



# **Integrated Marine Observing System**

## **Facility 10**

### **Australian Ocean Data Network**

#### **Implementation Plan 2017-2019**

**Version 3: 2017/08/03**



**Australian Ocean Data Network**



IMOS is a national collaborative research infrastructure, supported by Australian Government. It is operated by a consortium of institutions as an unincorporated joint venture, with the University of Tasmania as Lead Agent.



## Executive Summary

This document contains the Australian Ocean Data Network (AODN) Implementation Plan for July 2017 to June 2019, utilising the funding of NCRIS to support IMOS. External funds from other sources may arise for associated activities.

On 1 September 2016 the Australian Ocean Data Network was adopted by the National Marine Science Committee (NMSC) as the main conduit for publishing Australian marine data. AODN now has a responsibility to report regularly to the NMSC on AODN activity. NMSC also has overview of the AODN Implementation Plan; this Plan has the Committee approval.

The AODN portal (<https://portal.aodn.org.au>) currently includes data collections served by 11 organisations (with many more contributing) and 170 data collections can be discovered, accessed, sub-setted and downloaded. The metadata for these collections are described in the metadata catalog for AODN (<https://catalogue.aodn.org.au>). The catalog also holds descriptions of many more data collections (14,261 in total) most of which do not yet meet the specifications for inclusion in the AODN portal.

In 2016-17 our primary activity was on delivering the IMOS data collections. This resulted in several improvements to the information infrastructure now hosted on Amazon Web Services. In particular the sub-setting and aggregation of gridded data products (satellite Sea Surface Temperature and Ocean Colour; HF radar surface currents) was significantly improved with a near order-of-magnitude improvement in aggregation time; the export of single point timeseries was also implemented. The ingest of IMOS facility data has been improved by the establishment of 'pipeline' processes which test for conformance of a range of standards and speed the delivery of conforming datasets to publication. Enhancements to the Moorings Toolbox (<https://github.com/aodn/imos-toolbox/wiki>), now used by all IMOS moorings sub-facilities, improve processing and quality control.

More AODN-related, the development of controlled vocabularies and their management and publication using the Research Vocabularies Australia service (<http://www.ands.org.au/online-services/research-vocabularies-australia>) now sees AODN vocabularies available for community use in metadata creation; their use being encouraged through the new on-line AODN Metadata Submission Tool (<https://metadataentry.aodn.org.au/submit/>). In addition, all data collections and services (IMOS and AODN contributors) are now monitored for 'good health' and comprehensive analytics are collected across the whole infrastructure and published monthly.

This Implementation Plan embraces the new status accorded by the NMSC and seeks to explore the opportunities this brings, through the AODN Technical Advisory Group, to add new data collections to the AODN (including many of those in the AODN catalog which currently do not meet portal requirements). The changing landscape of e-infrastructure support provided by ANDS, NeCTAR and RDS, to align their investments with NCRIS capabilities and domain research data, although not quantifiable at this point in time, is likely to offer AODN with significant opportunities to develop information infrastructure in support of the marine community. Some ideas for these developments are itemised in the Plan, should funding become available.

All activities in this Plan are seen as enabling the transition to the IMOS 5-year plan (<http://imos.org.au/about/plansreports/5yearplan17-22/>). In addition, of course, IMOS data collections will continue to grow and AODN staff will endeavour to improve data delivery, enhance the information infrastructure, develop new reporting metrics, and explore product developments.

## **AODN Implementation Plan 2017-2019**

The objectives of this plan are to:

- Consolidate the AODN infrastructure and improve discovery and access to data
- Establish a number of specific AODN community developments
- Complete existing contractual obligations to NeCTAR/ANDS and GBRMPA.
- Consider ideas for aligning certain AODN activities, or establishing new ones, with the ANDS/NeCTAR/RDS Alignment Plan

## **AODN Infrastructure**

The major activities will be:

- 1) Operational activities, utilising AWS services
- 2) Database improvements including AODN vocabulary integration, and Animal Tracking acoustic database efficiencies
- 3) Continuing to improve IMOS facility data delivery workflows through pipeline improvements
- 4) Improving business reporting on access to and uptake of data collections
- 5) Providing improved assistance and information to enable integration of additional partner data collections into the AODN portal

Explicit IMOS-related activities:

These are considered the critical elements of the workplan, keeping IMOS data flowing is paramount. Focus on four areas: data ingest; portal enhancements and features; information reporting; other activities.

- **Data ingest:** the focus of this work will be consolidating the end-to-end pipeline data ingestion including conformance checking and logging of all activity to provide comprehensive information for all IMOS data. Additional work includes improvements to the ATF data ingestion to enhance reliability, and implementation of QA/QC protocol in the Moorings Toolbox
- **Portal enhancements and features:** some of the activity here is related to improvement of back-end services, such as upgrading the portal catalog to Geonetwork-3 and implementing core Geoserver with MCP plugin; some activity is directed to improving user experience of obtaining data through the portal – examples include implementing a batch queueing service for gridded data requests, better configuration of metadata in downloaded csv files, additional facets and vocabularies to improve search options
- **Information reporting:** we can identify several areas where improvements to reporting are possible. Building on the implementation of Google Analytics and Sumo Logic we expect to consolidate reporting of data uptake; there is scope for improved automation of the monthly data holdings reports; in association with the IMOS Office there are expected (planned) enhancements to the publications database and the development of a scheduling/inventory reporting tool
- **Other activities:** It is planned to develop easily spun-up test environments by making use of the special features provided by AWS, these will enable our data scientists to speed up configuration of and access to new datasets. OceanCurrent is the popular viewpoint for IMOS real-time data and there has been a call for closer relationships between the AODN portal and OceanCurrent, so a project will explore the options for doing so. Animal tracking is becoming an increasingly popular activity internationally; IMOS is leading the development of an international data exchange metadata template and is exploring potential collaborations

on database development with the major players (US IOOS-ATN, OTN, ETN) which may lead to development of a new construct.

Developments for wider AODN.

- Upgrading the portal catalog to Geonetwork-3 and implementing core Geoserver with MCP plugin will enable easier adoption of the AODN software stack
- The IMOS gridded data aggregator has the capability to work with external THREDDS servers, so we can consider offering this service to additional partner THREDDS catalogs. We will be looking to identify partner catalogs (e.g. TPAC, IMAS) which can be included
- We have developed a new data format for ingest/output, i.e. shapefile for NSW OEH bathymetry. This is of use to the wider AODN community and can enable pipelining ingest from contributors. Wider use of this will be explored
- The AODN Cookbook provides the community with detailed information on “Contributing Data”; last updated in 2015 it is time for a review
- Communications (e.g. newsletters, outreach, education involvement) will continue. We publish a regular newsletter and operate facebook/twitter feeds, and have an increasing role in contributing to higher education courses. This year we will be looking to run targeted workshops with potential AODN contributing organisations.

**Operational activities – ensuring the infrastructure is reliable, ensuring all available data is discoverable, accessible and downloadable – is the largest element in the team activities and consumes around 65% of the 16.3 EFT AODN staff resources.**

#### **AODN Community developments**

At the AODN Technical Advisory Group meeting in May 2017 certain developments were identified as important for community development. These are identified in the table below:

<b>Potential Activity Area</b>	<b>Potential Partners</b>	<b>Scoping paper or WG</b>
Improving the flow of biological data to the AODN portal. This would focus in the first instance on OBIS Australia and involve sorting out the workflow issues with respect to marine biological data between ALA, OBIS Node and AODN; sort out workflows and policies between agencies regarding the publication of the associated hydrographic data and its ingestion into a national hydrographic data collection (e.g. enhancing the MARVL shelf seas in-situ atlas); work with several committed agencies to get biological/hydrographic data into the Portal; work with those agencies and their users to identify required Portal enhancements that would make it easier to access these types of data in the Portal. This would call on WG activity for vocabularies (biological vocab & vocab mapping) and information modelling (standardisation of OGC web service responses and standardised query templates), and require infrastructure linkages between the portal and CSIRO Data	IMOS, CSIRO, IMAS, GA, SIMS, NIWA, ALA, ANDS	Scoping paper and WG topics

Trawler / ALA.		
Visibility of Australian research vessel underway data in the portal. CSIRO, AAD and AIMS hold data collections which could be published in the portal, as do universities and state organisations. Enabling a project to do this, which includes publishing of Cruise Summary Reports, would contribute to a recommendation of the National Marine Science Plan to have better coordination between research vessel activities. This would involve WGs on Vocab, Use cases, information modelling.	IMOS, CSIRO, AIMS, AAD, IMAS, SARDI, UWA, NSW OEH, WA DoT, etc.	Scoping paper and WG topics
Set up a WG to identify and scope the publication in the portal of High Profile Data Collections. The kind of questions the WG will address include: <ul style="list-style-type: none"> <li>• What is the current situation with an identified dataset (Where is it stored? download format ...)?</li> <li>• What could be the role of AODN to publish these datasets?</li> <li>• How can we add value to what is already provided?</li> </ul> Examples considered at the meeting included a National Wave Archive and the Southern Ocean Repeat Section, or SR3 transect. Industry data collections would also be considered and consolidate links to the Forum for Operational Oceanography.	IMOS, CSIRO, IMAS, ACE CRC, AIMS, ANDS, Pawsey, GA, BoM	Establish WG
Draft scoping papers on what it will take to integrate the data holdings of the major custodians of marine and ocean data. In these initial scoping papers we would aim to pick one dataset (maybe a high profile one) and list all the actions that need to happen in order to get this particular dataset available in the AODN portal. For each organisation and dataset, we could present one or two solutions (simple or more difficult) and we would seek feedback from the NMSC on their preference.	IMOS, CSIRO, GA, BOM, AAD, AIMS	Scoping papers

The Working Groups proposed will be established and scoping papers developed to identify technical and staff requirements to develop these activities. Some areas of work have already begun, for example:

- a) Sample IMOS acoustic tag data has been formatted for suitability of OBIS ingest
- b) AODN, through the ANDS-funded Seemap Australia project, has assisted IMAS to publish the [Seemap Australia National Benthic Habitat Classification Scheme](#) in the RVA
- c) work has started on establishing the National Wave Archive.

In addition, AODN TAG meetings will be held twice a year to review activity and propose avenues of work.

AODN continues to be involved in international best practice in data management and expects to have ongoing involvement in the following:

- International Oceanographic Data and Information Exchange ([www.iode.org](http://www.iode.org); AODN Director is the Australian national representative for oceanographic data management)
- International Quality Controlled Ocean Database (IODE-IQuOD, [www.iquod.org](http://www.iquod.org))
- International Ocean Tracking Data Standardization: International Metadata Standards and data transfer protocols
- Ocean Data Interoperability Platform ([www.odip.org](http://www.odip.org))
- Ocean Gliders Data Management Task team ([www.ego-network.org](http://www.ego-network.org))

### **Contractual obligations**

The Marine Virtual Laboratory (MARVL) contract (“increasing flexibility and utility”, a consortium of AODN, CSIRO and TPAC) runs until 30 September 2017, so contractual requirements to NeCTAR and ANDS have to be met in the first quarter. This contract enables MARVL to connect to OPeNDAP servers and to create provenance metadata records for export.

The NeCTAR marine sciences cloud contract (AODN, TPAC and sub-contract to GreyBits) runs until 30 September 2017, so contractual requirements have to be met in the first quarter. This contract enables the establishment of a national service for annotating and analysing underwater imagery and rapid deployment of virtual desktops for analytics.

Reef 2050 Integrated Monitoring, Modelling and Reporting (RIMReP) contract (to meet the requirements of DMS3 and DMS5) runs until April 2018, so contractual requirements have to be met in the first three quarters. This contract conducts an audit of existing observing systems and proposes best practice and protocols for the management of data services to meet RIMReP reporting requirements.

### **Future opportunities**

In response to the National Research Infrastructure Roadmap ANDS, NeCTAR and RDS have drawn up an alignment plan for 2017-2018 to work collaboratively towards the development of an Australian Research Data Cloud, described as “an integrated data-intensive infrastructure system, incorporating physical infrastructure, policies, data, software, tools and support for researchers”.

The alignment plan covers the period of 2017/18 only and articulates three proposed joint infrastructure investment programs:

- ❖ Research Domain Program
  - Responding to research domain and research community data infrastructure needs
- ❖ Research Data Platforms
  - Underpinning compute, storage and services infrastructure to support the data and informatics needs of Australian research and industry
- ❖ Sector-wide Support and Engagement
  - Planning, coordination and leadership to further policy development, international engagement, progress a national skills strategy, and support the building and sharing of data assets across the sector.

IMOS/AODN has driven work through ANDS, NeCTAR and RDS funding in the past and under this Alignment plan NCRIS capabilities are seen as core participants of the domain activities.

Of particular interest to AODN are the activities under the Research Domain Program, i.e.

- ❖ Data-enhanced Virtual Labs
- ❖ Research Data Cloud Program

AODN will be involved in defining the content of these activities and looks to develop work programs in each. Opportunities are seen to

- a) further enhance and evolve the Marine Virtual Laboratory into a national modelling infrastructure as a pre-cursor to the major activity of a national shelf seas re-analysis program – designed to provide data products in support of research, management and industry activity in Australia’s coastal and shelf seas
- b) build on the existing marine sciences cloud and consolidate with activities identified by the AODN TAG (see above); align a number of cloud-based activities with international engagement (e.g. marine animal tagging); develop pipelines for the creation of new data products.

AODN, as an e-research service, will endeavour to be involved in other e-research activities where appropriate, and will continue to advise/direct on ways to increase AODN community engagement through the addition of new data collections, including model data collections.

The timing of activities will be such as to meet the 2017-2019 milestones agreed and documented in the IMOS Office spreadsheet, as itemised below. It is envisaged that effort in the first half of 2017-18 will concentrate on infrastructure development and the second half of 2017-18 on data delivery enhancements. ANDS/NeCTAR/RDS engagement is expected towards the end of Q1, beginning of Q2.

**Milestones & Deliverables (2017-2018) (tentative)**

Activity	Start date	Finish date
<b>Q1 (1/7/2017-30/9/2017)</b>		
WPS Provenance service	1/7/2017	30/9/2017
Geonetwork3 backing the portal	1/7/2017	30/9/2017
MARVL contract completed		30/9/2017
<b>Q1 Milestone: TAG WG scoping documents – 30/9/2017</b>		
<b>Q2 (1/10/2017-31/12/2017)</b>		
Test Environments on AWS	1/7/2017	31/12/2017
QA/QC protocol implemented in Matlab Toolbox	1/7/2017	31/12/2017
Improved efficiency of the Animal Tracking database	1/7/2017	31/12/2017
<b>Q2 Milestone: Pipeline dataset logging operational – 31/12/2017</b>		
<b>Q3 (1/1/2018-31/3/2018)</b>		
Download improvements (Batch processing; metadata)	1/11/2017	31/3/2018
Additional data products (e.g. SST climatology)	1/10/2017	31/3/2018
RIMReP contract completed		31/3/2018
<b>Q3 Milestone: Updated AODN Cookbook – 31/3/2018</b>		

<b>Q4 (1/4/2018-30/6/2018)</b>		
Additional portal search facets	1/4/2018	31/6/2018
Enhanced connectivity between AODN Portal and OceanCurrent	1/1/2018	30/6/2018
<b>Q4 Milestone:</b> national surface wave observations archive – 30/6/2018		

Milestones & Deliverables for 2018-2019 will depend on developments in 2017-2018, in particular the outcomes of the TAG WGs and the ANDS/NeCTAR/RDS alignment with the NCRIS Roadmap.

#### List of acronyms

AAD	Australian Antarctic Division
ACE-CRC	Antarctic Climate and Ecosystems-Collaborative Research Centre
AIMS	Australian Institute of Marine Science
ALA	Atlas of Living Australia
ANDS	Australian National Data Service
AODN	Australian Ocean Data Network
AWS	Amazon Web Services
BoM	Bureau of Meteorology
CSIRO	Commonwealth Scientific and Industrial Research Organisation
DoT	Department of Transport
GA	Geoscience Australia
GBRMPA	Great Barrier Reef Marine Park Authority
IMAS	Institute for Marine and Antarctic Studies
IMOS	Integrated Marine Observing System
MARVL	Marine Virtual Laboratory
MCP	Marine Community Profile 2.0
NCRIS	National Collaborative Research Infrastructure Strategy
NeCTAR	National eResearch Collaboration Tools and Resources
NIWA	National Institute for Water and Atmosphere (NZ)
NMSC	National Marine Science Committee
OEH	Office of Environment and Heritage
RDS	Research Data Services
RIMReP	Reef 2050 Integrated Monitoring, Modelling and Reporting Program
SARDI	South Australian Research and Development Institute
SIMS	Sydney Institute for Marine Science
TPAC	Tasmanian Partnership for Advanced Computing
WPS	Web Processing Service
US IOOS-ATN, OTN, ETN	United States Integrated Ocean Observing System-Animal Tracking Network, Ocean Tracking Network, European Tracking Network
UWA	University of Western Australia