An in-situ observation data harmonization and integration that benefits to the wider Atlantic Ocean community

Valérie Harscoat (Ifremer, France)
Christoph Waldmann (MARUM, Germany)
Sylvie Pouliquen (Euro-Argo ERIC, Ifremer, France)

In-situ ocean observation data can be difficult to access and use for scientists and end users who have not been involved in their acquisition. Moreover, there are no 'on the shelf' standards easily implementable by anybody to make the data easily usable, interoperable and re-usable through common features. The risk of 'mixing apples with oranges' can highly impact the quality of products that integrate in-situ observation data and derive information for end users or for research activities, especially in an operational context. Furthermore, data that is not managed properly and archived in long-term repositories will be lost for the community after 10 years.

The first concern for end users is where to find or how to access the available in-situ data acquired in the Atlantic Ocean. They want to know which systems, networks or infrastructures in place at European or international level, are collecting and distributing such data. Then they want to access their data services (viewing, downloading and monitoring). Within the AtlantOS project, a catalogue tool was set up to document and facilitate the access to the existing systems. From this central web interface, the user can discover the existing systems and readily access their data services and products related to the Atlantic Ocean. Sustainability of this tool is ensured by its connection to the Global Earth Observation System of Systems (GEOSS).

Furthermore, for end users or infrastructures ingesting data from in-situ networks:

- It is much easier and reliable to go to a single entry point to access the data, rather to connect to multiple individual scientist servers. The access to in-situ network data is enhanced by providing a unique entry point to discover and download existing data and products. Within AtlantOS, the main advances are more data in the existing global data centres (EGO for gliders, OceanSITES for fixed point platforms and transport arrays, ICOS-Ocean for some VOS and GO-SHIP carbon data), improved access to ADCP data for GO-SHIP and a new Global Data Assembly Centre (GDAC) for drifters (endorsed by DBCP/JCOMM) for data access to NRT drifter data plus best copy selection of Delayed Mode products.
- Before using such in-situ data, it must be ensured that the same data type from different in-situ networks is interoperable to avoid 'mixing apples and oranges'. Therefore, the AtlantOS community agreed to adopt and implement a minimum set of mandatory and common features relying on existing international standards and protocols. That includes metadata for platforms and data providers, vocabularies for metadata and data, Near Real Time QC procedures for a core of 7 Essential Variables (Temperature, Salinity, Current, Sea Level, Oxygen, Chlorophyll, Nitrate and Carbon).

Moreover, infrastructure managers need to enhance their tools for data ingestion and need to integrate better quality data from in-situ networks. Within AtlantOS, each infrastructure chose an adapted strategy: setting up new nodes (ICOS-Ocean, physical data from CPR) or direct GDAC data flow (on-going for Argo and to follow for Gliders, Drifters and OceanSITES) to SeaDataNet, new marine biological data flow to related infrastructures (Fish Acoustics to ICES, ETN to EMODnet-Biology) and more data in Copernicus Marine In Situ Thematic Centre through GDACs harvesting.

Finally, data providers would like to keep track of the data usage for the platforms they operate, even when shared with other centres. This is the aim of the traceability of use service set up in the framework of the AtlantOS project. This integrated service relies on minimal and common tracking information shared by the systems and computes data usage statistics in a central web dashboard.

In the last four years, all the systems involved in AtlantOS have entered an improvement loop to ensure that in-situ data from different and diverse observing networks operating in the Atlantic Ocean are readily accessible and useable to the wider community, including the international ocean science community and other stakeholders. Relying on existing infrastructures that will last after the end of the project lifetime, AtlantOS has moved forwards along the implementation of the FAIR principles for Atlantic observations (Findable, Accessible, Interoperable, and Re-Usable).