IMOS Integrated Marine Observing System

Australian Ocean Data Network (AODN)

Ships of Opportunity – Australian Plankton Survey WORKFLOW

Version 5.0

26th June 2024

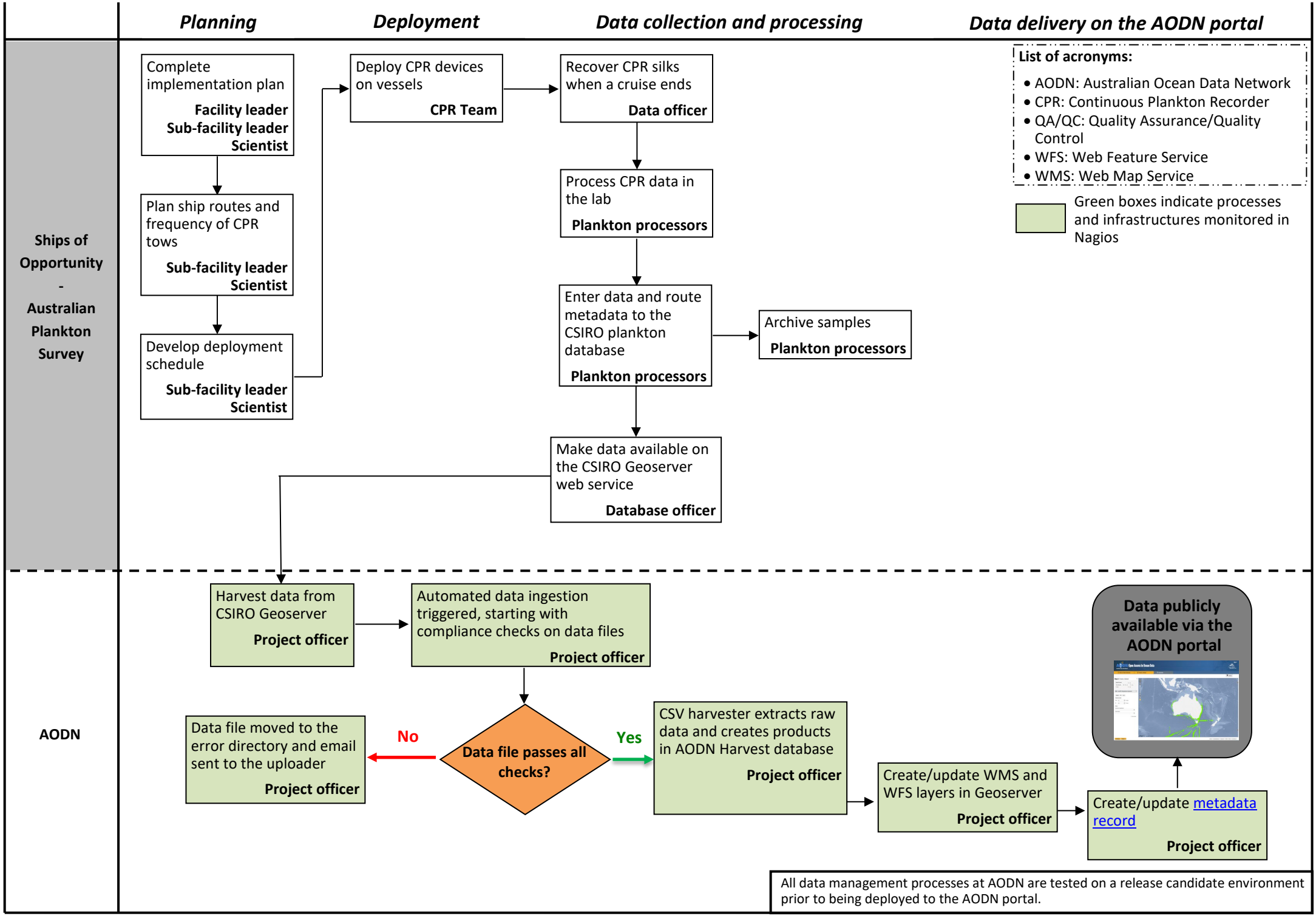
Data Workflows

The AODN, in managing the data for IMOS, has developed workflows for each IMOS facility to describe the flow of data from planning through data collection to data delivery and public data access.

The primary goals of these workflows are to:

- Improve data flow and data handoff, 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 and the AODN.
- 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 workflow is available on the next page of this document. Additional information (*i.e.* timeline, input, output, step description) for each operation step is available in the 'Supporting Information' section. The role and contact details of people involved in the workflow are summarised in a table and suggested potential improvements are listed at the end of the document.


Planning
Deployment
Data collection and processing
Data delivery on the AODN portal
Ships of Opportunity - Australian Plankton Survey
Complete implementation plan
Facility leader
Sub-facility leader
Scientist
Plan ship routes and frequency of CPR tows
Sub-facility leader
Scientist
Develop deployment schedule
Sub-facility leader
Scientist
Deploy CPR devices on vessels
CPR Team
Recover CPR silks when a cruise ends
Data officer
Process CPR data in the lab
Plankton processors
Enter data and route metadata to the CSIRO plankton database
Plankton processors
Archive samples
Plankton processors
Make data available on the CSIRO Geoserver web service
Database officer
List of acronyms:
• AODN: Australian Ocean Data Network
• CPR: Continuous Plankton Recorder
• QA/QC: Quality Assurance/Quality Control
• WFS: Web Feature Service
• WMS: Web Map Service
Green boxes indicate processes and infrastructures monitored in Nagios
AODN
Harvest data from CSIRO Geoserver
Project officer
Automated data ingestion triggered, starting with compliance checks on data files
Project officer
Data file passes all checks?
No
Data file moved to the error directory and email sent to the uploader
Project officer
Yes
CSV harvester extracts raw data and creates products in AODN Harvest database
Project officer
Create/update WMS and WFS layers in Geoserver
Project officer
Create/update metadata record
Project officer
Data publicly available via the AODN portal
All data management processes at AODN are tested on a release candidate environment prior to being deployed to the AODN portal.

Supporting information

Phase	Operation step	Timeline	Input	Output	Step description	Step operator
Planning	**Complete implementation plan**	**6 monthly milestone report**	IMOS funding	Implementation plan	• Annual plan for Australian Continuous Plankton Recorder deployments. • Evaluate required resources.	**Ships of Opportunity Facility leader, Sub-facility leader and Scientist**
	Plan ship routes and frequency of CPR tows	**On-going**	Implementation plan		• Identify scientifically valuable routes. • Contact shipping companies and research providers (AIMS, AAD, NMF) to find out about ship availability.	**Ships of Opportunity Sub-facility leader and Scientist**
	Develop deployment schedule	**On-going**	• Implementation plan • AAD shipping schedule • Merchant shipping schedule	Deployment schedule and plan	Agree on deployment schedule with shipping providers (dependent on when ships are running, especially research vessels).	**Ships of Opportunity Sub-facility leader and Scientist**
Deployment	**Deploy CPR devices on vessels**	**One-two weeks lead in**	• Deployment plan • CPR devices	CPR deployed	• Get machine and prepare cassettes. • Meet requirements from shipping companies (*e.g.* training when there are new crew). • Arrange transportation to ship departure point.	**Ships of Opportunity CPR Team**
Data collection and processing	**Recover CPR silks when a cruise ends**	**At the end of each voyage**	Deployed CPR	Silks containing plankton samples	Arrange transportation of equipment from ship arrival point back to lab.	**Ships of Opportunity Plankton processors**
	Process CPR data in the lab	**One week up to 12 months, depending on priority of samples and type, most data within 6 months**	Silks containing plankton samples and underway data from ships log	• Route information (waypoints and cutting information for each silk segment) • PCI values • Phytoplankton and zooplankton counts for each sample • Biomass for each sample	• Measure and mark silks. • Do PCI and cut silk at the same time (colour varies, and can see graduation as move through silk). • Process each segment individually for phytoplankton count (on silk). • Wash zooplankton off silk and count zooplankton Filter phytoplankton and zooplankton off silk and do dry weight biomasses.	**Ships of Opportunity Plankton processors**

Data collection and processing	**Enter data and route metadata to the CSIRO plankton database**	**One-two days**	• PCI values • Phytoplankton and zooplankton counts for each sample • Biomass data for each sample • Voyage and CPR deployment/recovery information	Database tables containing route information, sample information and scientific variables	• Enter route information from logsheets.or ships track • PCI, phytoplankton and zooplankton and biomass data are all entered separately, as the samples are processed.	**Ships of Opportunity Plankton processors**
	Archive samples	**Within a week of silk being fully processed**	Silk segments not processed	Archiving of segments for future collaborations	Samples preserved in a fixative, each segment stored in a separate bag.	**Ships of Opportunity Plankton processors**
	Make data available on the CSIRO Geoserver web service	**Immediate upon data entry**	Database tables containing route information, sample information and scientific variables	CSIRO Geoserver web service available for harvest by AODN	Create/update Geoserver layers (and underlying database views) so that AODN can periodically harvest all the raw data without direct access to the CSIRO database.	**Ships of Opportunity Database officer**
	Harvest data from CSIRO Geoserver	**Automatic (when files uploaded)**	Raw data layers in CSIRO Geoserver	Harvested data in CSV files in AODN incoming directory	A weekly scheduled process, downloads raw data from CSIRO Geoserver, into CSV files, and pushes them into the AODN ingestion pipeline.	**AODN Project officer**
	Automated data ingestion triggered, starting with compliance checks on data files		Harvested data in CSV files in AODN incoming directory	Valid CSV files ready to be published, non-compliant files moved to error directory	Handling of incoming files is automatically triggered when a file arrives in the AODN incoming directory. Checks are applied to verify that each file • is a correct data product for that upload location; or • Is valid CSV data file according to schemas agreed with CSIRO Any file that fails any of these checks is moved to an error directory and the provider is notified by email. The steps below are only performed for files that pass all checks.	**AODN Project officer**

	CSV harvester extracts raw data and creates products in AODN Harvest database		Valid raw data in CSV files	Raw data and data products in AODN database	Harvester loads all raw data into AODN Harvest database and creates database views for the derived products to be served via AODN Geoserver.	**AODN Project officer**
Data delivery on the AODN portal	**Create/update WMS and WFS layers in Geoserver**	**Automatic update**	Populated AODN Harvest database tables	WMS and WFS layers created in Geoserver	• Database tables and/or views are used to create a WMS and WFS layer in Geoserver. • Configure the pop-up window (content.ftl) and filters. • Create style for WMS visualisation.	**AODN Project officer**
	Create/update metadata record	**Automatic update**	Populated AODN Harvest database tables	Geonetwork record created and configured to support data discovery, visualisation and download via the AODN portal	• Create a metadata record with a new UUID. • Configure the newly created record (*e.g.* abstract, point of contact, parameters, timeframe). • Fill out the distribution section with links to the corresponding Geoserver WMS and WFS layers and other AODN's download services. • Talend harvester automatically updates bounding box.	**AODN Project officer**

Contact details (as at June 2024)

	Role	Name	Email address	Phone number
Ships of Opportunity - Australian Plankton Survey	Facility leader			
	Sub-facility leader	Anthony Richardson	a.richardson1@uq.edu.au	(07) 3833 5958
	Scientist			
	IMOS Liaison and plankton processor	Claire Davies	Claire.Davies@csiro.au	(03) 6232 5273
	Plankton processors	Julian Eribe Palomino	Julian.Uribepalomino@csiro.au	
		Felicity McEnnulty	Felicity.Mcennulty@csiro.au	
		Ruth Eriksen	Ruth.Eriksen@csiro.au	
		Frank Coman	Frank.Coman@csiro.au	(07) 3826 7103
		Anita Slotwinski	Anita.Slotwinski@csiro.au	(07) 3833 5964
		Mark Tonks	mark.tonks@csiro.au	(07) 3833 5973
		Sahan Jayasinghe	sahan.jayasinghe@csiro.au	
	Database engineer	Steven Edgar	Steven.Edgar@csiro.au	(07) 38335932
	Data officer	Margaret Miller	Margaret.Miller@csiro.au	(07) 3833 5944
AODN	Project officer	Leonardo Laiolo	Leonardo.Laiolo@utas.edu.au	
	Data services team leader	Benedicte Pasquer	Benedicte.Pasquer@utas.edu.au	(03) 6226 1927

Suggested improvements

- Give IMOS facilities read access to Nagios, the AODN infrastructure monitoring web application (**AODN suggestion**).
- Improve data download from web services provided by CSIRO. Currently, every time the Talend harvester run, the content of the database at IMOS is deleted and all data is downloaded again from web services available at CSIRO. Incremental update to only download new or modified data would be preferable.

Supporting links

AODN portal: http://portal.aodn.org.au

AODN portal help page: https://help.aodn.org.au/
IMOS - Phytoplankton Abundance and Biovolume (CPR) layer on AODN portal: https://portal.aodn.org.au/search?uuid=8e594cd1-363a-4b58-b741-9fa3624adb77

IMOS - Zooplankton Abundance and Biomass Index (CPR) layer on AODN portal: https://portal.aodn.org.au/search?uuid=bf287dfe-9ce4-4969-9c59-51c39ea4d011

IMOS - Phytoplankton Abundance and Biovolume (CPR) metadata record: https://catalogue-imos.aodn.org.au/geonetwork/srv/api/records/8e594cd1-363a-4b58-b741-9fa3624adb77

IMOS - Zooplankton Abundance and Biomass Index (CPR) metadata record: https://catalogue-imos.aodn.org.au/geonetwork/srv/api/records/bf287dfe-9ce4-4969-9c59-51c39ea4d011

IMOS user code library: https://github.com/aodn/imos-user-code-library
IMOS website: http://www.imos.org.au/
Ships of Opportunity Australian Plankton Survey sub-facility: https://imos.org.au/facilities/shipsofopportunity/auscontinuousplanktonrecorder