Continuous ocean monitoring from sensor arrays on the UK large research vessels

Louise Darroch¹, Juan Ward², Alexander Tate³, and Justin Buck¹
¹British Oceanographic Data Centre, National Oceanography Centre, Liverpool, United Kingdom (louise.darroch@bodc.ac.uk)
²National Marine Facilities, National Oceanography Centre, Southampton, United Kingdom
³UK Polar Data Centre, British Antarctic Survey, Cambridge, United Kingdom

More than 40% of the human population live within 100 km of the sea. Many of these communities intimately rely on the oceans for their food, climate and economy. However, the oceans are increasingly being adversely affected by human-driven activities such as climate change and pollution. Many targeted, marine monitoring programmes (e.g. GOSHIP, OceanSITES) and pioneering observing technologies (e.g. autonomous underwater vehicles, Argo floats) are being used to assess the impact humans are having on our oceans. Such activities and platforms are deployed, calibrated and serviced by state-of-the-art research ships, multimillion-pound floating laboratories which operate diverse arrays of high-powered, high-resolution sensors around-the-clock (e.g. sea-floor depth, weather, ocean current velocity and hydrography etc.). These sensors, coupled with event and environmental metadata provided by the ships logs and crew, are essential for understanding the wider context of the science they support, as well as directly contributing to crucial scientific understanding of the marine environment and key strategic policies (e.g. United Nation's Sustainable Development Goal 14). However, despite their high scientific value and cost, these data streams are not routinely brought together from UK large research vessels in coordinated, reliable and accessible ways that are fundamental to ensuring user trust in the data and any products generated from the data.

The National Oceanography Centre (NOC) and British Antarctic Survey (BAS) are currently working together to improve the integrity of the data management workflow from sensor systems to end-users across the UK National Environment Research Council (NERC) large research vessel fleet, making cost effective use of vessel time while improving the FAIRness of data from these sensor arrays. The solution is based upon an Application Programming Interface (API) framework with endpoints tailored towards different end-users such as scientists on-board the vessels as well as the public on land. Key features include: Sensor triage using real-time automated monitoring systems, assuring sensors are working correctly and only the best data are output; Standardised digital event logging systems allowing data quality issues to be identified and resolved quickly; Novel open-source, data transport formats that are embedded with well-
structured metadata, common standards and provenance information (such as controlled vocabularies and persistent identifiers), reducing ambiguity and enhancing interoperability across platforms; An open-source data processing application that applies quality control to international standards (SAMOS, or IOOS Qartod); Digital notebooks that manage and capture processing applied to data putting data into context; Democratization and brokering of data through open data APIs (e.g. ERDDAP, Sensor Web Enablement), allowing end-users to discover and access data, layer their own tools or generate products to meet their own needs; Unambiguous provenance that is maintained throughout the data management workflow using instrument persistent identifiers, part of the latest recommendations by the Research Data Alliance (RDA).

Access to universally interoperable oceanic data, with known quality and provenance, will empower a broad range of stakeholder communities, creating opportunities for innovation and impact through data use, re-use and exploitation.