



Australian Ocean Data Network

Notes Before Completing the Application

We have read and understood the notes concerning our application submission.

True

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

Yes

CORE TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS

Background & General Guidance

Glossary of Terms

BACKGROUND INFORMATION

Context

R0. Please provide context for your repository.

Repository Type. Select all relevant types from:

Domain or subject-based repository, National repository system; including governmental, Research project repository

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
concur

Brief Description of Repository

Australia's Integrated Marine Observing System (IMOS) has been operating a wide range of observing equipment throughout Australia's coastal and open oceans since 2006. IMOS makes all its data openly and freely accessible to the marine and climate science community, other stakeholders and users, and international collaborators. IMOS is one of the national research infrastructure capabilities currently supported under the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS; <https://education.gov.au/national-collaborative-research-infrastructure-strategy-ncris>). It is operated as a consortium of institutions as an unincorporated joint venture (IMOS Principal Participants Agreement (IMOS UJV; http://imos.org.au/fileadmin/user_upload/shared/IMOS_General/Framework_Policy/IMOS_Participants_Agreement_UJV_signed_23Nov2016.pdf)), with the University of Tasmania (UTAS) as Lead Agent.

The operation of IMOS is enabled by a series of facilities (<https://imos.org.au/facilities>) including the Australian Ocean Data Network (AODN), the data management facility for IMOS. More broadly, the AODN strives to make accessible, as much marine data as possible, including publicly funded data from other organisations, as well as data from private industry and not-for-profit organisations.

The AODN is a "National repository system, including governmental" because it serves as the data management facility for IMOS.

The AODN is a "Domain or subject-based repository" because the AODN only deals with marine and climate data resources.

The AODN is a "Research project repository" because the Marine Biodiversity Hub of the National Environmental Science Programme (NESP) (recently superseded by the Marine and Coastal Hub; <https://www.awe.gov.au/science-research/nesp/hub-marine-coastal>), in its Data Management Framework (<https://www.nespmarine.edu.au/document/data-management-framework>) identifies that the AODN is the aggregation point for all spatial and non-spatial Hub data.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
No comment

Brief Description of the Repository's Designated Community.

The AODN's primary role as the data management facility for IMOS, designates its target community and intended users from the marine and climate science community, other stakeholders and users, and international collaborators. As our data is freely available to anyone, there is a small pool of users that would fall outside our target community, such as sailors, who can require extra support in accessing and using our data.

The AODN is an interoperable online network of marine and climate data resources, exposing not only IMOS funded data, but other marine and climate data throughout Australia's coastal and open oceans. IMOS and the following six Australian Commonwealth agencies: Australian Institute of Marine Science (AIMS), Geoscience Australia (GA), Australian Bureau of Meteorology (BOM), Royal Australian Navy (RAN), Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the Australian Antarctic Division (AAD) initially formed the core of the AODN, which has since grown to encompass organisations and individual members of the Australian, New Zealand and Pacific marine research community (<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-partners>). Increasingly, though, universities and state government offices are offering up data resources to the AODN, and delivery of data to the AODN is being written into significant research programs e.g., the NESP Marine Biodiversity Hub (<https://www.nespmarine.edu.au/>).

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:
No comment

Level of Curation Performed. Select all relevant types from:

A. Content distributed as deposited, B. Basic curation – e.g. brief checking; addition of basic metadata or documentation, C. Enhanced curation – e.g. conversion to new formats; enhancement of documentation, D. Data-level curation – as in C above; but with additional editing of deposited data for accuracy

Reviewer Entry

Reviewer 1

Comments:
Accept. All levels supported.

Reviewer 2

Comments:

Concur

Comments

The majority of data hosted by the AODN, is IMOS funded (henceforth referred to as IMOS data), and designated Project Officers work with individuals/organisations to ensure the appropriate and required level of curation is enacted upon data deposition. The AODN provides a metadata authoring and data submission tool to aid in the submission of non-IMOS funded data for hosting (henceforth referred to as AODN data) - <https://metadataentry.aodn.org.au/submit/>. For all such submissions of metadata, the Metadata Officer will assess the metadata record and liaise with the metadata author to alter or add extra information where required.

By order of importance, these are the curation options that can be adopted at the AODN:

- Option C. “Enhanced curation” is followed the majority of the time, via two different pathways:
 - o Most of the time, Project Officers and data providers agree on data standards (see R8), as well as file formats to be provided to the AODN. Project Officers will offer support to the data providers so they can deliver their data to the AODN in the agreed standards and file formats.
 - o On a rare occasion, the AODN needs to convert original data files to another file format (see R7) that is more suitable for distribution to the broader audience (or to enable inclusion into a database), and that the data provider is not able to produce or maintain.
- Option D. “Data-level curation” is followed in some exceptions, for example, that the AODN as a central repository of a particular type of data, produces higher level products from a baseline dataset that encompasses multiple data providers. The benefit is shared across data providers irrespective of their capabilities and capacities. These higher-level products are designed and implemented in partnership with the relevant data providers.
- Options B. “Basic curation”, and A. “Content distributed as deposited” are exceptionally applied to AODN data that are not IMOS funded and deemed of lower importance.

Note that preservation goals and actions are the same for all data, regardless of the curation level.

Reviewer Entry

Reviewer 1

Comments:

Accept. Rationale is well described.

Reviewer 2

Comments:

Answer is complete and appropriately detailed.

Insource/Outsource Partners. If applicable, please list them.

Insource Partners

- Institute for Marine and Antarctic Studies (IMAS) – metadata is harvested to the AODN metadata catalogue

(<https://catalogue.aodn.org.au/> see R13 for further information) and a subset for discovery via the AODN Portal (<https://portal.aodn.org.au/> see R13 for further information). Future application for CoreTrustSeal (CTS) certification.

- CSIRO Oceans and Atmosphere (specifically – the Information and Data Centre (IDC) at CSIRO National Collections and Marine Infrastructure (NCMI)) - metadata is harvested to the AODN catalogue and a subset for discovery via the AODN Portal.
- Australian Institute of Marine Science - metadata is harvested to the AODN catalogue and a subset for discovery via the AODN Portal. Currently drafting application for CTS certification.
- National Institute of Marine and Atmospheric Research (NIWA), New Zealand - metadata is harvested to the AODN catalogue and a subset for discovery via the AODN Portal.
- Geoscience Australia - metadata is harvested to the AODN catalogue and a subset for discovery via the AODN Portal. Currently drafting application for CTS certification.
- Australian Antarctic Division - metadata is harvested to the AODN catalogue and a subset for discovery via the AODN Portal. CTS certified.
- Deakin University - metadata is harvested to the AODN catalogue and a subset for discovery via the AODN Portal.
- National Computational Infrastructure (NCI) - metadata is harvested to the AODN catalogue.
- eAtlas - metadata is harvested to the AODN catalogue.

Outsource Partners

- Primarily, Amazon Web Services (AWS; <https://aws.amazon.com/>) infrastructure (hardware/support/storage). The AWS shared responsibility model (<https://aws.amazon.com/compliance/shared-responsibility-model/>) provides details on agreed responsibilities.
- The National eResearch Collaboration Tools and Resources project (Nectar; <https://nectar.org.au/>), provides servers for development environments (hardware).
- Condense (<https://condense.com.au/>) software consultancy (software support)

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
Answer is appropriately detailed.

Summary of Significant Changes Since Last Application (if applicable).

N/A

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

Other Relevant Information.

Reviewer Entry

Reviewer 1

Comments:

Reviewer 2

Comments:

ORGANIZATIONAL INFRASTRUCTURE

1. Mission/Scope

R1. The repository has an explicit mission to provide access to and preserve data in its domain.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

In the IMOS Principal Participants Agreement (http://imos.org.au/fileadmin/user_upload/shared/IMOS_General/Framework_Policy/IMOS_Participants_Agreement_UJV_signed_23Nov2016.pdf), the following is identified: “The core mission of IMOS is to routinely operate a wide range of observing equipment throughout Australia’s coastal and open oceans, making all of its data accessible to the marine and climate science community, other stakeholders and users, and international collaborators”.

The aim of the AODN, more broadly, is to make as much marine data as possible – publicly funded as well as data from private industry and not-for-profit organisations - accessible and freely available over the internet (as outlined in the AODN Data Policy). The AODN's Mission Statement (<https://imos.org.au/facilities/aodn>) is outlined as follows: "Providing efficient and effective ingestion and trusted publication of data, value added data aggregation and products for IMOS and national users."

The IMOS Data Policy (Version 2.1, <https://imos.org.au/imos-data-policy>) was developed in conjunction with the AODN Data Policy (Version 3.1, <https://imos.org.au/facilities/aodn/aodn-data-management/aodn-data-policy>). As IMOS participants are required to lodge data to the AODN, large parts of the AODN Data Policy are directly related and many aspects of the two policies are naturally similar. The contractual constraints on IMOS facilities are described within the IMOS Data Policy. Both data policies have been reviewed and endorsed by IMOS Management.

The IMOS Data Policy clearly states what the responsibilities of the AODN are for IMOS data.

This section of the data policy is copied below:

"IMOS AODN has accepted responsibility for preserving IMOS digital assets including data and derived products hosted by the IMOS AODN. Staff at the AODN will ensure that:

- IMOS scientific data is managed for the long-term and made available in an easily accessible form,
- Metadata records are managed for all IMOS scientific research data and are made available for public searching in an effective form,
- Where possible, and appropriate, Australian sourced data are included in international data systems,
- Current best practice is maintained in relation to its repository management functions and related systems,
- Confidentiality of data is maintained during embargo periods,
- The usage of custodian's data and the custodian's right to be cited as the source of published data is monitored,
- Information is provided to scientists on the resources that are available to support their work and advice is available on the design of data collection programs and effective data management strategies through assistance with developing Project Plans."

The AODN Data Policy clearly states what the responsibilities of the AODN are for AODN data. The AODN hosts data by exception. Exceptions will be made to assist organisations short-term while they develop their own systems or are working toward developing their own hosting arrangement. The AODN will also assist individuals who are unable to store their own data and metadata but are interested in making their data publicly available.

This section of the data policy is copied below:

"The AODN acting for and on behalf of contributors of data hosted by the AODN has accepted responsibility, for the duration of IMOS funding, for preserving the digital assets including data and derived products. Staff at the AODN will ensure that:

- Data is managed for the long-term and made available in an easily accessible form,
- Metadata records from all participants are available for public searching,
- It maintains best practice in relation to the systems that it uses, and
- It provides appropriate guidelines and tools and assistance for participants.”

Also outlined in the AODN Policy are the requirements around data that is made accessible via the AODN, but where the data is not hosted by the AODN.

This section of the data policy is copied below:

“A condition of participation in the AODN is that all data provided to the program is unencumbered – in that it is freely accessible at no charge to third parties. It is a requirement of the AODN that all data provided to the AODN is adequately documented with metadata and arrangements are made for data to be held by the custodian organisation or an alternate organisation for long term access. The AODN aims to be a federated system where data and metadata is stored with the originating organisation and harvested by the AODN metadata catalogue.”

Further to information provided in the two data policies, the AODN Preservation Strategy (<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-preservation-strategy>) sets out the approach to the long-term retention and preservation of the data hosted by the AODN and gives guidance on implementing those principles. The purpose of the strategy is to document initiatives to improve access to and re-use of data hosted by the AODN, and to facilitate services to enable the efficient and effective management of the assets.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
Appropriate detail and useful external references.

2. Licenses

R2. The repository maintains all applicable licenses covering data access and use and monitors compliance.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

Internal (IMOS data)

All IMOS generated data (exception for some data from the Animal Tracking Facility, see below), is public and freely available with a Creative Commons “Attribution” licence (CC-BY: <https://creativecommons.org/licenses/by/4.0/>), see the Creative Commons website for more detail at <https://creativecommons.org.au/learn/licences/>). Whenever a work is copied or redistributed under this type of Creative Commons licence, the original creator (and any other nominated parties) must be credited as the source of the data. The licence has no other restrictions on use, although does include the accompanying statement: “Data, products and services from IMOS are provided "as is" without any warranty as to fitness for a particular purpose”.

Each metadata record describing a dataset, provides alongside the CC-BY licence, a default template for citation, requiring acknowledgement of IMOS, and if relevant, to credit other partners. Both the licence and the suggested citation are also highlighted when users download data from the AODN Portal.

The IMOS Animal Tracking Facility (<https://imos.org.au/facilities/animaltracking>) provides a central data repository for the Australian acoustic telemetry scientific community to manage and access data related to IMOS receiver deployments as well as receivers and transmitters deployed by contributing researcher groups (Hoenner et al. 2018). The IMOS Animal Tracking Facility Data Policy (http://content.aodn.org.au/Documents/IMOS/Facilities/animal_tracking/LATEST/IMOS_Animal_Tracking_Facility_Data_Policy_LATEST.pdf) details how data contributed may be stored under one of three access levels. The majority of data is Open Access, albeit with reduced accuracy of receiver and transmitter deployment locations. A small amount of sensitive data (1% of all detections) is embargoed or protected. Data embargo can be applied at the transmitter level for a set period of five years from the deployment data and is not renewable. After expiration of the period, the associated data and metadata will be made available as Open Access. Protected Data status provides security where publicly available transmitter metadata and associated detection data may present threats to tagged animals or projects. Projected Data status can be applied to approved specific projects for a set period of 99 years.

External (AODN data)

All data made available to the AODN will be made publicly available – to facilitate this, the AODN recommends that data is

licensed through an appropriate Creative Commons license, preferably the By Attribution (CC-BY) license. Each metadata record describing a dataset, hosted by the AODN, is created using the AODN Data Submission Tool (<https://metadataentry.aodn.org.au/submit/>) which requires selection of an appropriate CC licence. Similarly, to IMOS data, alongside the licence, a default template for citation is provided in the metadata record.

Whilst the AODN recommends that users cite their data sources according to the guidelines of the CC-BY licence, it is difficult to make this mandatory. Currently data users are not required to register for access to the data, which makes compliance monitoring infeasible. In the current AODN Implementation Plan 2021-2022, a key milestone is the documentation of options for implementing user registration in the AODN Portal. User registration will be encouraged in the first place, and then eventually made mandatory. This information will be used to determine usage, rather than to monitor future compliance of said usage.

IMOS maintains a record of the uses (publications, reports, data products, projects, and postgraduate research projects) and users (for example: authors, students and principal investigators) of IMOS data within the IMOS Impact Database. This information is collected through a variety of mechanisms including quarterly milestone reporting from IMOS facilities and targeted keyword searches of publicly available material in online repositories. IMOS does not currently have a formal mechanism for following up on non-compliance in appropriate acknowledgements for the use of IMOS data. Acknowledgement practices however are recorded for each use of IMOS data and the IMOS office is currently in the process of establishing a plan regarding increased acknowledgement. We expect by the time of CTS renewal (2025) that IMOS would have a process to track citation compliance.

Hoenner, X., Huveneers, C., Steckenreuter, A. et al. Australia's continental-scale acoustic tracking database and its automated quality control process. *Sci Data* 5, 170206 (2018). <https://doi.org/10.1038/sdata.2017.206>

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
The licensing is well described reasonable.

3. Continuity of access

R3. The repository has a continuity plan to ensure ongoing access to and preservation of its holdings.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:

3 – The repository is in the implementation phase

Reviewer 2

Comments:

3 – The repository is in the implementation phase

Response:

The AODN has a Preservation Strategy

(<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-preservation-strategy>) that sets out the approach to the long-term retention and preservation of the data hosted by the AODN (for more details see R10). This strategy applies to the dataset collections and derivatives hosted by the AODN, all data contributors including IMOS staff and researchers, and non-IMOS and affiliate contributors and contributions for the duration of IMOS funding.

The AODN is the data management facility of IMOS, and its funding is considered at the program level within IMOS, i.e. that IMOS needs a properly resourced AODN Facility to enable open access to data

(<https://imos.org.au/about/imos-five-year-plan-2022-2027>). The AODN budget for the last several years has averaged 11% of the IMOS NCRIS budget, that is in addition to primary data management performed before the data reach the AODN (Lara-Lopez, A., et. al. (2016), which is more than the common rule of thumb of 5 to 10% of funding of a project should be devoted to data management (Tanhua et. al. (2019). IMOS funding through NCRIS currently extends to June 2023. The NCRIS program is funded until June 2028 in Commonwealth budgets and it is anticipated that as a foundational NCRIS capability IMOS will continue to be funded for the life of the current program (i.e. June 2028). IMOS is highly regarded within the NCRIS program and international arenas as a leader in ocean observing and data delivery. We are often used as a test case or model infrastructure for other programs. As indicated above, our strong track record and status suggest Commonwealth funding support will continue.

IMOS does not have a formal agreement with another institution if funding is withdrawn. In the next two years we plan to investigate options for an exit strategy from our current cloud provider (AWS) and formalise a continuity plan in relation to cessation of funding, with an appropriate federally funded government organisation.

IMOS is operated by a consortium of institutions (as an unincorporated joint venture (UJV)), with the University of Tasmania as Lead Agent. In the unlikely event that something was to happen, or funding was to be discontinued, our metadata and data holdings would be secure through the following means:

1) The IMOS and AODN metadata catalogues are currently harvested to Research Data Australia (RDA) (managed by the Australian Research Data Commons (ARDC), <https://researchdata.edu.au/>), so copies of the metadata would be available here. IMAS at the University of Tasmania operates similar metadata infrastructure, and transfer of the content of each of

the metadata catalogues, to be hosted by IMAS, would be arranged.

2) The initial seven parties of the IMOS UJV are UTAS, AIMS, Australian Bureau of Meteorology, CSIRO, South Australian Research and Development Institute (SARDI), Sydney Institute of Marine Science (SIMS) and University of Western Australia (UWA); with associate participants being Curtin University and the Australian Antarctic Division. The breadth of organisations involved in the running of IMOS, provides scope for hosting of IMOS funded data by one or several of these organisations, of which several maintain their own metadata and data infrastructure.

3) For the small amount AODN data that is hosted by the AODN, with a portion of this covering datasets covered by Digital Object Identifiers (DOI), the logical arrangement would be alongside the related metadata that will be hosted by IMAS.

Note as the AODN mints DOIs for both IMOS and AODN data, and IMOS documentation, these links would need to be maintained.

Lara-Lopez, A., et. al. (2016). Australia's Integrated marine observing system (imos): data impacts and lessons learned. *Mar. Technol. Soc. J.* 50, 22–33. doi: 10.4031/MTSJ.50.3.1

Tanhua et. al. (2019). Ocean FAIR Data Services. *Front. Mar. Sci.* 6:440. doi: 10.3389/fmars.2019.00440

Reviewer Entry

Reviewer 1

Comments:

Accept at level 3. Demonstrated thought and considerations for a continuity plan are provided, with an anticipated completion prior to the next certification renewal. The near term is adequately supported.

Reviewer 2

Comments:

The organization has considered short- to medium-term risks. Addressing the long-term funding risks is a difficult problem. As a minor note, the ownership of the DOIs (in terms of enabling a redirect to a different organization) would need to be addressed as part of the transition plan. The organization is in the process of further improving their good start on a continuity plan.

4. Confidentiality/Ethics

R4. The repository ensures, to the extent possible, that data are created, curated, accessed, and used in compliance with disciplinary and ethical norms.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The AODN (for its own IMOS funded data this is clearly understood, see below) does not require evidence that research has been carried out legally and ethically. All applicable data has been produced from within universities and recognised Australian organisations, so upon deposition (or harvesting to) with the AODN, we can be confident that legal and ethical guidelines have been followed.

The majority of IMOS data has no ethics risk, as there is no human or animal involvement. The only IMOS data with ethical considerations, fall under the Animal Tracking Facility. Animal Tracking consists of 2 Sub-Facilities: Acoustic Telemetry which uses a large array of strategically placed acoustic receivers around Australia to detect and track tagged animals, and Animal Tagging, deploying CTD trackers and bio-loggers on Australian and Southern Ocean seal and bird species, collecting behavioural and physical oceanic data.

As stated in the current Acoustic Telemetry Project Plan:

“Animal Ethics: Animal Research Authorities are held individually by researchers through their home institutions and subject to State Government legislation, or Federal Government legislation for research in federally managed waters. Animal Research Authorities are not required for the deployment or retrieval of acoustic receivers, nor for use of data retrieved, but are required for tagging of fish and other vertebrates, and for cephalopods in some jurisdictions. All users of the network are individually responsible for obtaining Animal Research Authorities for their tagging work.”

As stated in the current Animal Tagging Project Plan:

“Animal ethics and welfare - All procedures used to capture seals and attach instruments are conducted under the auspices of institutional Animal Ethics Committees (e.g., Macquarie University, UTAS) using well developed Standard Operating Procedures and world best practice. Deployments at Kerguelen, Dumont D’Urville and Scott Base have Animal Ethics clearance to June 2022.”

Data contributed to the IMOS Animal Tracking Facility database may be stored under one of three access levels:

- Open Access data (subject to the condition of a CC-BY licence).
- Embargoed data (5 years maximum). Allows scientists to publish their research within a reasonable timeframe prior to their data becoming publicly available. Embargoed data is only available for viewing and downloading by members of the project (in addition to IMOS and AODN staff).
- Protected data (99 years). This status provides security where publicly available transmitter metadata and associated detection data may present threats to tagged animals or projects. Currently one project in the database has this status.

The acoustic telemetry data comprises three data groups: Receiver metadata which includes deployment date and location; Transmitter metadata which includes deployment date and location; and Detection data which includes events recorded by the receiver and detections of transmitters.

With the presence of data with disclosure risk (and to additionally manage receiver location exposure), for users accessing the IMOS Animal Tracking Database, there are two types of visibility based on user type:

- Public and Registered users can only see location GPS coordinates truncated to two decimal places.
- Project members, Animal Tracking Facility team, AODN team can see GPS coordinates as provided by the data owners.

The AODN Team, alongside the IMOS Animal Tracking Facility data manager, manage the access and integrity of the database, including back up security. All data are stored on secure servers (detailed further in R16).

For complete details, including details of roles with access to, and/or responsibilities for data – see the IMOS Animal Tracking Data Policy (https://s3-ap-southeast-2.amazonaws.com/content.aodn.org.au/Documents/IMOS/Facilities/animal_tracking/LATEST/IMOS_Animal_Tracking_Facility_Data_Policy_LATEST.pdf).

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
The response clearly addresses the more sensitive data held by AODN.

As an organization which makes data publicly available, with minimal restrictions (CC-BY, in most cases), there is little that the organization can do, apart from the high quality documentation already provided, to ensure that the data are used in compliance with disciplinary and ethical norms. The organization is, in this reviewer's opinion, taking all reasonable steps possible.

5. Organizational infrastructure

R5. The repository has adequate funding and sufficient numbers of qualified staff managed through a clear system of governance to effectively carry out the mission.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

Australia's Integrated Marine Observing System (IMOS) is enabled by the National Collaborative Research Infrastructure Strategy (NCRIS) (<https://education.gov.au/national-collaborative-research-infrastructure-strategy-ncris>). It is operated by a consortium of institutions as an unincorporated joint venture, with the University of Tasmania (<https://www.utas.edu.au/>) as Lead Agent. Both the IMOS and AODN offices are hosted by the University of Tasmania.

- The AODN is reliant on NCRIS funding. IMOS funding through NCRIS currently extends to June 2023. The NCRIS program is funded until June 2028 in Commonwealth budgets and it is anticipated that as a foundational NCRIS capability IMOS will continue to be funded for the life of the current program (i.e. June 2028).
- On average 11% of the IMOS budget is for data management. The AODN is the data management facility of IMOS, and its funding is considered at the program level within IMOS, i.e. that IMOS needs a properly resourced AODN Facility to enable open access to data (see the IMOS Five Year Plan (2022-27) at <https://imos.org.au/about/imos-five-year-plan-2022-2027>).
- AODN Full-time equivalents (FTE) - 19.9 for 2021/22 and forecasted to be similar for 2022/23.
- All staff are employed under fixed-term contracts.

AODN staff are divided into the following areas, and managed through a clear system of governance:

Management Team

- 1 Director, who is responsible for the overall running and direction of the AODN
- 1 Executive Officer, who is responsible for the effective management of operational and financial aspects of the AODN, collaboration agreement and contract management, contribution to reporting arrangements, project management support and assistance to the AODN Director
- 1 Team Leader, who is responsible for the Data Services team
- 1 Team Leader, who is responsible for the Information Infrastructure team

General Staff

- 9 Project Officers (8 managed by the Data Services Team Leader, and 1 by the Executive Officer), who are responsible for marine data and information management and for setting up services for data ingestion, discovery and access. (roles further defined as either Metadata Officer or Data Officers)
- 8 Software Engineers (all managed by the Information Infrastructure Team Leader), who perform DevOps, software development, system administration and maintenance activities.

All staff are able to access a budget for ongoing training and professional development (discussed with their team leader

as part of their annual performance review), as well as a budget for attending national/international meetings/conferences.

The 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 (<http://www.iode.org>; AODN Director is the Australian national representative for oceanographic data management)
- International Quality Controlled Ocean Database (IODE-IQuOD, <http://www.iquod.org>; a Project Officer is task team leader for formats of data products)
- Ocean Gliders Data Management Task team (<http://www.ego-network.org/>)
- Southern Ocean Observing System (SOOS) Data Committee (<http://soos.aq/>)
- Research Data Alliance (RDA) Working and Interest Groups (<https://rd-alliance.org/groups>)
- Federation of Earth Science Information Partners (ESIP; <http://www.esipfed.org/>)
- AODN is also a member of the Open Geospatial Consortium (OGC) (<http://www.opengeospatial.org/>).

AODN staff are members of around 15 national working groups or task teams, evidencing their national and international specialist skills.

For example, the following:

- IMOS Ocean Glider User Group
- IMOS moorings long time-series working group
- Forum for Operational Oceanography (FOO) Surface waves working group
- Australia and New Zealand Land Information Council (ANZLIC)/Intergovernmental Committee on Surveying and Mapping (ICSM) metadata working group
- ARDC Australian Vocabulary Special Interest Group (AVSIG)
- National Environmental Information Infrastructure – Monitoring Sites Register (NEMSR).

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
Appropriately detailed answer with useful external references.

6. Expert guidance

R6. The repository adopts mechanism(s) to secure ongoing expert guidance and feedback (either inhouse or external, including scientific guidance, if relevant).

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The AODN has access to varied sources for expert guidance, internally via IMOS staff, and externally from both national and international contacts in the marine community and more broadly from the research community.

Internally the AODN can leverage advice from staff within the IMOS office, and as IMOS is operated by a consortium of institutions, the AODN also has access to the various IMOS facility staff that are spread across institutions all over Australia.

The IMOS Science and Technology Advisory Committee (<http://imos.org.au/nodes/imosstac/>) plays a critical role for IMOS as it represents the scientific opinion of the marine and climate research community, and will work to advise the IMOS Office on the scientific rationale and direction of a national observing system.

- The committee will provide advice and assessments to the IMOS Office on the technical implementation by facilities and sub-facilities, and on the scientific merit of research undertaken with IMOS data.
- The committee will also work to advise the IMOS Office on the development, application, and implementation of new technologies within the national observing system.

In 2020/2021 IMOS commissioned an external review of the AODN, and based on the findings, a plan has been established with incorporation into upcoming AODN annual implementation plans.

The AODN Technical Advisory Group (TAG) provides a forum for developing data management and data publishing standards for the Australian marine community. It provides technical advice to the AODN and to organisations wishing to publish marine data. Part of the TAG's role is to identify areas for development and to establish working groups to pursue the development. The TAG is comprised of science and disciplinary experts from within the IMOS community and the broader AODN community (from both Australia and New Zealand).

The AODN actively engages in national and international activities that promote the use and uptake of standardised interoperable data. Information technology is a rapidly changing area and new approaches are being developed to meet

the challenges of the explosion in data volume and its delivery methods. It is important that the AODN staff keep abreast of these, with one of the main partners being the ARDC. The involvement of AODN staff in various national working groups or task teams is outlined in R5.

We regularly participate in international conferences, and initiatives to foster collaboration and relationships with other experts in data management. We are members of various international fora for information exchange on developing standards, for example (further outlined in R5):

- Research Data Alliance (RDA; <https://rd-alliance.org/>) e.g., working group on Data Citation, and interest group on Vocabulary Services
- Federation of Earth Science Information Partners (ESIP; <http://www.esipfed.org/>)
- The AODN is a member of the Open Geospatial Consortium (OGC; <http://www.opengeospatial.org/>) and a partner in
- Ocean Data Interoperability Platform (ODIP; <http://odip.org/>), the international (EU-USA-Australia) project which contributes to the removal of barriers hindering the effective sharing of data across scientific domains and international boundaries.

The AODN seeks feedback from users, scientists and the broader research community and public using three avenues – IMOS website, IMOS newsletter, and varied outreach activities. The AODN communicates with its scientific users through various outreach activities – such as giving seminars/lectures (e.g. regular presentations to IMAS Undergraduate/Honours/Masters students), running workshops (e.g. the annual Sydney Institute of Marine Science (SIMS) Topics in Australian Marine Science (TAMS) Course Practical - An Introduction to IMOS) on how to use the metadata/data tools, regular news articles (<https://imos.org.au/facilities/aodn/aodn-news>) and IMOS social media posts.

The AODN also provides a 'Help desk' for users, and people can make general enquiries, obtain help or provide feedback. These contact details are provided on both the AODN webpage (<https://imos.org.au/facilities/aodn>), and other online infrastructures. Users are provided with both a phone number and email address. Furthermore, users accessing our GeoNetwork catalogues are able to provide specific 'reviews' of individual metadata records, both by rating according to categories and/or by provision of text.

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
Clearly meets the guideline. External reviews are particularly useful in meeting this item, as completed by this organization.

DIGITAL OBJECT MANAGEMENT

7. Data integrity and authenticity

R7. The repository guarantees the integrity and authenticity of the data.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:

3 – The repository is in the implementation phase

Reviewer 2

Comments:

3 – The repository is in the implementation phase

Response:

The AODN is committed to ensuring the integrity and validity of the data held in its repository (see the AODN Preservation Strategy (<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-preservation-strategy>), further detailed in R10). Processes to ensure this, include ingest controls, audit trails and metadata management.

Ingestion

Providers who regularly upload data register for access via File Transfer Protocol (FTP) over Transport Layer Security (TLS) or rsync. This involves direct communication with AODN staff, so that the identity of the depositor and the nature of the data they will provide are known. Only registered users are able to upload data to the AODN. When using rsync, checksums are always used to verify that a file was transferred correctly. When using FTP over TLS, there is negligible chance the file contents could get corrupted while transferring. The risk of data corruption is about the same as if copying the files between two local drives. There is a significant portion of our data providers that require to use FTP rather than rsync.

A minority of data providers provide and update their data sets on their own institutional FTP site or other service, which AODN periodically downloads to maintain an up-to-date archive.

The data ingestion process is fully automated using Python code that is version-controlled, peer-reviewed and Open-source (on GitHub: [python-aodncore](https://github.com/aodn/python-aodncore) (<https://github.com/aodn/python-aodncore>), [python-aodndata](https://github.com/aodn/python-aodndata) (<https://github.com/aodn/python-aodndata>), [python-aodntools](https://github.com/aodn/python-aodntools) (<https://github.com/aodn/python-aodntools>)). The process includes standard steps for compliance checks, harvesting data and metadata into a database, and storing data in our archival storage (AWS S3). It can be tailored for each data stream/collection.

Data files published in NetCDF (network Common Data Form) format (with few exceptions), are validated for completeness and compliance with the Climate and Forecast (CF) metadata conventions (both IMOS and AODN data), and IMOS conventions (an extension of CF) (IMOS data only). Compliance checks (see R8) are performed using the US IOOS Compliance Checker software (<https://github.com/ioos/compliance-checker>) (both IMOS and AODN), and our own IMOS plugin (<https://github.com/aodn/cc-plugin-imos>) (IMOS only).

Received data files are not modified, except to assist data providers to meet our format and compliance requirements (see R8) by

- Converting from a supplied format (e.g. comma-separated values (CSV) to NetCDF for archival); or
- Fixing minor metadata compliance issues.

These modifications, where necessary, are put in place in consultation with each data provider and performed automatically by peer-reviewed code. Changes are only made to file format, structure and metadata, not to any measured or computed data values. The original data, as provided, is also archived.

Some data is entered by users via a web-interface and managed in databases. If data entry errors are identified (either by data contributors or AODN staff), corrections are discussed with data contributors (and the discussion archived in a GitHub repository), and corrections are applied to the database where necessary. Ongoing improvements to the web-app/s aim to reduce data entry errors.

Data files are stored on Amazon Web Services (AWS) Simple Storage Service (S3) for archival and public access. See R9, R10 and <https://docs.aws.amazon.com/AmazonS3/latest/userguide/DataDurability.html> for the benefits this provides, including

- Reliability
- Long-term preservation: data will not be affected by data degradation or other digital preservation concerns, because it is stored in multiple locations with a custom distributed filesystem that does checksums and parity while managing the data.
- Versioning (backup of all previous versions).

Data can be downloaded from a variety of services that all use Hypertext Transfer Protocol (HTTP) by default currently. However, we already have Hypertext Transfer Protocol Secure (HTTPS) enabled which reduces the chance of data content being corrupted during transfer. We are planning to soon set up automatic redirects so that HTTPS is always used by default.

Provenance / Audit trails

Supplying provenance information is the responsibility of each data provider. Any provenance information supplied by data providers is retained and archived with the data (e.g. in the form of NetCDF attributes), or recorded in metadata records.

Detailed logs of each ingestion event - from a file's first appearance at AODN through to its publication - are recorded on the AODN production server.

For data products generated during ingestion, specific information (such as links to the original data, and to the version of the code that generated the product) is included in each output NetCDF file.

Metadata

Data collections are described by ISO 19115-3 metadata records (see R13), and include data access links (specific to each collection) in the form of

- Open Geospatial Consortium (OGC) standard Web Map Services and Web Feature Services (WMS/WFS)
- Relevant folders on the AODN THREDDS server
- Relevant folders on the AODN public file server.

The Metadata Officer creates all IMOS funded metadata records, and ensures they are comprehensively described.

External researchers who create metadata using the AODN Data Submission Tool (see R8) are unable to submit their work until all mandatory elements are filled, ensuring a complete metadata record is created. Guidance in metadata creation (and workflows are covered in R12) is documented in online guides -

<https://help.aodn.org.au/public-documents/aodn/metadata-guides/>, and for the usage of the AODN Data Submission Tool

- https://metadataentry.aodn.org.au/submit/media/guide/AODN-DaST-UserGuide_V2.pdf.

Metadata records are updated when required by the Metadata Officer. Small changes not relevant to the usage of data are made with no reference, however if changes are made to the data and/or its availability, this information is added to the metadata with reference to the timing of the change. Substantial metadata edits (and timing) are documented under the metadata maintenance element. Versioning management is facilitated by monthly back-ups of the database backing GeoNetwork. Metadata changes relevant to data stored in databases, are archived in GitHub.

Reviewer Entry

Reviewer 1

Comments:

Accept.

Reviewer 2

Comments:

Reviewer thinks the compliance level might even be 4 because:

File hash checks are a useful adjunct to the controls that the provider has in place, particularly to validate the integrity of the transfer and as a guard against inadvertent modification and some forms of malicious actions. The transfer checks provided by https and within S3 storage are valuable controls and do serve as protection for many transfer errors and for "bit rot".

The controls for authenticity of the data are strong. It is not clear to me that anything critical is missing, so this response reads to me that the repository is compliant. There are improvements that could be made (like fixity checks based on hashes), but there are complementary controls in place that accomplish most of the same purposes.

8. Appraisal

R8. The repository accepts data and metadata based on defined criteria to ensure relevance and understandability for data users.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The AODN's primary role is to host, manage and archive data produced by the IMOS facilities, in effect, all IMOS data. In regard to AODN data, the aim of the AODN is to make as much marine data as possible, publicly funded as well as data from private industry and not-for-profit organisations, accessible and freely available. In the latter case, on occasion, a researcher or an organisation is also looking to archive their data with the AODN, due to the lack of a suitable repository within their own organisation or research institution. In addition, the AODN will host data if they have been requested to provide a DOI (see R10, R12). On the rare occasion that the AODN has been approached to host data that falls outside our mission/collection profile, we have assessed the data, and with our knowledge of the larger data community within Australia have been able to provide relevant suggestions for the researcher/organisation to contact.

All IMOS data have corresponding metadata records created by the Metadata Officer. IMOS metadata is created by the Metadata Officer in conjunction with IMOS facility staff and the relevant AODN Project Officers, to ensure it is of the highest quality, and adequately describes the associated data collection. As the majority of IMOS data is stored in NetCDF format, additional metadata is provided within the files to aid the interpretation and utilisation of the data.

Researchers or organisations looking to archive AODN data with the AODN, are unable to do this without a complementary metadata record. Contributors are encouraged to use the online AODN Data Submission Tool (<https://metadataentry.aodn.org.au/submit/>), to create the metadata record. The metadata tool has been designed to provide an intuitive user interface to guide the user in metadata creation. In order to successfully submit they need to have successfully completed all mandatory fields, with these elements clearly identified during the process of metadata creation. The metadata records submitted via the tool are to a minimum standard, however the Metadata Officer will still

manually check the metadata record for completeness and ensure that it adequately describes its associated dataset. Once the Metadata Officer is satisfied (often after negotiation with the submitting user), the metadata record is published (either in the IMOS or AODN metadata catalogues) and becomes publicly available. Owing to the diverse range of scientific disciplines covered by the AODN, the Metadata Officer is unable to comprehensively judge whether a metadata record completely describes a dataset. However, every care is taken to ensure that the metadata record is as detailed as possible.

Whilst metadata creation by the AODN is a manual process, users are unable to submit a metadata record via the online tool until all mandatory elements have been completed, which in turn ensures that the metadata will validate against the relevant ISO 19115-3 metadata schema in the GeoNetwork catalogue, which meets requirements for usability and integration. IMOS metadata is created by the Metadata Officer, which whilst a manual process, is done to a consistent standard ensuring adherence to the schema.

The AODN exposes data in a wide variety of formats, with a preference for durable formats that are endorsed and published by standards agencies, and they are accompanied by the appropriate metadata. The AODN can accept data in any recognised format, see AODN Preservation Strategy (<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-preservation-strategy>) for further information. The majority of IMOS data (and many AODN contributors) is stored using the NetCDF system, a set of software libraries and machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data. For NetCDF data submissions, we use the IOOS Compliance Checker (<https://github.com/ioos/compliance-checker>), together with our own IMOS plugin (<https://github.com/aodn/cc-plugin-imos>) to check incoming NetCDF files against the CF-1.6 and IMOS-1.4 conventions (currently, for IMOS data only), with this process peer reviewed (also see R7). Any file that fails the checks is rejected, and feedback is sent to the data provider. If the data provider attempts to deposit in a non-preferred format, the Data Officer will work with them to transition to a suitable format or provide enough supporting documentation or access to software if they are unable to do so.

To date, the AODN has not had to remove datasets with associated persistent identifiers. In regard to IMOS datasets that have been made unavailable via the AODN Portal, whilst the method to access the data has changed (the data remains accessible still, or has been archived), the AODN documents this change in the metadata record, so users are able to understand its absence.

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
Some evidence for this is also described in the previous response, including checks for CF compliance (a form of automated check for adherence to relevant schemas).

9. Documented storage procedures

R9. The repository applies documented processes and procedures in managing archival storage of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository
Response is complete.

Response:

We maintain an internal wiki to document our processes and procedures for the benefit of current and future staff. Changes are peer reviewed and version controlled. Diagrams and step by step instructions on the data-harvesting process are maintained in editable formats. These pages are unable to be made public, but the index of the wiki is annexed to this form.

Data is stored using Amazon Web Services (AWS). As per the AWS shared responsibility model (<https://aws.amazon.com/compliance/shared-responsibility-model/>), AWS has responsibility for the physical storage of the files, including ensuring consistency across archival copies and handling deterioration of storage media. The two AWS services used are AWS Simple Storage Service (S3) and Amazon Elastic Block Store (EBS). S3 is used to store files, while EBS is used for databases. Our primary database is replicated to a server on the Australia Government supported Nectar Research Cloud.

We use AWS S3's object versioning functionality to keep backups of all files in all previous versions. Each version of the data is replicated multiple times across the S3 system to mitigate redundancy. Databases are backed up as SQL dump files to S3 daily for short-term recovery, monthly backups are kept for long-term recovery. In the future we hope to replicate our backups to non-AWS infrastructure as an extra level of protection. One of the responsibilities of our dedicated AODN DevOps team is to be aware of potential risks, however, we don't follow a specific Risk Management framework.

ANNEX A details the content of the AODN Internal Discussions wiki

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:

10. Preservation plan

R10. The repository assumes responsibility for long-term preservation and manages this function in a planned and documented way.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:
4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:
4 – The guideline has been fully implemented in the repository

Response:

The AODN Preservation Strategy (<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-preservation-strategy>) sets out the approach to the long-term retention and preservation of the data hosted by the AODN and gives guidance on implementing those principles. The AODN hosts data collected by IMOS and AODN (non-IMOS) data, by exception. The purpose of the strategy is to document initiatives to improve access to and re-use of data hosted by the AODN, and to facilitate systems and services to enable the efficient and effective management of the assets.

Unless otherwise stated the preservation strategy applies to the dataset collections and derivatives hosted by the AODN, all data contributors including IMOS staff and researchers, and non-IMOS and affiliate contributors and contributions for the duration of IMOS funding. The level of responsibility for the preservation of each item is equal. The AODN manages a small number of DOIs (<https://imos.org.au/facilities/aodn/does-and-data-deposits-with-the-aodn>) and assumes the commitment to preserve the resource and the capability to ensure continuity of access.

The AODN makes use of Amazon Web Services (AWS) as a cloud computing platform and for scalable data storage in

the cloud. Most data hosted by the AODN will be on S3 (object) storage allowing multiple versions of the data to be stored. Versions are stored using the S3 Glacier storage class for long-term, secure, durable storage for data archiving, providing 99.999999999% (11 nines) of data durability. Each version of the data is replicated multiple times across the S3 system to mitigate redundancy. Multi-factor authentication (MFA) delete is enabled on our data bucket which adds an additional layer of security for changing Bucket Versioning settings and permanently deleting object versions. We intend to keep our data on AWS in the long-term and so have no plans for future migrations. We successfully migrated our data to AWS in 2015 and are confident we'll be able to migrate back out again if needed. We however, do have plans to replicate our backups to non-AWS infrastructure, which could be regarded as a form of migration.

The AODN pursues data in durable formats that will be accessible into the future; that are publicly documented with complete authoritative specifications with a preference for formats that are endorsed and published by standards agencies. To mitigate against formats becoming obsolete over time, non-proprietary, open, documented formats most likely to be accessible in the future are considered for the long-term preservation of data. Preference, at least for IMOS data files and databases, are NetCDF and PostgreSQL respectively, both widely used and accepted within the discipline of data science.

In the event that data are only available in unsuitable, proprietary formats, then details on software required will also be provided. Archival in such formats is strongly discouraged.

The contractual constraints on IMOS facilities are clearly described within the IMOS Data Policy, and joint responsibilities articulated there and in the IMOS Workflows (<https://help.aodn.org.au/public-documents/imos/facilities/workflows/>). The roles and responsibilities in regard to contributors of non-IMOS data that is hosted by the AODN and/or exposed via the AODN are clearly articulated in the AODN Data Policy. All data that is hosted by the AODN is owned by the originator (individual or organisation). The AODN has no rights to the data, with its role primarily to host the data and meet its responsibilities.

In regard to the IMOS data managed, contractually actions relevant to preservation are ensured. For the small amount of non-IMOS data that the AODN has agreed to host, significant work is undertaken during the submission process (refer R8) to address the necessary preservation requirements. Active management of the IMOS data ensures that AODN data is managed to the same standard as a by-product.

Additional documents that promote the AODN data preservation intentions and the role preservation plays in fulfilling the AODN mission include:

- IMOS Data Policy (<https://imos.org.au/imos-data-policy>)
- AODN Data Policy (<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-data-policy>)
- IMOS Five Year Plan (2022-27) (<https://imos.org.au/about/imos-five-year-plan-2022-2027>; refer pg. 1 and 17)

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:

A complete and appropriately detailed response.

11. Data quality

R11. The repository has appropriate expertise to address technical data and metadata quality and ensures that sufficient information is available for end users to make quality-related evaluations.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

The AODN repository comprises a full-time Metadata Officer, with the remainder of the AODN Project Officer team being data focused with a wide breadth of skills. When required, the team can source advice from the IMOS Senior Science Officer. The team are involved in both national and international groups, for example relevant ARDC Communities of Practice (<https://ardc.edu.au/resources/communities-of-practice/>) and the Research Data Alliance (RDA), to ensure the repository is up to date with quality standards on data archiving and data curation.

IMOS data

All IMOS metadata is created and managed by the Metadata Officer, a process which ensures a consistent approach. If further information is required to complete the metadata record, the Metadata Officer will consult with other members of the IMOS office or may reach out to relevant IMOS facility staff. Any information relevant to data quality is included in the metadata record, or links are provided (this is in addition to any data quality information included in the metadata within individual NetCDF files). Vocabularies are utilised in metadata creation (also available in the AODN Data Submission Tool, see below) for a range of fields (e.g., keywords, data parameters and organisations) to decrease input error and ensure consistent entry.

As a central repository for a wide range of data products, we do not have the required expertise to assess the technical quality of all these products. Instead, the providers of each data product are required (contractually, in the case of IMOS-funded data) to perform a minimal level of quality assurance / quality control.

As described in R7/R8, self-describing NetCDF data files are checked for completeness and compliance with the Climate and Forecast (CF) metadata conventions, and IMOS conventions (an extension of CF) before publication. These conventions include a detailed description of data quality using standard attributes and status flags. Any files that do not meet the conventions are rejected, and the provider notified with detailed reasoning so that they can fix the issues. We also publish various (mostly real-time) data streams where no quality assessment has been performed. In these cases, the uncertainty about data quality is made clear in the metadata, and through file naming conventions.

AODN data

AODN contributors wanting to store their data on AODN infrastructure will in most cases contact the AODN office to first enquire about the process. If they are also seeking a DOI, information is available at <https://imos.org.au/facilities/aodn/does-and-data-deposits-with-the-aodn> (for more information see below). These are often one-off submissions and vary by submitter.

Contributors will need to create a metadata record to accompany the data and users are directed to utilise the AODN Data Submission Tool (<https://metadataentry.aodn.org.au/submit/>) (for further details on this workflow see R12). After completion, the user will submit the metadata record (only possible after a suite of mandatory elements have been entered). This metadata record will be reviewed by the Metadata Officer, and like the data workflow below, this may involve an iterative process between the contributor and the Metadata Officer to bring the metadata up to the required standard, also ensuring that it adequately describes its associated dataset.

In relation to the dataset, a relevant Data Officer will contact them and arrange transfer of the first draft/subset of files for initial assessment. For data stored in NetCDF files, these are validated for completeness and compliance with the Climate and Forecast (CF) metadata conventions. An iterative process will assess and document the changes required, to be made by the contributor, for the files to pass the checks.

The data that is submitted to the AODN under these circumstances can cover a broad range of scientific disciplines. Where possible, the Data Officer assigned to check the data will be within their recognised area of expertise. Outside this, Data Officers are unable to comment on the quality of the scientific work undertaken.

Should a member of the general scientific community (designated community) wish to contact the AODN about a particular dataset, there are various mechanisms by which this can be achieved. On our tools, such as the AODN Portal, 'Contact' is the link to be used, otherwise the AODN's contact details are listed under metadata contact on all our metadata records. Each metadata record also contains a point of contact (often the primary researcher), and occasionally

a second contact, sometimes another researcher involved with the data, or the generic contact for the AODN.

All AODN contacts are made to our general enquiries email address: info@aodn.org.au, which doubles as the Help Desk. Comments and feedback about specific datasets made to the AODN, are forwarded onto the relevant researcher/organisation. These requests are documented in an internal GitHub repository for tracking and resolution.

IMOS is currently submitting best practice documentation (data quality manuals and reports) to the Ocean Best Practices System (OBPS) repository (<https://www.oceanbestpractices.org/>). Users of the repository are able to search and discover these community ocean best practices, and are able to evaluate and provide feedback on specific documentation. Access to this documentation is available on the relevant facility and sub-facilities' pages on the IMOS websites and links added to relevant metadata records.

To assist the data user's understanding of the data, the Metadata Officer, and researchers submitting data, can add links to related websites, code repositories, publications (journal articles, reports) and other related data collections. If required, researchers can add file(s) to further explain the data, such as a readme file.

Every metadata record has a suggested citation (either generic or specific) that researchers can use to cite that metadata record and associated dataset. The AODN has a DOI minting service (set up with assistance from the ARDC) that gives it the capability to create dataset DOIs and DOIs for internal IMOS documentation. If a dataset has a DOI, it is included in the suggested citation in the metadata record. If users discover IMOS/AODN datasets with DOIs, through DataCite (<https://search.datacite.org/>), the number of citations recorded against that DOI are displayed. The AODN is also included in Clarivate Analytics' Data Citation Index (<https://clarivate.com/webofsciencegroup/solutions/webofscience-data-citation-index/>) on the Web of Science platform.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
A complete and appropriately detailed response.

12. Workflows

R12. Archiving takes place according to defined workflows from ingest to dissemination.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

IMOS data

The AODN, the data management facility for IMOS, has developed workflows (<https://help.aodn.org.au/public-documents/imos/facilities/workflows/>) for each IMOS facility/sub-facility to describe the flow of IMOS data from planning, through data collection, to data delivery and public data access.

The primary goals of the workflows are to:

- Improve data flow and data hand-off, making tracking of data status easy and preventing data loss
- Identify and delimit precisely the responsibilities of each person involved
- Improve communication at the interface between IMOS facilities (i.e. between AODN and other IMOS facilities)
- Improve transparency for end users by providing more details to populate metadata records (i.e. limitations and processing methods applied to datasets)
- Assist in reporting planned deployments against actual deployments and data delivery.

The majority of IMOS workflows are documented, with plans to comprehensively update all in this year's AODN Implementation Plan. Regular yearly reviews are planned after the updates.

All IMOS data are described by metadata records in the IMOS metadata catalogue (<https://catalogue-imos.aodn.org.au/geonetwork/srv/eng/>). All IMOS metadata is created and managed by the AODN Metadata Officer, a process which ensures a consistent approach.

AODN data

As the AODN is recognised as the 'go to place for marine data', and as an approved Nature Scientific Repository (<https://www.springernature.com/gp/authors/research-data-policy/repositories-earth/12327148>), we also occasionally handle data archival for Australian marine scientists who do not have an appropriate repository for archiving their data, and/or are seeking to acquire a DOI for their dataset, which requires the AODN to manage the data (see R10 for further information). A very small percentage of the data held by the AODN falls under this category.

To assist metadata creation in these situations, the AODN directs users to utilise the AODN Data Submission Tool (<https://metadataentry.aodn.org.au/submit/>). The tool provides a user-friendly interface to guide users in the creation of both metadata to describe their data, and where applicable, data submission in the one process. Help documentation is provided to users in the use of the tool, and if further assistance is required, the Metadata Officer is the point of contact. The AODN Data Submission Tool requires that users create a log in, and after a user has successfully submitted their metadata and data for upload, this process is not automated but the submission will be checked over by the Metadata Officer and a relevant Data Officer (dependent on the data type and source). Both the metadata and data will not be uploaded to the relevant metadata catalogue (IMOS or AODN), and data storage location, until both have been qualitatively and quantitatively checked.

The AODN maintains a data policy (<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-data-policy>) which clearly documents the obligations for both scientists and AODN staff with regards to data management, with more specific information regarding DOI related data submissions covered here – <https://imos.org.au/facilities/aodn/does-and-data-deposits-with-the-aodn>.

The data that is submitted to the AODN under these circumstances can cover a broad range of scientific disciplines. Where possible, the Data Officer assigned to check the data will be within their recognised area of expertise. Outside this, Data Officers are unable to comment on the quality of the scientific work undertaken. The Data Officer will work with the user to upload the data in the most appropriate format. Where possible, files may be saved in different formats to increase their likelihood of long-term preservation. Whenever this is done, copies of the original file are also kept.

AODN partners are encouraged to make their data holdings discoverable via the AODN Portal, alongside the IMOS data. Clear guidance to assist partners is provided at the following location - <https://help.aodn.org.au/contributing-data/>.

All internal AODN related workflows and processes (relevant to both IMOS and AODN data) are documented in an internal document (relevant pages of which have been annexed and are available on request to the AODN: contact info@aodn.org.au). Should a workflow or process change, then it is updated in the document. The bulk of the metadata management work is carried out by one staff member (the Metadata Officer), which simplifies the process of managing change.

ANNEX B details the content of the AODN Internal Discussions Issue: Project Officer's workflow

Reviewer Entry

Reviewer 1

Comments:

Accept. It is great to see an effort to review and maintain the documented workflows.

Reviewer 2

Comments:

A complete and appropriately detailed response.

13. Data discovery and identification

R13. The repository enables users to discover the data and refer to them in a persistent way through proper citation.

Compliance Level:

3 – The repository is in the implementation phase

Reviewer Entry

Reviewer 1

Comments:

3 – The repository is in the implementation phase

Reviewer 2

Comments:

3 – The repository is in the implementation phase

Response:

The IMOS and AODN metadata and data holdings are discoverable via several different local methods:

- The AODN Portal (<https://portal.aodn.org.au/>) enables users to discover data collections (both IMOS and AODN), create a subset of the data, and download it. Collections on the AODN Portal can be discovered using the following facets: 'Parameter', 'Organisation', 'Platform', 'Date', 'Geographic Boundary' and text ('Keyword') search.
- The IMOS metadata catalogue (<https://catalogue-imos.aodn.org.au/>) enables users to search all IMOS metadata records in its GeoNetwork catalogue. Users can filter/search by the following means: 'Type of resources', 'Topics', 'Keywords', 'Contact for the resource', 'Provided by', 'Years', geographically and free text.
- The AODN metadata catalogue (<https://catalogue.aodn.org.au/>) enables users to search its full AODN holdings: IMOS metadata, hosted AODN metadata, and harvested AODN metadata. This GeoNetwork catalogue, offers the same filtering and search capabilities of the IMOS catalogue.

All IMOS metadata holdings and hosted AODN metadata were recently transformed (from the Marine Community Profile (MCP) of ISO19115/19139) to the international standard, ISO 19115-3 to allow for greater interoperability. Harvested AODN metadata comprises a variety of metadata standards/profiles, with the majority already ISO 19115-3, or planning to transform their MCP to ISO 19115-3.

Both the IMOS and AODN catalogues are harvested by several other organisations:

- Australian Research Data Commons (ARDC) – Research Data Australia (<https://researchdata.edu.au/>)

- Australian Government - central source of Australian open government data: <https://data.gov.au/>

Each harvested metadata record includes a 'Point of Truth' URL to direct the user back to the "original copy" in the IMOS/AODN catalogue (or source metadata catalogue if it is harvested itself into either the IMOS or AODN catalogues).

The AODN repository is registered with re3data.org (<https://www.re3data.org/repository/r3d100010914>), DataCite (<https://search.datacite.org/repositories/ardcx.aodn>), it has an entry in FAIRsharing, the online registry of scientific data standards, databases and policies - <https://fairsharing.org/FAIRsharing.j5eden>, and it is a Nature recommended repository – Earth, Environmental and Space sciences (<https://www.springernature.com/gp/authors/research-data-policy/repositories-earth/12327148>).

All IMOS and hosted AODN metadata records (and the majority of AODN harvested metadata records) have a recommended data citation provided alongside a Creative Commons licence (and we plan to make this more prominent based on the outcomes of our current Architecture Review). This information is also provided to users of the AODN Portal when the user selects data to download and accepts the licence agreement and terms. If data users choose to download data directly from the IMOS THREDDS Data Server (<http://thredds.aodn.org.au/thredds/catalog.html>), the suggested citation is provided in the global attribute 'citation' in the NetCDF file/s that they access. Further, the AODN provides guidance on 'Data Licencing and Acknowledgment'

(<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-data-licencing>) on its website, and through its online Portal User Guide - <https://help.aodn.org.au/user-guide-introduction/aodn-portal/data-use-acknowledgement/>.

The AODN has the capacity to mint DOIs for its datasets. The AODN manually mints DOIs (via DataCite; <https://datacite.org/>) upon request in relation to AODN data submissions (some include IMOS data), with the majority of these related to planned data journal submissions (mainly Nature's Scientific Data; <https://www.nature.com/sdata>), and which represent static data or snapshots of dynamic data collections. As the IMOS data collections are of a dynamic nature, we don't currently assign DOIs to the data. The AODN are considering the options available (namely following the recommendations of the Research Data Alliance (RDA) Data citation working group; <https://www.rd-alliance.org/groups/data-citation-wg.html>), with plans to scope the work in the next financial year, with an implementation in place in a few years.

Reviewer Entry

Reviewer 1

Comments:

Accept. It is good to see plans for applying persistent identifiers to the dynamic time-series datasets.

Reviewer 2

Comments:

Concur with being in the implementation phase with appropriate actions defined.

14. Data reuse

R14. The repository enables reuse of the data over time, ensuring that appropriate metadata are available to support the understanding and use of the data.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

As noted in R13, all IMOS metadata holdings and hosted AODN metadata were recently transformed to the international standard, ISO 19115-3 to allow for greater interoperability. All IMOS metadata records hosted by IMOS (some records are produced by CSIRO and AIMS) are curated by the Metadata Officer, ensuring that the records are richly populated with information. For metadata hosted by the AODN, creation of this metadata occurs via the AODN Data Submission Tool (<https://metadataentry.aodn.org.au/submit/>), which requires a minimum amount of metadata (mandatory elements dictated by the metadata schema) content before the user is able to submit (User guide - https://metadataentry.aodn.org.au/submit/media/guide/AODN-DaST-UserGuide_V2.pdf). Curation of this metadata occurs after submission by the Metadata Officer, to ensure a rich metadata record. The AODN also provides guidance on metadata creation, in both a short 'Quick Guide' format and a more comprehensive GeoNetwork guidance (<https://help.aodn.org.au/public-documents/aodn/metadata-guides/>) – note there are plans, in the near future, to update the 'Quick Guide' to reflect the new ISO 19115-3 metadata schema in use, and due to users utilising the AODN Data Submission Tool now for primary metadata creation, the full guide will be reduced.

The AODN updates their IMOS metadata holdings as required, e.g., change in data collections sites, new contact details for 'point of contact' etc. See earlier requirements for details on the process of archiving data and metadata in the AODN.

As data can be accessed from several different points in the repository, the provision of access to full metadata records does differ. If data is accessed via the IMOS or AODN metadata catalogues, users already have the full metadata record on display. If a user accesses data through the AODN Portal, links to the collection level metadata record is provided on each step of the Portal, with plans in the future to provide the metadata as part of the package of downloaded data. However if a user accesses data via our S3 storage (<http://data.aodn.org.au/>) or our IMOS THREDDS Data Server (<http://thredds.aodn.org.au/thredds/catalog.html>), there are no direct links between the data files and the collection level

metadata record. However, for most IMOS data types (and many AODN contributors), the NetCDF system is used, which provides metadata information bundled in with the data. Metadata entry for IMOS NetCDFs utilise the IMOS NetCDF Conventions (https://s3-ap-southeast-2.amazonaws.com/content.aodn.org.au/Documents/IMOS/Conventions/IMOS_NetCDF_Conventions.pdf) (and AODN submissions are encouraged to consult these conventions where applicable) which describes the NetCDF format and the format of filenames that will be used to distribute IMOS data. NetCDF is widely used (<https://www.unidata.ucar.edu/software/netcdf/usage.html>) and has a number of contributors (<https://www.unidata.ucar.edu/software/netcdf/credits.html>) ensuring its continued development. Unidata reports that access to all earlier forms of NetCDF data will be supported by current and future versions of the software.

NetCDF is a well-accepted and long-lived data format, however not all data is requiring of this standard. The AODN can expose data in a wide variety of recognised formats (see R8 and R10 for further information), as long as it's accompanied by the appropriate metadata. One-off data submissions of AODN data, are rarely large datasets, so submitters are encouraged to supply data in "long-lived" formats (e.g. Excel to CSV), and if not, formats are updated to more modern versions (e.g. .xls to .xlsx). There are currently no plans to audit AODN data on a regular basis to account for evolution of formats, IMOS data on the other hand, is actively managed, and storage locations are updated as required.

Further information can be found in the AODN Preservation Strategy (<https://imos.org.au/facilities/aodn/aodn-data-management/aodn-preservation-strategy>) and see R10 for more information on the preservation plan.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
A clear and appropriately detailed response.

TECHNOLOGY

15. Technical infrastructure

R15. The repository functions on well-supported operating systems and other core infrastructural software and is using hardware and software technologies appropriate to the services it provides to its Designated Community.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

Our software runs on two well-supported operating systems: Ubuntu and Amazon Linux. We don't maintain physical infrastructure, instead using cloud-based servers on Amazon Web Services (AWS) and the Australian Government supported Nectar Research Cloud.

As much as possible we use pre-existing community standards. Our primary external facing application stack, the portal stack (that underpins our AODN Portal), which is used for the public discovery and consumption of our data, is largely made up of Open-source Geospatial Foundation applications: GeoNetwork, GeoServer, Geowebcache, PostGIS and Openlayers. These applications implement standards of the Open Geospatial Consortium (OGC) such as Web Map Service (WMS) and Web Feature Service (WFS). Our primary internal application stack, the pipeline stack, which is used for the ingestion of internal and external data into our systems, primarily uses open-source python libraries. Our preference is to use off-the-shelf open-source products where possible. When the pre-existing functionality of open-source software is not fit for our purposes, we either develop our own systems or contribute enhancements to pre-existing community-supported software. We maintain our own partial implementation of the OGC Web Processing Service (WPS) standard since we require a more serverless and scalable solution than is available elsewhere. It deviates from the standard in its lack of compulsory features, but the features which have been implemented conform correctly. External community software we have made contributions towards include GeoNetwork, Geowebcache, GeoServer, Unidata's Thredds Data Server and the ncWMS implementation of OGC Web Map Service.

Our gridded data are stored as NetCDF files which are automatically tested for compliance to the "Climate and Forecast" (CF) metadata conventions (see R7/R8 for further details), which are designed to promote the processing and sharing of files and are widely used within earth-science domains.

Our metadata holdings (plus harvested content from other Australian marine organisations) are maintained and managed in GeoNetwork catalogues. With our recent upgrade to version 3, we now ensure we are utilising the latest stable version and remain up to date with the latest features and fixes. This upgrade also allowed us to transition to the new ISO 19115-3 metadata standard, which improves interoperability with national and international partners, and was done in collaboration with other AODN partners.

Every financial year, the AODN plans work based on the objectives of the current AODN Implementation Plan. Every year, operational and maintenance activities are incorporated to ensure the continual improvement and upgrading of our infrastructure. In conjunction to this, we maintain a list of upgradable software with estimations of how much effort the upgrade will take and the urgency of the upgrade. Our software development team follows AGILE processes, specifically the SCRUM methodology using 2 or 3-week iterations. Potential future improvements to our infrastructure are logged as backlog items which are then prioritized and eventually included in an iteration. After the addition of new functionality, a system testing process is undertaken to confirm that nothing has been broken. We don't maintain an overarching inventory of all our systems, instead opting to maintain inventories of each application stack separately.

We generally intend to move more of our systems to a serverless framework and improve the genericness and re-usability of our in-house software. We monitor the latest AWS developments to identify opportunities to use the latest cloud technologies. We generally maintain system documentation on an internal wiki and internal GitHub repositories. External-facing documentation for our open-source software is available on our public GitHub repositories (<https://github.com/aodn/>).

We have monitoring set up for all our production infrastructure and public-facing services as well as for most of our internal services and non-production infrastructure. This monitoring allows us to confirm that our services are providing sufficient availability, bandwidth, and connectivity to our users. Where possible our services are shared across AWS availability zones, meaning that failures in one data centre should be automatically dealt with. Most of our infrastructure is defined in CloudFormation files, AWS's Infrastructure-as-code solution. Our config is maintained via Ansible and Chef, both configuration-as-code solutions. Thanks to having our infrastructure and configuration stored as code, systems lost in a disaster should be able to be re-deployed with confidence that the new system will replicate the old. With these mechanisms in place, we are confident of our ability to recover from most conceivable disasters, despite not maintaining a disaster recovery plan.

To confirm that our primary software stack can meet periods of higher-than-usual user demand, we undertake regular stress testing of our systems. Due to the serverless nature of most of our software it is easy for us to scale up and down depending on demand. For this reason, financial resources are our main consideration when demand for our services increases. Our serverless services are configured to scale up to a maximum level which we feel represents a realistic maximal level of legitimate use.

Reviewer Entry

Reviewer 1

Comments:
Accept

Reviewer 2

Comments:
A clear and appropriately detailed response.

16. Security

R16. The technical infrastructure of the repository provides for protection of the facility and its data, products, services, and users.

Compliance Level:

4 – The guideline has been fully implemented in the repository

Reviewer Entry

Reviewer 1

Comments:

4 – The guideline has been fully implemented in the repository

Reviewer 2

Comments:

4 – The guideline has been fully implemented in the repository

Response:

Our AODN DevOps team manages the security of our systems and data. They undertake half-yearly security audits based on a checklist which is kept up to date with new systems and processes. AWS Trusted Advisor is also used for automated auditing of AWS services. Changes to system access configuration require peer-review from a second DevOps team member before being implemented. Changes to code require peer-review from a second staff member before being merged into main branches.

Users are assigned roles representing their positions within the organisation and receive a level of access based on the needs of these organisational roles, with DevOps team members being given the highest level of access. In assigning access, we follow a policy of “least privilege”, where a user is given the minimum level of access needed to perform his/her job functions. When an employee leaves the organisation, the DevOps team follows a carefully documented offboarding process to ensure that all access is revoked.

Authentication and authorisation for AWS is managed via the AWS Identity Access Management service. Two-factor authentication is required for administrative access to production services on AWS. Authentication and authorisation for the Nectar Research Cloud uses the Australian Access Federation (AAF). Employees log in via their University of Tasmania accounts, which also requires two-factor authentication. We also use GitHub single sign-on for authentication and authorisation management.

Firewalls are used to restrict resource access to those who are known to need access. When possible, we try to remove resources from public internet access to limit exposure risks. Physical production infrastructure is maintained by AWS under the shared responsibility model and some non-production infrastructure is maintained by the Nectar Research Cloud. None of our physical infrastructure is maintained by us or our host institution.

Logs from our production systems are sent to a central location and monitored for unusual access patterns. AWS costs are monitored using the AWS Cost Explorer and billing pages to ensure that no unusual activity is taking place. The DevOps team monitors the state of production systems and investigates any unexpected states. During our half-yearly security audit, we search for signs of security incidents we might have failed to detect if they occurred. We adjust our processes accordingly if such signs are discovered.

Any response to a security incident would be undertaken by the DevOps team. We have tested and documented our backup restoration process so that we can be confident in our ability to recover any data lost or corrupted in a security incident. Since our systems are defined using configuration-as-code and infrastructure-as-code we are also confident in our ability to restore our systems to their correct state after a security incident. None of our systems are essential so we could accept significant downtime if it were required to give us time to respond to an incident.

Reviewer Entry

Reviewer 1

Comments:
Accept.

Reviewer 2

Comments:
I concur that the guideline has been fully implemented. The use of two factor authentication is commendable. Actually testing the backup strategy is also commendable.

APPLICANT FEEDBACK

Comments/feedback

These Requirements are not seen as final, and we value your input to improve the CoreTrustSeal certification procedure. Any comments on the quality of the Requirements, their relevance to your organization, or any other contribution, will be considered as part of future iterations.

Response:

Reviewer Entry

Reviewer 1

Comments:

Overall, this was a comprehensive application with well-described processes and procedures. Areas such as dynamic dataset identifiers, a continuity plan and workflow updates are anticipated to show progress in the future certification renewal.

Reviewer 2

Comments: