovery level metadata content from internet search provider query results**
- **Ability to execute simple and advanced searches against internet search provider data, along with assessment of limitations in query support**

It is also recommended that a cost benefit analysis be performed to determine potential infrastructure, development, implementation, and maintenance costs of the Ocean Data and Information System prior to initiating the project in full.
If the results of the technical feasibility and cost/benefit analyses are determined to be acceptable, the project will then proceed.

Establish scope for the Ocean Data and Information System:
Following the technical assessment and cost/benefit analysis, the full scope of the Ocean Data and Information System will be defined. The planned user experience and capabilities and user experience will be defined and this in turn will drive the underlying technical environment.

Working with the knowledge platform teams, the scope of the Ocean Data and Information System will be defined in areas such as:

- **Content coverage** – Scope of information to be integrated into the Ocean Data and Information System versus that which will simply be discoverable from the Ocean Data and Information System.
- **System coverage** – Systems for which interoperability arrangements will be established and relative priorities for implementation.
- **Search, access, and visualization capabilities** – Core (mandatory) search, access, and visualization capabilities to be employed in the Ocean Data and Information System will be defined, considering the results of feasibility and cost/benefit analyses completed at the start of the project.

Ocean Data and Information System implementation:
The Ocean Data and Information System will be implemented by teams responsible for IT infrastructure, data and information brokering, and the Ocean Data and Information System. These efforts will be supported through stakeholder (data and information providers and users, knowledge platform initiatives, IODE) engagement and feedback.

Enabling support:
The successful implementation of the Ocean Data and Information System providing harmonized access to data and information across systems will require development in several supporting areas relevant to enabling data and information providers to participate in the Ocean Data and Information System:

- **Semantic development** – Although there may a need to develop specific ontologies to support the classification of resources within the Ocean Data and Information System, it is expected that most semantic development will focus on adopting existing semantic resources and providing input into their ongoing development.
- **Data and information publication best practices** – As interoperability arrangements are established with individual data and information systems, feedback will be provided back to contributing systems in areas such as search and access capabilities, limitations in terms of ability to provide harmonized access to content, and so on. This feedback is seen as a mechanism improving overall accessibility of data and information through improvements at both the system level and to content standards and publication practices employed by the individual data and information systems.
- **Data and information provider support** – As an ongoing supporting activity, efforts must continue to be made to support data and information providers and their ability to participate in the Ocean Data and Information System. This work is not directly linked to this project, but falls within the role of IODE NODCs, ADUs, and other non-IOC structures. As stated in the audit, there are regional disparities which may require some support from the IOC.
7. Conclusion

The marine data and information landscape is broad and includes stakeholders linked and not linked to the IOC. Although disparities exist on many levels, and current levels of integration are mixed, with the advances in technical and content standards, there is a basis on which a level of harmonized discovery can be explored and developed. Re-casting the role of IOC in the engagement of marine data and information system providers is a pathway to seeking collaboration to improving the openness and interoperability of marine data and information. Investigating the capabilities of search provider service offerings is a key enabler to improving the discovery and linking of structured and unstructured data.

Undertaking this project will raise other related challenges to be resolved:

- The duplication of resources (and different versions of individual resources) across systems will become obvious within the Ocean Data and Information System, requiring the development of strategies to handle this complexity.
- The improvement of semantics and use of PIDs will be required to provide effective cross system query and content linking.
- There will be a need to provide support for data and information producers with limited/no capacity to participate.
ANNEX 1
Ocean Data and Information System – Conceptual Architecture
ANNEX 2
Current and past activities relevant to the Ocean Data and Information System

Information (knowledge) portals

OceanKnowledge Platform

The IODE OceanKnowledge Pilot Project commenced work in 2015. In the first two year phase it proposes to harness the power of Semantic Web to provide a demonstrator single point discovery platform across the diverse IODE products and services that vary from information on individuals and documents to physico-chemical and biological data. The system architecture is proposed to be scalable to accommodate the ultimate OceanKnowledge vision to offer single point discovery across all global ocean-related websites, databases and repositories.

The complexity of the underlying semantic technology, standards, interoperability, linked data and information exchange between content providers and their microdata layers requires innovative high tech solutions to pull together the information model.

An initial Project Plan has been produced that defines use cases and provides a start point system diagram of linkages between underlying data stores with a discovery layer on top. Interest in the project from IOC and LME Learn has identified specific reporting needs that have derailed OKn progress, but the main curtailment of progress is the need for project resource investment (additional ICT developers) to take this forward to at least a prototype stage during 2017/8.

The OceanKnowledge Project vision maps to the IOC Audit concept of the Ocean Data and Information System and the proposed demonstrator could contribute or be one of its building blocks.


ICES leads work package 11 for AORA-CSA, the Atlantic Ocean Research Alliance Coordination and Support Action.

**Lead:** Neil Holdsworth, Head of data and information, ICES

**Objective:**
- to establish a long-term knowledge sharing platform in six identified research priority areas
- to allow for long-term usability of the data, information and knowledge products thereby generating real value from the investment in infrastructure
- to include a classification system, which allows for a simple, focused and reliable use and analysis of the information made available through the knowledge platform.
- to enhance the exploitability of the platform for policy making and stakeholder purposes, representatives from funding agencies and these stakeholder communities will be consulted in the design
- the principle of open access will underpin the platform.
Cataloguing efforts

**IODE OceanPortal** was a web-based high-level directory database of ocean related websites. Its objective was to promote improved access by assisting a variety of users (policy makers – scientists - citizen scientists etc.) interested in ocean science, to locate ocean data and information websites and databases. Additionally, to promote exchange of information and experience of this type of web information between users. Records were indexed under eight super-categories and 251 specific subject categories. In 2007 the following metrics were offered:

- 4600+ records
- +/- 1000 records added
- +/- 130,000 visits per year, +/- 500 visits per day
- 92,000 citations in Google

It was a highly regarded international source-of-choice for marine site information, and was one of the most visited of all IOC websites.

OceanPortal exposed global web sources but there were also regional ocean portals: PorticoOceanico (product of ODINCARSA) & African OceanPortal (product of ODINAfrica)

OceanPortal was closed in 2009 because it was decided that there was then sufficient other web-based search engines to seek ocean related web sites. In addition, the work of keeping the inventory current by discovering new ocean related websites and the maintenance of web links became labour intensive.

Annex of IODE website/portal/database inventories (from audit)

### Annex 4: Databases and information systems managed by or related to IODC

<table>
<thead>
<tr>
<th>NOC</th>
<th>Structural Element/Project</th>
<th>Number</th>
<th>Database/Information System Type</th>
<th>Database/Information System Name</th>
<th>URL/Online Access</th>
<th>Modeled by NOC (Project Office)</th>
<th>Date of Creation</th>
<th>Funding Source</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOCDE</td>
<td>National</td>
<td>63</td>
<td>Online databases</td>
<td>Online database</td>
<td>IOCDE</td>
<td>Yes</td>
<td>National</td>
<td>1061</td>
<td>National</td>
</tr>
<tr>
<td>IOU</td>
<td>National</td>
<td>20</td>
<td>Online databases</td>
<td>Varied</td>
<td>Varied</td>
<td>Varied</td>
<td>National</td>
<td>2013</td>
<td>National</td>
</tr>
<tr>
<td>ISIC</td>
<td>National</td>
<td>2</td>
<td>Online databases</td>
<td>Varied</td>
<td>Varied</td>
<td>Varied</td>
<td>National</td>
<td>2013</td>
<td>National</td>
</tr>
<tr>
<td>SeeCaribbean (IOC-Caribbean)</td>
<td>Regional</td>
<td>1</td>
<td>Online databases</td>
<td>SeeCaribbean</td>
<td><a href="http://seecaribbean.org">http://seecaribbean.org</a></td>
<td>Yes</td>
<td>International</td>
<td>2007</td>
<td>International</td>
</tr>
</tbody>
</table>

**European Marine Observation and Data Network (EMODnet)**

EMODnet is a European Commission funded project that aims to coordinate marine data collection and management across Europe. The project is designed to improve the accessibility, usability, and interoperability of marine data across Europe, facilitating better decision-making and supporting research, innovation, and environmental management. EMODnet is a collaborative initiative involving multiple stakeholders from the European Union and beyond, working together to create a comprehensive, accessible, and sustainable marine data infrastructure. The project focuses on integrating and sharing marine data from diverse sources, ensuring that data is available and accessible to all stakeholders, including scientists, policymakers, and the general public. By doing so, EMODnet aims to support the sustainable use of marine resources and promote the development of new technologies and innovative practices in the field of marine observation and data management.

**WODC project**

WODC is a project led by the UNESCO World Oceans database (WOD) and hosted by the IOC. It aims to establish a unique and comprehensive dataset of oceanographic observations from various sources, including satellite data, ocean models, and in situ measurements. The project focuses on integrating and sharing marine data from diverse sources, ensuring that data is available and accessible to all stakeholders, including scientists, policymakers, and the general public. By doing so, WODC aims to support the sustainable use of marine resources and promote the development of new technologies and innovative practices in the field of marine observation and data management.
<table>
<thead>
<tr>
<th>Project Name</th>
<th>Database Type</th>
<th>Database Name</th>
<th>URL</th>
<th>Date</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aquatic Command</td>
<td>Online database</td>
<td>Aquatic Command</td>
<td><a href="http://www.aquaticcommand.org">http://www.aquaticcommand.org</a></td>
<td>2017</td>
<td>IOC/IODE Data Portal</td>
</tr>
</tbody>
</table>

The IOC/IODE Data Portal project will be the environmental framework for the compilation and dissemination of biological sciences. The data will be made available via the IOC database.

The IOC/IODE Data Portal project has been developed to meet the needs of the IOC/IODE Data Portal, which is a comprehensive set of databases and services that provide access to marine and oceanographic data. The project is funded by the European Commission (EU) and the IOC/IODE Data Portal. The project aims to provide a comprehensive set of databases and services that facilitate access to marine and oceanographic data.

The IOC/IODE Data Portal project has been developed to meet the needs of the IOC/IODE Data Portal, which is a comprehensive set of databases and services that provide access to marine and oceanographic data. The project is funded by the European Commission (EU) and the IOC/IODE Data Portal. The project aims to provide a comprehensive set of databases and services that facilitate access to marine and oceanographic data.
<table>
<thead>
<tr>
<th>Project Name</th>
<th>Type</th>
<th>Description</th>
<th>Key Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caribbean Marine Atlas</td>
<td>Online atlas with comprehensive and interactive</td>
<td>FOA</td>
<td>Development of a comprehensive digital atlas of the Caribbean marine environment, providing information on marine resources, biodiversity, and conservation areas.</td>
</tr>
<tr>
<td>Ocean Knowledge Portal</td>
<td>Online portal</td>
<td>FOA</td>
<td>Development of an online portal for marine knowledge sharing, designed to facilitate data access and collaboration among scientists.</td>
</tr>
<tr>
<td>Virtual Field Learning Platform</td>
<td>Online platform for interactive learning</td>
<td>FOA</td>
<td>Development of an interactive learning platform for virtual field trips, enhancing educational experiences for students and professionals.</td>
</tr>
<tr>
<td>Harmful Algae Information System (HARMIS)</td>
<td>Online database</td>
<td>FOA</td>
<td>Development of an online database for harmful algae information, supporting research and monitoring efforts.</td>
</tr>
<tr>
<td>Oceanography and Information Exchange System (ODISY)</td>
<td>Online initiative</td>
<td>FOA</td>
<td>Development of an initiative for exchanging oceanographic data and information, facilitating international collaboration.</td>
</tr>
<tr>
<td>Integrated Coastal Area Management (ICAM)</td>
<td>Online platform and data</td>
<td>FOA</td>
<td>Development of an integrated platform for coastal area management, integrating data from various sources for comprehensive planning and decision making.</td>
</tr>
<tr>
<td>Global Ocean Observation System (GOOS)</td>
<td>Integrated GOOS overview</td>
<td>FOA</td>
<td>Overview of the Global Ocean Observation System, a collaborative framework for global ocean data and information.</td>
</tr>
</tbody>
</table>

**Key Contributions:**
- FOA: includes national and international contributions.

**Notes:**
- The table above provides a summary of various initiatives and projects focused on oceanography and information exchange. Each entry describes the type of initiative, its description, and the key contributions made through national and international collaborations.
<table>
<thead>
<tr>
<th>Metadata system</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOC/IODE</td>
<td>online database</td>
<td>Used to store oceanographic and meteorological data, including forecasts and real-time data.</td>
</tr>
<tr>
<td>IOC/IODE</td>
<td></td>
<td>- Interactive data visualization, allowing users to access and analyze data in real-time.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Interoperability</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOC/IODE</td>
<td>- Interoperability is crucial for data sharing and analysis across different systems.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual Network Data Systems</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOC/IODE</td>
<td>- Integration of individual network data systems allows for a unified view of global data.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TAPF Ocean Port</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOC/IODE</td>
<td>- TAPF Ocean Port serves as a central hub for accessing oceanographic data.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tsunami Data</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOC/IODE</td>
<td>- Tsunami data is critical for predicting tsunamis and preparing for response.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tsunami Science Institutions</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOC/IODE</td>
<td>- Tsunami science institutions are responsible for monitoring and predicting tsunamis.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tsunami Early Warning Centers</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOC/IODE</td>
<td>- Early warning centers provide timely and accurate tsunami alerts to coastal areas.</td>
</tr>
</tbody>
</table>
### ANNEX 3

#### Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADU</td>
<td>Associated Data Unit</td>
</tr>
<tr>
<td>AIU</td>
<td>Associated Information Unit</td>
</tr>
<tr>
<td>BODC</td>
<td>British Oceanographic Data Centre</td>
</tr>
<tr>
<td>CSV</td>
<td>Comma-separated values</td>
</tr>
<tr>
<td>CSW</td>
<td>Catalog Service for the Web</td>
</tr>
<tr>
<td>EDMO</td>
<td>European Directory of Marine Organizations</td>
</tr>
<tr>
<td>EDMODNet</td>
<td>European Marine Observation and Data Network</td>
</tr>
<tr>
<td>EML</td>
<td>Ecological Markup Language</td>
</tr>
<tr>
<td>FTP</td>
<td>File Transfer Protocol</td>
</tr>
<tr>
<td>GCMD</td>
<td>Global Change Master Directory</td>
</tr>
<tr>
<td>GI Cat</td>
<td>NOT ACRONYM</td>
</tr>
<tr>
<td>Gisc</td>
<td>Global Information System Centre</td>
</tr>
<tr>
<td>GEOSS</td>
<td>Global Earth Observation System of Systems</td>
</tr>
<tr>
<td>GOOS</td>
<td>Global Ocean Observing System</td>
</tr>
<tr>
<td>HTML</td>
<td>HyperText Markup Language</td>
</tr>
<tr>
<td>ICES</td>
<td>International Council for the Exploration of the Sea</td>
</tr>
<tr>
<td>INSPIRE</td>
<td>Infrastructure for Spatial Information in Europe</td>
</tr>
<tr>
<td>ODINAfrica</td>
<td>Ocean Data and Information Network Africa</td>
</tr>
<tr>
<td>ODINCARSA</td>
<td>Ocean Data and Information Network Caribbean and Latin and South America</td>
</tr>
<tr>
<td>IOC</td>
<td>Intergovernmental Commission</td>
</tr>
<tr>
<td>IODE</td>
<td>International Data and Information Exchange</td>
</tr>
<tr>
<td>ISNI</td>
<td>International Standard Name Identifier</td>
</tr>
<tr>
<td>ISO</td>
<td>International Standards Organization</td>
</tr>
<tr>
<td>IT</td>
<td>Information technology</td>
</tr>
<tr>
<td>LOD</td>
<td>Linked Open Data</td>
</tr>
<tr>
<td>JCOMM</td>
<td>Joint Technical Commission for Oceanography and Marine Meteorology</td>
</tr>
<tr>
<td>NCEI</td>
<td>National Centers for Environmental Information (USA)</td>
</tr>
<tr>
<td>NetCDF</td>
<td>Network Common Data Form</td>
</tr>
<tr>
<td>NODC</td>
<td>National Oceanographic Data Centre</td>
</tr>
<tr>
<td>OAI-PMH</td>
<td>Open Archives Initiative Protocol for Metadata Harvesting</td>
</tr>
<tr>
<td>OBIS</td>
<td>Ocean Biogeographic Information System</td>
</tr>
<tr>
<td>ODIP</td>
<td>Ocean Data Interoperability Platform</td>
</tr>
<tr>
<td>OGC</td>
<td>OpenGIS Catalogue Service</td>
</tr>
<tr>
<td>ORCiD</td>
<td>Open Researcher and Contributor ID</td>
</tr>
<tr>
<td>PANGAEA</td>
<td>NOT ACRONYM</td>
</tr>
<tr>
<td>PHP</td>
<td>scripting language (originally ‘Personal Home Page’)</td>
</tr>
<tr>
<td>PID</td>
<td>Persistent Identifier</td>
</tr>
<tr>
<td>RDF</td>
<td>Resource Description Framework</td>
</tr>
<tr>
<td>SCOR</td>
<td>Scientific Committee on Oceanic Research</td>
</tr>
<tr>
<td>SPARQL</td>
<td>NOT ACRONYM</td>
</tr>
<tr>
<td>SRU</td>
<td>Search/Retrieve via URL</td>
</tr>
<tr>
<td>TT-MOWIS</td>
<td>Task Team for Integrated Marine Meteorological and Oceanographic Services within WIS (JCOMM)</td>
</tr>
<tr>
<td>URI</td>
<td>Universal Resource Identifier</td>
</tr>
<tr>
<td>VIVO</td>
<td>NOT ACRONYM</td>
</tr>
<tr>
<td>W3C</td>
<td>World Wide Web Consortium</td>
</tr>
<tr>
<td>WDS</td>
<td>Wireless distribution system</td>
</tr>
</tbody>
</table>
WIS - WMO Information System
WMO - World Meteorological Organization
WOD - World Ocean Database
WoRMS - World Registry of Marine Species
XML - Extensible Markup Language
ANNEX 4
Reference Documents

