

CHALLENGES IN AUTONOMOUS PLATFORM DATA MANAGEMENT AT BODC



JAK JONES, ROBYN OWEN, JORDAN
ATHERTON, OLIVIA PAMPALONI

AUTONOMOUS PLATFORMS AT BODC

- BODC manage data from Argo floats, Gliders and Autosub Long Range (ALR) also known as Boaty McBoatface.
- BODC handle data in three modes:
 - Near-real time (NRT)
 - Recovery (glider and ALR)
 - Delayed mode
- The focus of this talk will be near-real time data management



AUTONOMOUS PLATFORMS CHALLENGES

**Collating all metadata
before a deployment**

**Ensuring we have systems
running 24/7 365 days a
year**

PI Liaison

**Aligning with community
standards**

**New advancements such as
new sensors**

Making data FAIR



METADATA SOLUTIONS

Controlled Vocabularies

Metadata Portal

Buy-in from manufacturers



CONTROLLED VOCABULARIES

- Argo have a comprehensive collection of vocabularies that are applied to Argo float data
- The OceanGliders community also have their own vocabulary (OG1)
- To avoid ambiguity we map the parameters to controlled vocabularies from the NERC Vocabulary Server (NVS)
- This allows both users and machines to understand the metadata and data



The screenshot shows the NERC Vocabulary Server (NVS) interface. At the top, there are logos for UKRI, NERC Environmental Data Service, and the National Oceanography Centre. The main heading is "The NERC Vocabulary Server (NVS)" with a "Service Status" button. A navigation bar includes links for NVS Home, Vocabularies, Thesauri, Search NVS, SPARQL, Other Tools, and About NVS. The main content area is titled "Vocabulary" and "OceanGliders Parameter Usage Vocabulary". It lists metadata such as URI, Description, Creator, Modified, Version Info, Identifier, Register Manager, and Register Owner. On the right, there are sections for "Alternate Formats" (RDF/XML, Turtle, JSON-LD) and "Alternate Profiles" (Alternate Profiles ?). Below this is a "Members" section with a table of parameters.

ID ↑	Preferred Label ↑	Definition ↑	Date ↑
DPHDOXY	Calibrated phase shift reported by oxygen sensor	D phase is calculated as a third degree polynomial of (the phase angle measured by blue excitation light minus the phase angle measured by red excitation light).	2021-06-25
CHLA	Chlorophyll-A	In-situ fluorometer with either manufacturer, laboratory or sample calibration applied.	2018-06-01
FLUOCHLA	Chlorophyll-A signal from fluorescence sensor	Chlorophyll-A signal from fluorescence sensor.	2018-06-01
CDOM	Concentration of coloured dissolved organic matter in sea water	The quantity of coloured organic compounds in solution in unit volume of any body of water	2018-06-01

<https://vocab.nerc.ac.uk/collection/OG1/current/>

METADATA PORTAL – GLIDERS AND ALR

- Receiving metadata prior to a deployment is essential for BODC processing.
- Previous deployment setup involved manually editing text using the Oracle SQL interface (not user friendly)
- This is now done using a custom-built web-based UI
- Features a series of dropdowns constrained to controlled vocabulary lists hosted on the NVS (<https://vocab.nerc.ac.uk/>), in addition to free text entry fields to drive metadata setup, and ultimately database population
- This implementation resulted in a sixfold reduction in the time it takes to set up a deployment



Contextual metadata provided when a selection is made

New platform or sensor models or instances can easily be created in the same place, and are immediately available – propagating through the workflow

The screenshot shows a web form for 'Deployment 1' with the following sections:

- Platform:** A search dropdown with 'unit_436: Stella' selected. Below it, contextual metadata is displayed: 'Model name: Teledyne Webb Research G2', 'Owner: National Oceanography Centre (NOC), Southampton', 'WMO code: 6800993', and 'ICES code: 74XK'. A '+ Create Platform' button is on the right.
- Logger:** A search dropdown with 'unit_436: Teledyne Webb Research Slocum G1+G2 Glider Navigation data logger' selected. A 'Remove' button is on the right. A '+ Create Logger' button is on the right.
- Sensors:** A search dropdown with '285: Aanderaa Oxygen optode 4831' selected. A 'Remove' button is on the right. A '+ Create Sensor' button is on the right.
- Parameters:** A list of parameters: 'OXYSDP1', 'OXYCPHAR', 'OXYCOPVR', 'OXYCOPVE', 'DOXYAAOE'. Below this is another search dropdown with 'ctd' entered. A 'Remove' button is on the right.

At the bottom right, there is a 'Submit' button.

Search dropdowns to drive database population, lists are constrained against controlled vocabularies maintained by vocabulary experts

Variables (parameters) associated with the chosen sensor are presented

BUY-IN FROM MANUFACTURERS – ARGO

- We have worked with manufacturers in the Argo community to provide machine readable JSON objects containing sensor metadata
- BODC have created a metadata validator for these files and upcoming work will focus on integrating the JSON objects into workflows
- This means we receive metadata from the source and do not need to hassle PIs and engineers before a deployment
- This is allowing us to work towards automating end-to-end workflows



MONITORING SOLUTIONS – GLIDER AND ALR DASHBOARD

- Previous monitoring of glider and ALR data archival and processing was manual and involved checking multiple applications
- Monitoring can now take place at a glance in a single application
- Issues in the flow of data are visually highlighted
- Glider data managers can easily see the server and active deployment status
- A snapshot of data is plotted from BODC's ERDDAP instance (<https://linkedsystems.uk/erddap/index.html>), highlighting any problems in the data



Server status can be viewed at a glance

Links to glider operator portals in the header bar allow for a quick comparison against operator files

The screenshot shows a dashboard with a teal header bar. On the left is the logo for the National Oceanography Centre and British Oceanographic Data Centre. On the right is a navigation menu with links: Home, Glider inventory, GliderLab, SAMS Dashboard, UEA Dashboard, and Fleet Finder. Below the header is a sub-header with 'Server', 'Active Deployment Status', and 'Plots'. The main content area features a central white panel with a red border, titled 'Last updated: 20/11/2025, 15:01:45 (automatically refreshes every 30 seconds)'. This panel contains four server status cards: 'livbodcrt', 'livbodcords', 'livbodcwebin', and 'api.linked-systems.uk'. Each card lists metrics like CPU usage, Disk usage, Load, and SLOCUM status, all marked as 'OK' with green indicators. A copyright notice at the bottom reads 'Copyright © National Oceanography Centre 2025.'



Timestamps demonstrating the flow of data through the system at the various processing steps

The dashboard displays several sections related to deployment and data processing. A red box highlights the 'APDS V2 Status' and 'Data Delivery' sections, which are color-coded to show the status of each deployment. A red arrow points from the text above to the 'Data Delivery' section.

APDS V2 Status

Deployment	Serial No	Name	Source	RFI	RFI OC	OC	ConvJSON	Number of files ingested
708	unit_305	Dolomite	2025-11-20 14:34:48	2025-11-20 14:34:50	2025-11-20 14:34:51	2025-11-20 14:37:59	2025-11-20 14:34:52	490/492
694	unit_631	Unit_631	2025-11-07 19:14:42	2025-11-07 19:15:45	2025-11-07 19:17:43	2025-11-07 19:23:44	2025-11-07 19:18:48	567/573
695	Unit_632	Unit_632	2025-10-23 19:47:55	2025-10-23 19:47:58	2025-10-23 19:48:04	2025-10-23 19:52:10	2025-10-23 19:48:06	357/362
700	sg638	Ziggy	2025-11-20 13:32:41	2025-11-20 13:32:43	2025-09-29 12:59:59	2025-11-20 13:38:59	2025-11-20 13:32:44	720/750

Data Delivery

Deployment	Serial no	Name	Met Office last submitted file	Time submitted (ConvJSON)	Raw zip ERDDAP timestamp	OC file ERDDAP timestamp	OC ingested ERDDAP timestamp
694	unit_631	Unit_631	unit_631-2025-310-0-0[slocum_g1-g2_glider_navigation_data_logger].json	2025-11-07 19:20:23	2025-11-07 19:18:18	2025-11-07 19:23:45	2025-11-20 10:49:24
695	Unit_632	Unit_632	unit_632-2025-294-0-4[oxygen_optode_4831].json	2025-10-23 19:48:07	2025-10-23 19:48:04	2025-10-23 19:52:11	2025-11-20 10:49:24
700	sg638	Ziggy	p6380771[umpumped_ct_sal_ctd].json	2025-11-20 13:32:45	2025-11-20 13:34:08	2025-11-20 13:38:00	2025-11-20 14:00:08
708	unit_305	Dolomite	unit_305-2025-315-0-150[slocum_g1-g2_glider_navigation_data_logger].json	2025-11-20 14:34:53	2025-11-19 08:53:29	2025-11-20 14:37:59	2025-11-20 14:44:10

V2 Slocum Status

Deployment ID	Serial no	Name	ASCII File	Cache Pending
708	unit_305	Dolomite	2025-11-20 14:34:49	OK
694	unit_631	Unit_631	2025-11-07 19:15:44	OK
695	Unit_632	Unit_632	2025-10-23 19:47:57	OK

GTS Status

Accession/Deployment ID	Serial no	Name	GTS Last Submitted File	GTS Time Submitted
SOC250954/ 694	unit_631	Unit_631	unit_631-2025-310-0-0.sbd.asc	2025-11-07 19:21:09
SOC250980/ 700	sg638	Ziggy	p6380771.nc	2025-11-20 14:15:16
SOC250950/ 695	Unit_632	Unit_632	unit_632-2025-294-0-4.sbd.asc	2025-10-23 21:15:19
SOC251008/ 708	unit_305	Dolomite	unit_305-2025-315-0-150.sbd.asc	2025-11-20 14:45:13

Metadata API Checks

Click the status code to view response JSON

Deployment	EGDEEP
694	200
695	200
708	200
700	200

Copyright © National Oceanography Centre 2025.

Colour coded to give a quick visual indication of status



Select an individual deployment for further inspection on the Plots tab

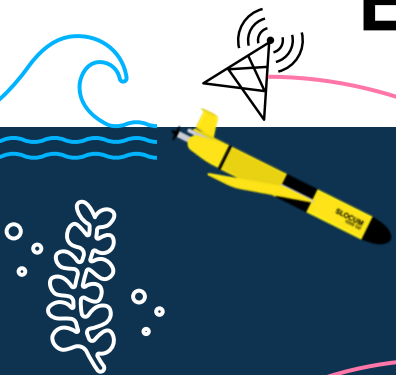
Plots are interactive, enabling greater inspection



Plots allowing inspection of data coming in



EXAMPLE: GLIDER NRT DATA PIPELINE



Data

Shore based server

Archive API

Recovery Delivery

QC, Modelling & DTO

Metadata requirements

File pusher app

Ingestion services

Decision-making

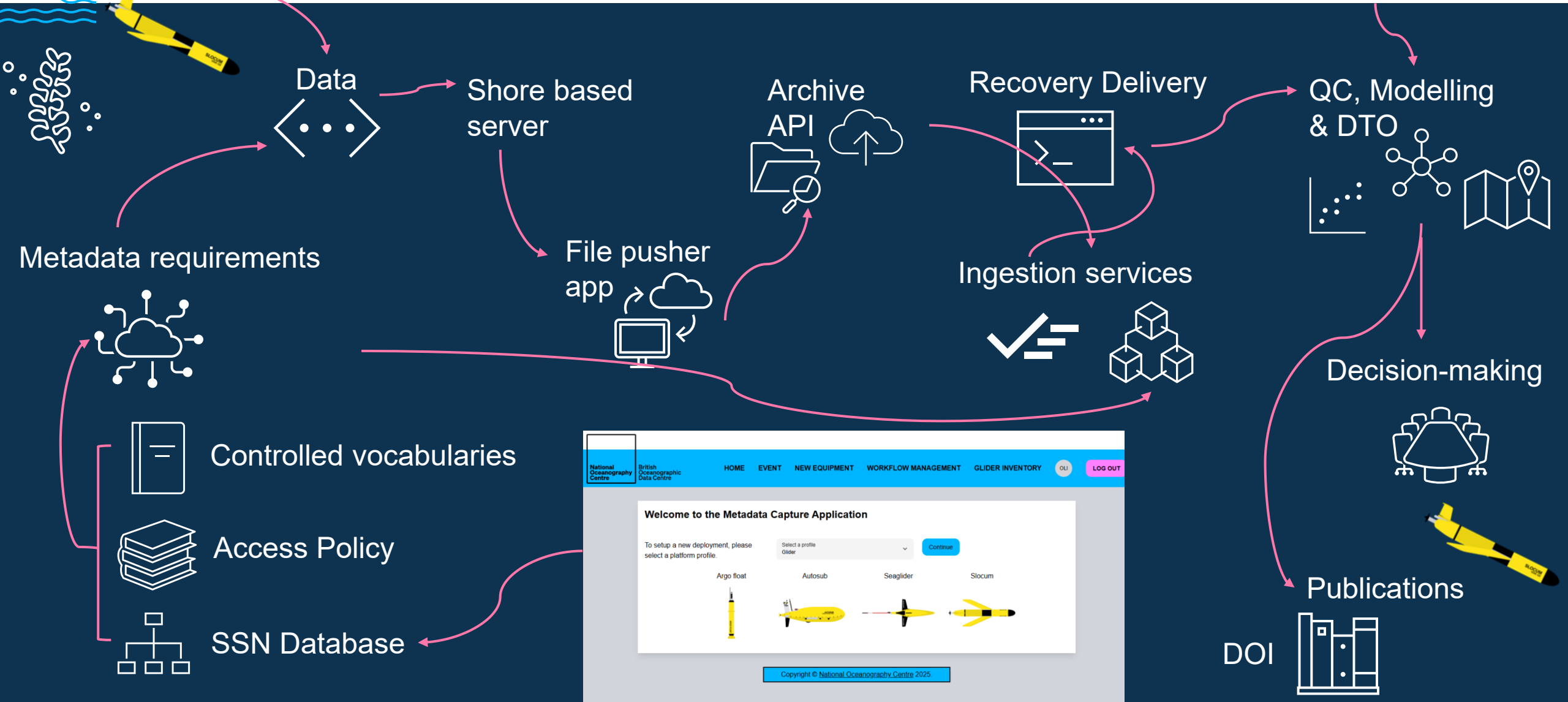
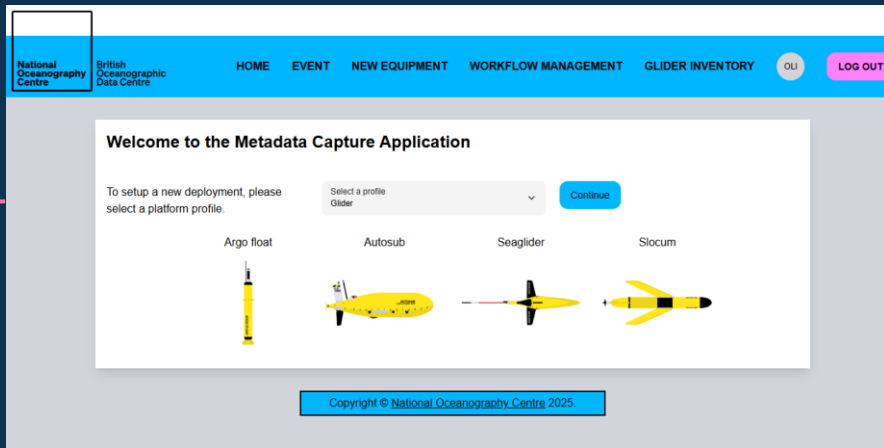
Controlled vocabularies

Access Policy

SSN Database

Publications

DOI



THANK YOU

GLIDERSBODC@BODC.AC.UK
ARGO@BODC.AC.UK



British
Oceanographic
Data Centre