

AUSTRALIAN OCEAN DATA NETWORK (AODN)

“Publicly funded data, publicly available”

Roger Proctor, Patrick Gorringer, Kate Roberts

Integrated Marine Observing System, University of Tasmania, Private Bag 21, Tas 7004, Australia
roger.proctor@, patrick.gorringer@, kate.roberts@utas.edu.au

Abstract - The Australian Ocean Data Centre Joint Facility (AODC-JF) was established to provide a modern, whole-of-government approach to ocean data management in Australia. The AODC-JF vision was “To put in place, by June 2011, an interoperable, online network of marine and coastal data resources, including data from the six AODC Joint Facility partner agencies, supported by standards-based metadata, which will serve data to support Australia’s science, education, environmental management and policy needs: Australia’s digital ocean commons.”

Building upon the publicly funded, open source, and scalable information infrastructure of the Integrated Marine Observing System (IMOS) which includes the software tools, the IMOS Ocean Portal and the GeoNetwork Metadata Entry and Search Tool (MEST), this vision is now achievable.

This paper will articulate progress in establishing the AODN.

Keywords: information infrastructure, marine data network

I. BACKGROUND

The BlueNet project (www.bluenet.org.au), a pilot project of the Australian Ocean Data Centre Joint Facility (AODC-JF, with partners the Australian Antarctic Division, Australian Institute of Marine Science, Bureau of Meteorology, Commonwealth Science and Industrial Research Organisation Marine and Atmospheric Research, Geoscience Australia and Department of Defence (Royal Australian Navy Directorate of Oceanography and Meteorology)), laid the foundations for a cultural change in the sharing of marine data in Australia. By developing an exemplar search and discovery metadata tool (the GeoNetwork Metadata Entry and Search Tool (MEST), www.geonetwork.org) international (ISO) standard metadata records can be created which describe and give access to a diverse range of datasets, ensuring that the datasets are accurately and securely archived whilst at the same time made available to the community with appropriate legal access.

The Integrated Marine Observing System (IMOS, <http://www.imos.org.au>), funded by the Australian Government through the National Collaborative Research

Infrastructure Strategy and the Education Investment Fund, has the primary objective to observe the oceans around Australia to meet the national and international research needs. The strategic focus of IMOS is on the role of the oceans in the climate system, and the impact of major boundary currents on the continental shelf, ecosystems and biodiversity. Australia has one of the largest marine jurisdictions of any nation on earth [1], covering in excess of 4% of the world’s oceans. At over 14 million km² Australia’s Exclusive Economic Zone is nearly twice the surface area of the Australian continent. It extends from the tropics to high latitudes in Antarctic waters and much of it is unexplored.

The surrounding Pacific and Indian Oceans strongly affect the continental climate-system at all time scales, from seasons to decades. The major ocean currents on its eastern, western, northern and southern boundaries, best known of these being the East Australian Current (EAC) and the Leeuwin Current (Fig.1) affect regional climatic conditions and help sustain the marine ecosystems. There is evidence that these currents are changing on decadal time scales [2] and have already impacted marine ecosystems.

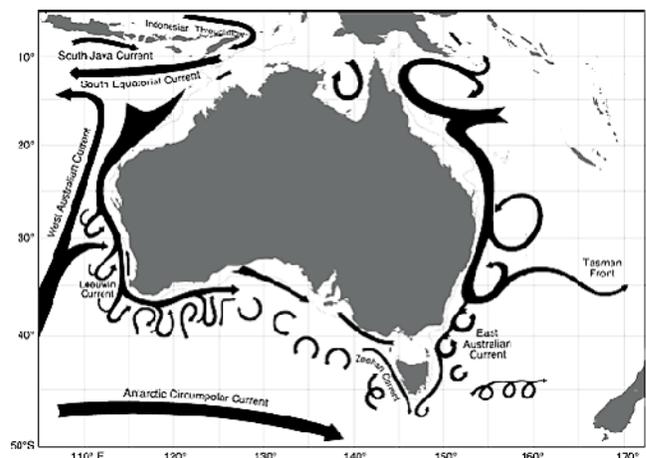


Figure 1 – boundary currents around Australia

The IMOS strategic research-goal is to assemble and provide free, open and timely access to streams of data that support research on:

- The role of the oceans in the climate system, and
- The impact of major boundary currents on continental shelf environments, ecosystems and biodiversity.

The scientific rationale for IMOS is set by five regional Nodes covering Queensland (QIMOS, including the Great Barrier Reef), New South Wales (NSWIMOS, southeastern Australia), Southern Australia (SAIMOS), Western Australia (WAIMOS), Tasmania and Bass Strait (TASIMOS), and the oceanic Bluewater and Climate (BWC) Node. The intention is to have a single, national IMOS with Node emphasis appropriate to regional characteristics of the marine system. Nodes are linked through integrated Science Plans (available on the IMOS website) addressing five major research themes

- Multi-decadal Ocean Change
- Climate Variability and Weather Extremes
- Major Boundary Currents and Inter-basin Flows
- Continental Shelf Processes
- Biological Responses.

streams in near real time is available through the IMOS Ocean Portal (Fig.2) accessible through the IMOS website.

Marine data and information are the main products of IMOS and data management is therefore a central element to the project's success. eMII provides a single integrative framework for data and information management that will allow discovery and access of the data by scientists, managers and the public. The initial strategy has focused on defining specific data streams and developing end-to-end protocols, standards and systems to join the related observing systems into an unified data storage and access framework.

The IMOS information infrastructure, a robust working infrastructure for end-to-end data management, provides search, discovery and access to IMOS data. The components consist of:

- A distributed data storage network
- A GeoNetwork metadata catalogue holding ISO 19115/19139 standard records
- An ocean portal for map-view access to the data (Fig.2),

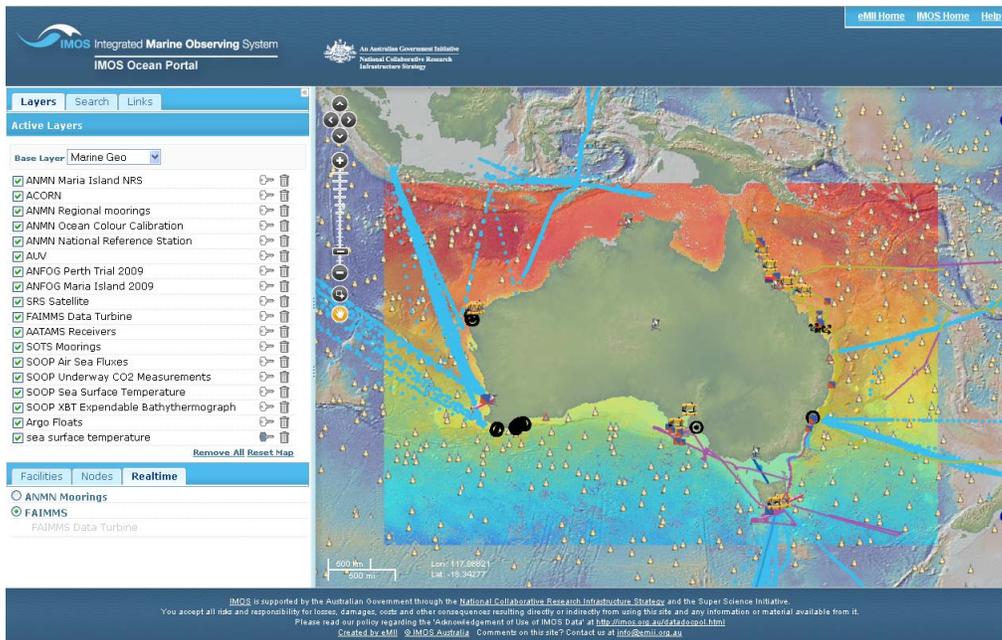


Figure 2 – View of the IMOS Ocean Portal

The eMarine Information Infrastructure (eMII), one of the original eleven IMOS Facilities, provides access to all IMOS data, enhanced data products, and web services in a searchable and interoperable framework. The value from this infrastructure investment lies in the nationally coordinated deployment of a wide range of equipment aimed at deriving critical data sets within a region that serve multiple applications. Free and open access to IMOS data

providing

- A national facility view of data
- A regional node view of data
- A real-time view of data
- A search and discovery view of data

Data may be downloaded either through the portal or from the MEST and all components are built on open source software and to international standards. A distributed data storage system has been developed in association with the

Australian Research Collaboration Service (ARCS). This has involved utilizing the ARCS Data Fabric, a ‘cloud’ storage system, i.e. the location of data across multiple platforms is invisible to the user. Data storage facilities have been set up at the regional high performance computing centres within five of the IMOS regional nodes (WAIMOS, SAIMOS, NSWIMOS, TASIMOS and QIMOS) linked through the Australian Academic and Research Network fibre optic backbone (10Gbit Bandwidth), with additional

Recognizing that a significant proportion of IMOS data is either gridded (satellite, high frequency (HF) radar) or in timeseries form (Argo, ships of opportunity, gliders, moorings, networked sensors) and could sensibly be written into a self describing format (CF compliant netCDF) meant that advantage could be taken of emerging web services to access these data through OPeNDAP/THREDDS servers. Both netCDF format data and non-netCDF format data (e.g. AUV imagery) can be accommodated within the ARCS

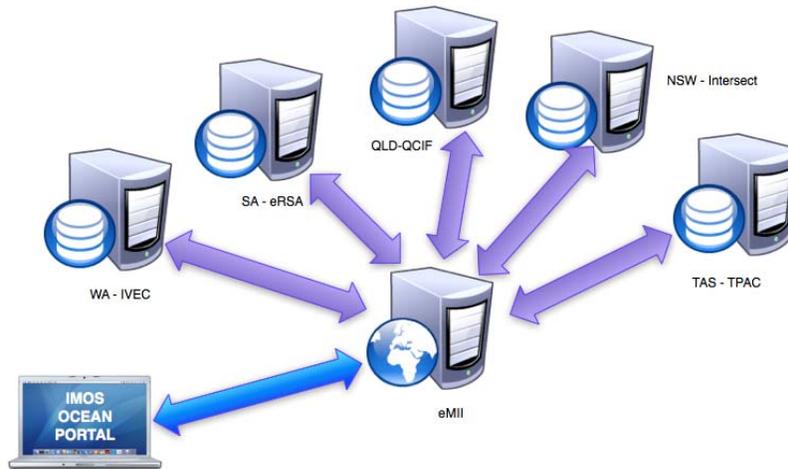


Figure 3 – IMOS distributed data system

storage and system monitoring in Tasmania at the Tasmanian Partnership for Advanced Computing. A schematic

Data Fabric.

The IMOS Ocean Portal was publicly released on 29 June 2009. The Portal, based on a Java ZK framework using OpenLayers, provides a map-centred view of IMOS data in

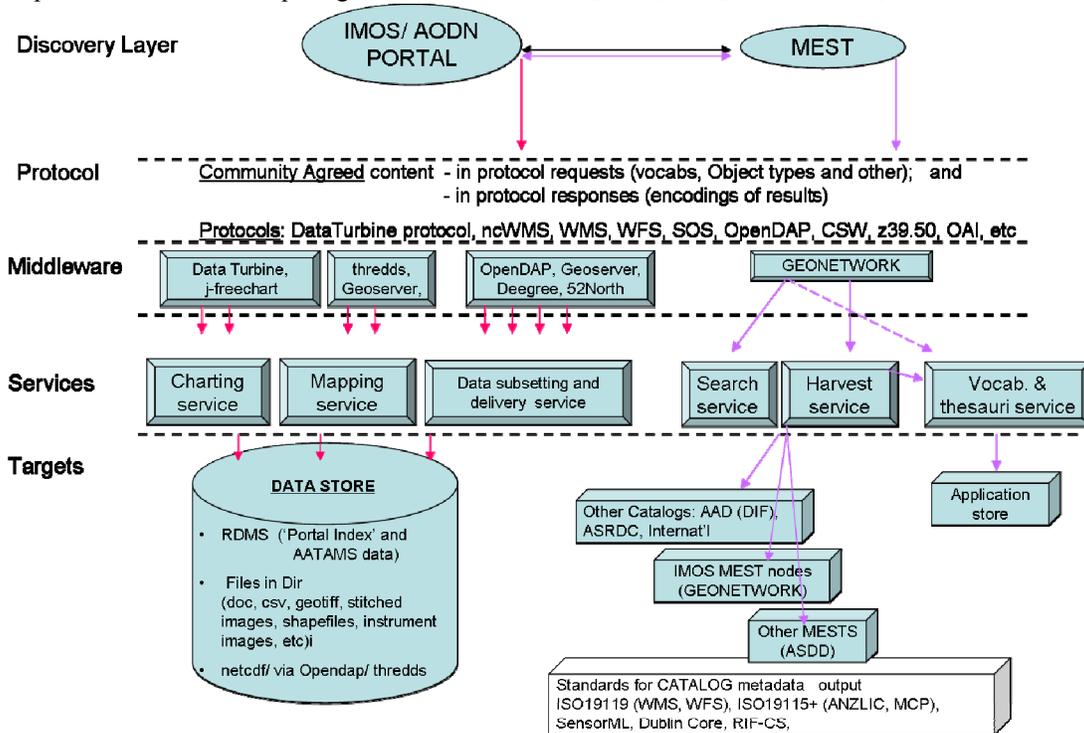


Figure 4 – IMOS Information Infrastructure

representation of the distributed IMOS data management system is shown in Fig. 3.

an intuitive way and provides easy access to the data, and access to search and discovery of IMOS and remote datasets via the IMOS MEST (a derivative of the BlueNet MEST). A detailed description of the IMOS Information Infrastructure is provided by [3] and the schematic shown in Fig. 4.

As can be seen in Fig. 4 the key services in this system are the charting and mapping services, data subsetting and delivery service, search service, harvest service and vocabulary and thesauri service. For the middleware we use a branch of the GeoNetwork trunk, the 'BlueNetMEST' branch v.1.3., locally known as the IMOS Metadata Entry and Search Tool (MEST) [<http://imosmest.aodn.org.au>] This allows harvesting of metadata from an array of sources (including WFS, fragments, files, OAI, THREDDS, OGC service metadata, thesauri), and is itself harvestable. Procedures for automated metadata creation, using the metadata standards: DC, ISO 19115-derived MCP -Marine Community Profile v.1.4, SensorML, ISO19119 (WFS, WMS), are being developed. Geoserver is the main middleware tool for Web Feature Services in the portal, although use is also made of Deegree (www.deegree.org). Real (or near-real) time data is being delivered using the DataTurbine (www.dataturbine.org) protocol and work is also in hand to deliver the timeseries data using the OGC Sensor Observation Service (e.g. via 52North, www.52north.org).

II. OPPORTUNITIES FOR DEVELOPING THE AUSTRALIAN OCEAN DATA NETWORK

Thus IMOS infrastructure offers the underlying architecture to provide a backbone for an expanding and evolving marine Australian ocean data network, a factor which inhibited the BlueNet project from fully contributing to the Australian Ocean Data Network (AODN).

The possibility of advancement of the AODN from its current (static) position was recognized by the Department of Innovation, Industry, Science and Research through the allocation of funds to establish an AODN Development Office, initially within eMII. This decision was subsequently reviewed by IMOS management with the outcome that the AODN Development Office is now the 12th Facility of IMOS with the AODC-JF Board acting as the AODN Development Office Steering Committee and the AODC-JF Technical Advisory Committee acting as the AODN Development Office Technical Advisory Committee.

The potential of the AODN Development Office for advancing the AODN has been recognised by the Australian National Data Service (ANDS) as the main conduit for delivering marine data to the Australian Research Data Commons (ARDC). ANDS directly supports the Development Office and additional funding has been secured from ANDS to enhance the delivery of real-time data from Research Vessels (see www.imos.org.au/aodn.html).

III. THE FUNCTIONS OF THE AODN DEVELOPMENT OFFICE

The AODN Development Office was established in July 2010 and intends to demonstrate, within twelve months, that it can operate as a central conduit for the exposure of marine data across Australia, thereby providing a firm foundation for the AODN to build on in future years. It is recognised that this is an ambitious task, requiring the support and goodwill of many agencies and individuals. However, there is reason to believe that the goodwill is forthcoming and that the time is right for publicly funded datasets to be publicly available. Utilising the information infrastructure developed by the IMOS project the AODN Development Office is, through developing contacts and interaction with the AODC-JF agencies, state government offices and agencies, universities and the industry sector, beginning the process of exposing marine data with fully supporting metadata to increase the utility of these data and enable the marine sector to make a significant step forward in establishing a marine data commons for Australia. This is seen as an important contribution towards the National Plan for Environmental Information announced in the April 2010 Federal Budget (www.environment.gov.au/npei/index.html) and IMOS / AODN Development Office are in communication with the relevant authorities.

The IMOS Information Infrastructure is based on national and international standards (many described above); a significant element of the Development Office work will be the promotion of these standards, and the adoption of other emerging standards. This will involve the Office in explaining to the community the importance of a standards-based approach, supporting the community in the adoption of those standards necessary to allow proper exposure of data, working with the community to develop standards where they do not exist or where the existing standards are lacking in respect to the data to be exposed. This essentially means working with agencies and individuals to identify what data is to be exposed, what resources are needed to expose it, and to work with agencies and individuals to put in place the systems software and metadata to enable the data to be published. Where-ever possible agencies will retain responsibility and custodianship of the data, the AODN Development Office will act as the 'oil' in the information publishing process. Where small groups or individuals are concerned, who perhaps do not have the capacity or resources to manage their own data securely, the AODN Development Office will offer to host the data. A critical role of the Development Office will be encouraging data providers to adopt common data models and rigorous quality control processes.

The AODN Development Office is constructing the AODN information system to enable access to these data. The main aims of the AODN information system are to

a) construct a catalogue of metadata records to enable efficient searching and discovery of marine data (this will primarily be populated by harvesting metadata records from distributed data repositories at host agencies, e.g. the AODC-JF agencies),

b) build a portal so that search, discovery and data download can be accomplished from a single access point, the ‘one-stop shop’ for marine data, and

c) ensure that these data are accessible as part of the Australian Research Data Commons, the ‘whole Australia’ data portal being developed by the Australian National Data Service.

The main objectives of the AODN Development Office are:

- Expand the quantity and scope of marine e-data in the data commons
- Promote and enhance the adoption of a data sharing and e-research culture
- Increase the number and diversity of contributors to and users of the network
- Enhance the infrastructure of AODC-JF and the visibility of AODC-JF data
- Enable the delivery of Research Vessel data to the AODN and hence the ARDC
- Provide an online repository of marine science data and information
- Consolidate a distributed archive facility, thereby consolidating the AODN

These objectives strive to:

- accelerate the availability of a significant portion of marine data available in Australian Commonwealth Agencies,
- accelerate the culture of data sharing away from the traditional self-interested and/or institutional approach towards a more open access e-Research approach;

thereby achieving the AODC-JF vision “To put in place, by June 2011, an interoperable, online network of marine and coastal data resources, including data from the six AODC Joint Facility partner agencies, supported by standards-based metadata, which will serve data to support Australia’s science, education, environmental management and policy needs: Australia’s digital ocean commons.”

Significant effort is being injected into the process of cultural change by demonstrating to the diverse marine community of Australia the benefits to be accrued from sharing data in a common and robust environment. At the same time it is recognised that support to enable groups or individuals to engage in this practice is required. This implies that personal interaction, as well as computing assistance, will be necessary.

Effort is also being directed at communication with the Australian marine science community. There are many aspects of communication which the Development Office

will embark upon. Initially the tasks will focus around developing engagement with stakeholders, linking data to users and users to data, and developing the culture of data sharing. These will involve: *Stakeholder engagement, and formulating an evaluation process for data ingest into the AODN; Establishing an AODN helpdesk; Establishing an AODN portal (<http://portal.aodn.org.au>); Conducting ‘Roadshows’ with AODN exemplars (e.g. IMOS data) and narratives of data sharing.*

Marine data is but one stream of environmental and socio-economic information. The Australian National Data Service seeks to encompass a common approach, and access to data, across this wide spectrum of information. The AODN Development Office, in partnership with ANDS, will

a) *further develop narratives and exemplars in promoting the “seeding of the commons” - the culture of sharing data is still novel to many disciplines in science, and some areas of marine science. Narratives and exemplars involving IMOS in cross-discipline studies will highlight the values of seeding the commons.*

b) *develop legal frameworks for deposit licences, use licences, and the responsibilities of data centres - whilst IMOS data is free without restriction to all users, it is recognized that not all datasets can be distributed in this way. Appropriate licensing for depositors and users needs to be on line. Creative Commons licences (<http://creativecommons.org.au/>) are increasingly seen as viable options for licensing the use of existing data sets; work with ANDS will focus on the deposit licences. The Development Office is working with ANDS to advise on choice of licence where necessary and will also assist in defining the responsibilities of data centres that hold shared data. The situation is made somewhat simpler by the recent announcement by the Australian Government Attorney General’s Department*

(http://www.ag.gov.au/www/agd/agd.nsf/Page/Copyright_CommonwealthCopyrightAdministration_Statement_ofPrinciplesforAustralianGovernmentAgencies) that all Australian Government agencies subject to the Financial Management and Accountability Act 1997 “should license their Public Sector Information (PSI) under Creative Commons BY license as the default license type and only adopt more restrictive licenses after a process of ‘due diligence and on a case-by-case basis’ ” (Principle 11(b)).

Influencing New Initiatives and Targeting Priority Legacy Datasets

Even with the resources currently available the opportunity exists to grow AODN and demonstrate the utility of a national, collaborative approach to marine data management. New initiatives are being established, such as the National Environmental Research Program (NERP), and it is imperative that AODN becomes the mechanism for

discovering and accessing the data they are collecting and using. The need for integration of marine data to better inform decision making by government, industry and the community is becoming increasingly clear at both Commonwealth and State levels, creating a strong ‘user pull’ for what AODN has to offer.

As an example, through discussion and negotiation with various partners, the establishment of a Western Australian regional view of the AODN (i.e. a WA-AODN) is being established as a pilot for a ‘State portal’ approach. This takes advantage of the existence of WAIMOS, the regional Node of IMOS, collaborative activity and strong user pull, and given the vast amount of ocean-based development planned for the WA region, makes a WA-AODN appear a real possibility and an ‘exemplar’ for other states. The plan is to publish distributed data in WA by creating ISO 19115/19139 standard metadata records (with embedded links to datasets) which can be harvested remotely by the AODN Metadata Entry and Search Tool (MEST – the Geonetwork catalogue tool). These records would reside in agencies MESTs (possibly installed by AODN Development Office for the purpose) or other catalogues harvestable because of the adoption of the same ISO standards: an example schema is shown in Fig. 5.

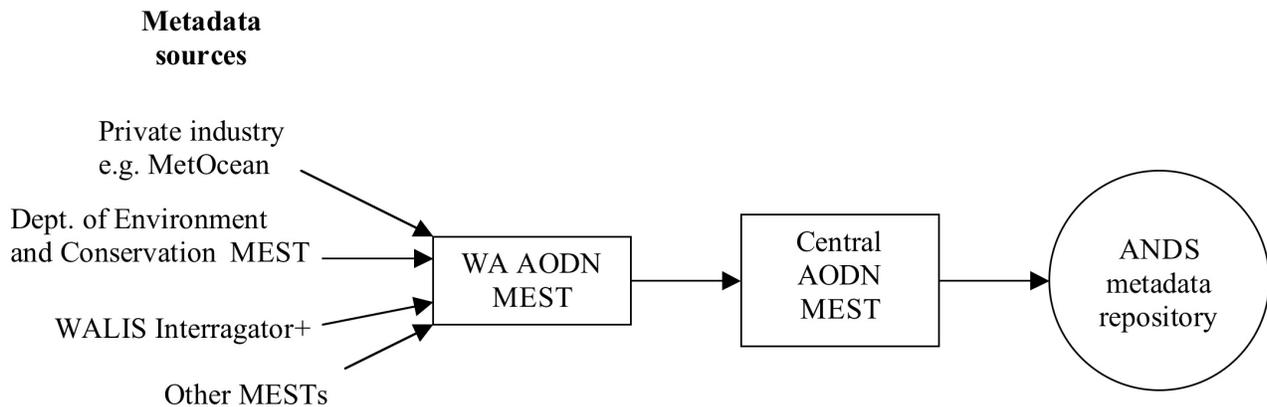


Figure 5 - Metadata flow for WA-AODN

To expose significant holdings of marine data into the ARDC, via the AODN, will require a good understanding of the effort required, including a) an assessment of the metadata required and its transformation into ISO 19115 form, b) the need for software systems installations to host the metadata records, and c) the harvesting requirements to populate the AODN catalogue and the ARDC, as well as assessing the state of the data itself. Technically this will lead to the publication of metadata via the central AODN MEST Catalogue using an agreed community metadata profile of ISO 19115 (either the current Marine Community Profile (modified) or the ANZLIC Profile), and is most likely to be done by installing a MEST instance, or some other catalogue search tool which can respond to Catalog

Services for the Web (CSW) queries for metadata. Common opensource OGC middleware tools (e.g. MapServer, GeoServer, THREDDS) that can publish data as OGC Web or OPeNDAP services may need to be installed as well as other associated local in-house systems to manage the publication of the marine data holdings. It will be important to agree on the use of common vocabularies and encoding standards for the marine datasets that are published and to expose and describe data at a consistent and appropriate level of granularity.

IV. SUSTAINABILITY OF THE AODN

The AODN is supported by the Australian Ocean Data Centre Joint Facility Board, and by the Oceans Policy Science Advisory Group (OPSAG), which is the peak body for Australian Government marine science agencies. Together, the AODC-JF Board and OPSAG are currently addressing the issue of the long-term sustainability of the AODN as a high priority.

Each of the agencies of the AODC-JF also undertakes to continue hosting those services developed through the AODN Development Office which expose their data,

critically important for the long-term sustainability of the AODN.

The information infrastructure, developed by IMOS, being used by the AODN to publish data (currently the MEST, Ocean Portal and Data Turbine) is based on open source software and therefore the code is freely accessible for bug fixes and enhancements. There are also active communities developing these systems to keep them current and IMOS/eMII are constantly monitoring other projects (e.g. the OGC project Sensor Observation Service) for possible improvements to the infrastructure.

The AODN portal also hosts substantial documentation on how to publish data which will remain as an enduring resource for those wishing to contribute to the AODN in the future.

Part of the successful sustainability of the AODN is seen as the regional adoption of the principles of data sharing.

IMOS has regional nodes and the AODN will need to create 'regional views' of the national system to fully engage the States.

The long term goal of the AODN Development Office will be to realise the AODN in a sustainable form which could be continued indefinitely given a suitable home and ongoing funding.

REFERENCES

- [1] Meyers, G. (2008). The Australian Integrated Marine Observing System. *Journal of Ocean Technology*, 3, 80-81.
- [2] Hill, K.L., Rintoul, S.R., Coleman, R. and Ridgway, K. R. (2008) Wind forced low frequency variability of the East Australia Current, RESEARCH LETTERS, VOL. 35, L08602, doi:10.1029/2007GL032912.
- [3] Proctor, R., Roberts, K., Ward, B. J. (2010) A data delivery system for IMOS, the Australian Integrated Marine Observing System. *Advances in Geosciences*, 8, 1-6. www.adv-geosci.net/8/1/2010/, doi:10.5194/adgeo-8-1-2010

ACKNOWLEDGEMENTS

The Integrated Marine Observing System is supported by the Department of Innovation, Industry, Science and Research through the National Collaborative Research Infrastructure Strategy and the Super Science Initiative. Thanks to Luke Edwards (Curtin University) for permission to show Figure 5. Thanks also to the anonymous reviewers for their comments.

LIST OF ACRONYMS

ANDS – Australian National Data Service
ANZLIC - Australia and New Zealand Land Information Council
AODC-JF – Australian Ocean Data Centre Joint Facility
AODN – Australian Ocean Data Network
ARCS – Australian Research Collaboration Service
ARDC – Australian Research Data Commons
AUV – Autonomous Underwater Vehicle
CF – Climate and Forecast
CSW – Catalog Services for the Web
DC – Direct Connect
EAC – East Australian Current
eMII – eMarine Information Infrastructure
IMOS - Integrated Marine Observing System
ISO – International Standards Organisation
MCP – Marine Community Profile
MEST – Metadata Entry and Search Tool
NERP - National Environmental Research Program
NPEI – National Plan for Environmental Information

OAI - Open Archives Initiative
OGC – Open Geospatial Consortium
OPeNDAP - Open-source Project for a Network Data Access Protocol
OPSAG - Oceans Policy Science Advisory Group
THREDDS - Thematic Real-time Environmental Distributed Data Services
WA – Western Australia
WFS – Web Feature Service
WMS – Web Map Service