

## Automated Data Ingestion for the Australian Ocean Data Network

Marton Hidas, Roger Proctor, Sebastien Mancini, Leigh Gordon, Angus Scheibner, Laurent Besnard, Peter Blain University of Tasmania, Integrated Marine Observing System, Hobart, Australia

**NCRIS** 

Infrastructure for Australia

An Australian Government Initiative

**ASSOCIATE PARTICIPANTS** 

Curtin University

National Research

Integrated Marine **Observing System** 

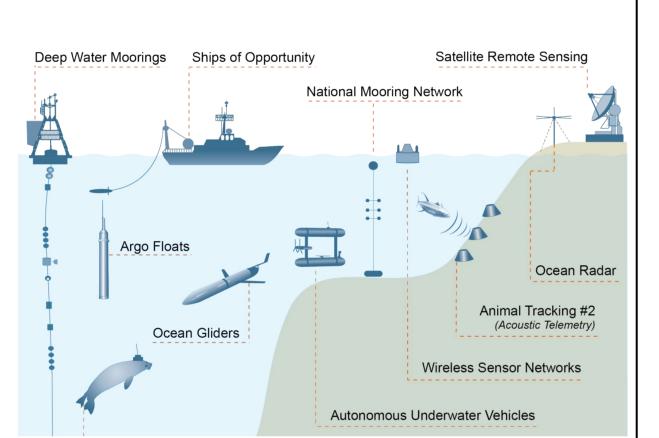
1 - Australian Ocean Data Network

The Australian Ocean Data Network (AODN) is an interoperable online network of marine and climate data resources. It is a collaboration between six major Australian Commonwealth Agencies, the Integrated Marine Observing System (IMOS), and a growing list of universities, state government offices and other organisations in Australia, New Zealand and the Pacific.

All data products contributed to the AODN are discoverable, and freely and openly available to the public via the AODN Portal (portal.aodn.org.au).

## 2 - Integrated Marine Observing System

Australia's Integrated Marine Observing System (IMOS, imos.org.au) is establishing a sustained observing program for Australian oceanic waters and shelf seas, covering physical, biological, and chemical variables.

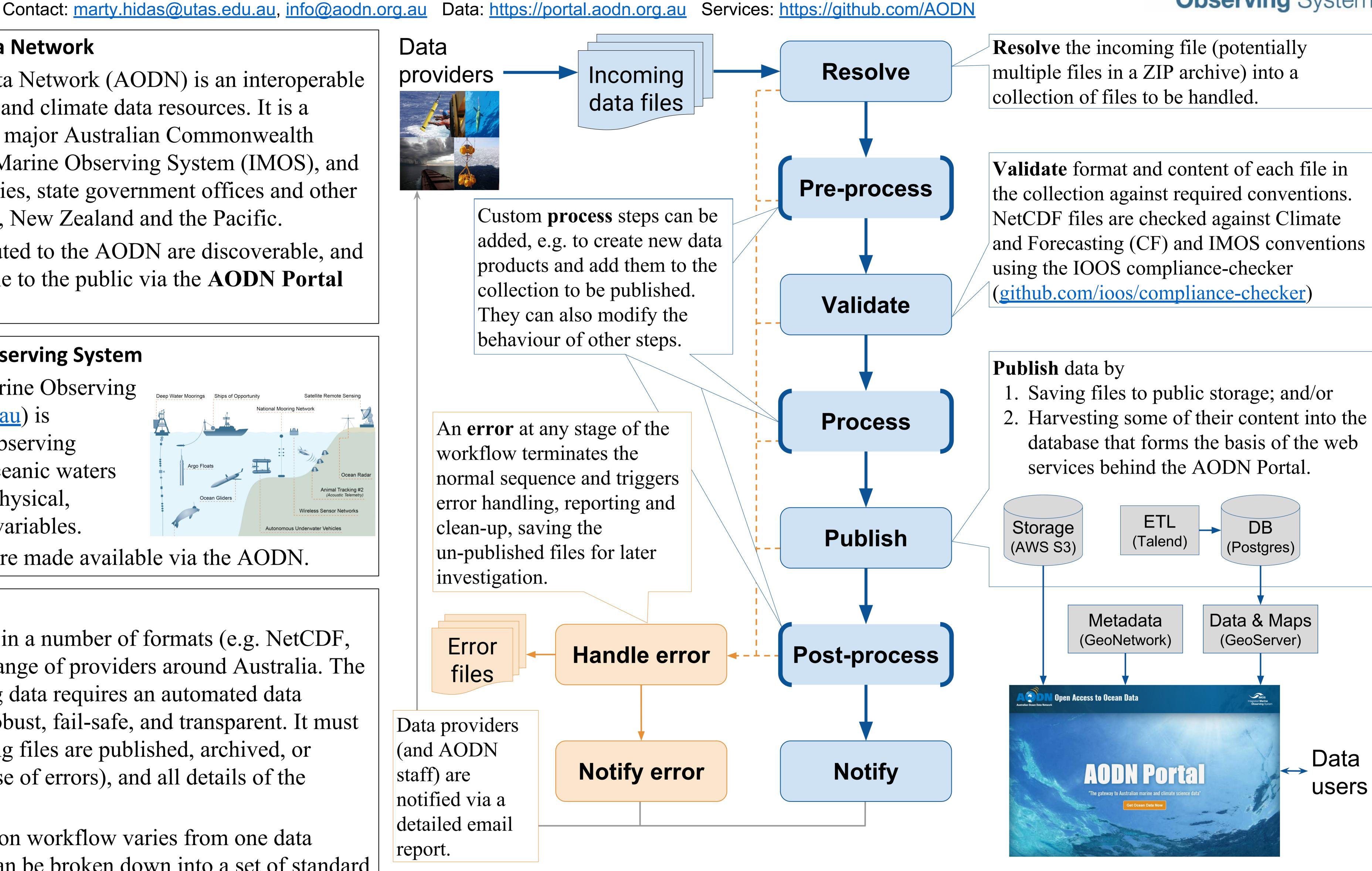


All IMOS data products are made available via the AODN.

## 3 - Data Ingestion

The AODN receives data in a number of formats (e.g. NetCDF, CSV, shapefile), from a range of providers around Australia. The sheer volume of incoming data requires an automated data ingestion system that is robust, fail-safe, and transparent. It must guarantee that all incoming files are published, archived, or retained for review (in case of errors), and all details of the process are logged.

While the detailed ingestion workflow varies from one data source to the next, each can be broken down into a set of standard tasks. At the core of our system is an open-source Python package (github.com/aodn/python-aodncore), which implements all these common tasks in a flexible way, so that individual workflows can be implemented with minimal effort. Another package contains code specific to individual data streams (github.com/aodn/python-aodndata).



PRINCIPAL PARTICIPANTS

PEOUTINE TRANS

IMOS is a national collaborative research infrastructure, supported by Australian

Government. It is operated by a consortium of institutions as an unincorporated

joint venture, with the University of Tasmania as Lead Agent. www.imos.org.au

UNIVERSITY of TASMANIA