



EuroGOOS DataMEQ Working Group Meeting

13 November 2024, Bologna, Italy (Hybrid)

Agenda

1. Introduction (T. Carval)
2. DataMEQ highlights and work plan review (D. Eparkhina)
3. Data traceability - DOI strategies and data aggregators integration (F. Merceur, A. Novellino, T. Carval, A. Matthews).
4. Cloud-native workflows for DAC/GDAC architecture - FAIR-EASE, Blue-Cloud Workbench, and other collaborative projects (T. Carval, V. Racapé, S. Simoncelli, L. Drouineau)
5. Artificial intelligence - Applications in anomaly detection, species identification, and sound signature cataloging (T. Carval, A. Novellino, V. Tosello)
6. Summary of decisions and actions (T. Carval)

List of participants

1. Thierry Carval, IFREMER, France (DataMEQ WG chair)
2. Dina Eparkhina, EuroGOOS (DataMEQ WG facilitator)
3. Andrea Casaucao, SOCIB, Spain
4. Andy Matthews, National Oceanography Centre, UK
5. Antonio Novellino, ETT, EMODnet, Italy
6. Callum Rollo, VOTO, Sweden
7. Carles Castro Muniain, FVON, ODN, USA
8. Christoph Waldman, IQuOD
9. Claudia Fratianni, INGV (will replace Simona in DataMEQ)
10. Dominique Obaton, IFREMER, France
11. Emma Reyes, SOCIB, Spain
12. Fred Merceur, IFREMER, France
13. Guilherme Castelhão, IQuOD, USA
14. Juan Gabriel Fernández, SOCIB, Spain
15. Jukka Seppala, Syke, Finland
16. Kate Collingridge, Cefas, UK
17. Ludovic Drouineau, IFREMER, France
18. Marie Jossé, CNRS, France
19. Marta de Alfonso, Puertos del Estado, Spain
20. Martin Kramp, OceanOPS/WMO
21. Pier Luigi Buttigieg, AWI/ODIS, Germany
22. Sebastian Ehrhart, SYKE, Finland, EU

23. Simona Simoncelli, INGV, Italy
24. Susanne Tamm, BSH, Germany
25. Tjerk Krijger, MARIS
26. Vanessa Tosello, IFREMER, France
27. Veselka Marinova, IO-BAS, Bulgaria
28. Virginie Racape, PokaPok, France

Key discussions points

DataMEQ Highlights

- The EuroGOOS Data Management, Exchange, and Quality (DataMEQ) Working Group focuses on enhancing the quality and interoperability of ocean data systems.
- New EuroGOOS Data Policy 2023 developed as the European implementation of the new IOC Data Policy 2023 has already been formally signed by 32 EuroGOOS member organisations, with a target of 48.
- Metadata for Essential Ocean Variables (EOV) data must align with the IODE Ocean Data and Information System (ODIS) schema to ensure global visibility in the GOOS/IODE Data Space.
- [OceanGliders NetCDF format version 1.0](#) published, its implementation is scheduled for release in Q4 2024.
- SeaDataNet progresses on NetCDF vocabularies for Ship-ADCP (Acoustic Doppler Current Profiler) data.
- Results from the EU EuroSea and AMRIT projects to be integrated into the DataMEQ activities. Note the EuroSea Tide Gauge Task Team deliverable on New Tide Gauge Data Flow Strategy DOI 10.3289/eurosea_d3.3.

Data Traceability

- Persistent Identifiers (PIDs) are critical for datasets, institutions, and tracking the uptake. Promoting PIDs across networks like Copernicus, EMODnet, and World Ocean Database is critical. Example: Argo float data integrates DOI for dataset granularity and traceability. Argo notes considerable issues with citation – a study showed [Argo datasets appear in 530 articles published in 2023](#), but only 77 were cited correctly.
- Misaligned citations in publications reduce traceability and proper credit to contributors.
- Education and enforcement are necessary to standardize citation practices. Need for DataMEQ to develop authoritative guidelines.
[Post-meeting note: EuroGOOS may review the IOC Data Cookbook guidelines 2013 – section 3.8 <https://unesdoc.unesco.org/ark:/48223/pf0000373775>]
- Encourage the adoption of RDA Complex Citation Working Group recommendations, with DataMEQ contributing to this working group.
- Need for a broker vision to decouple data systems from data itself, ensuring long-term data heritage through DOIs.

Collaboration with Global Programs/Frameworks

- IODE ODIS (Unesco [Ocean Data Information System](#)): Linking DataMEQ systems to ODIS through STAC ([SpatioTemporal Asset Catalogues](#)).
Reference the ODIS Node setup guide (<https://book.odis.org/gettingStarted.html>) as a resource for facilitating this integration.

- WMO WIS2: Enhancing integration with global frameworks for metadata and data dissemination (World Meteorological Organization Information System) .
- GOOS BioEco EOVS portal will serve as a model for replicating FAIR data practices across other EOVS (Essential Ocean Variables).
- Address the importance of properly aggregating the European Research Infrastructure Consortia (ERICs) data through the data architecture.

Artificial Intelligence Applications

- Examples of AI being used for species identification from acoustics and imagery in projects like EOVS iMagine in Azores and EMODnet Sound Signature Catalog. These projects enhance FAIR image/sound annotation practices.
- Expanding beyond image analysis and species identification, we explored the application of [YOLO](#) to scientific graphs to detect duplicate drifting buoys in a dataset of 30 000 platforms.
- Challenges include standardizing annotations and integrating marine species with OBIS (Ocean Biodiversity Information System), as well as storage, chunking, and integrating operational workflows.

Cloud-Native Workflows

- Emphasis on transitioning DataAssembly Centre/Global Data Assembly Centre (DAC/GDAC) architecture to cloud platforms like the European Open Science Cloud (EOSC) and implementing reproducible scientific workflows such as Copernicus Marine Data Store (MDS), FAIR-EASE and BlueCloud workbenches or Galaxy Europe.
- The shift to a cloud-based approach will improve the efficiency of data management systems (ex: Argo) and align it with contemporary technological standards: scalable, reliable, robust.

Addressing Data Flow Fragmentation

- Concerns raised about fragmented portals. Recommendation to enhance data flow design towards the various portals.
- Discussion on harmonizing sea level data flow through initiatives like GLOSS. Need for consolidation of GLOSS data and improvement of traceability.

Actions

1. Collaboration with ODIS and WMO:
 - 1.1. Coordinate with IODE ODIS on outputs and shared methodologies via STAC catalogues. Develop a roadmap for linking European data systems to an IOC-integrated data space. Consider strategies to decouple data systems from data itself, using DOIs to secure the heritage and long-term usability of marine data.
 - 1.2. Promote interoperability and link DataMEQ systems with WMO WIS2.
 - 1.3. Conduct periodic FAIR assessments for datasets and workflows.
2. Enhance citation standards:
 - 2.1. Develop EuroGOOS guidelines and advocate for stricter enforcement of proper citation by publishers.
 - 2.2. Ensure DataMEQ contributions to the RDA Complex Citation Working Group.
 - 2.3. Review mechanisms to make Galaxy workflows citable for wider adoption of FAIR practices.

3. Explore image annotation standards:
 - 3.1. Continue development of image/sound signature catalogues and AI models, with a focus on processing mixed image/sound categories and large datasets (ex: [Ecotaxa](#)).
 - 3.2. Investigate how deep species annotations can integrate with the Ocean Biodiversity Information System (OBIS).
4. Identify topics for Standard Operating Procedures (SOPs) and Best Practices (BPs):
 - 4.1. Develop SOPs and BPs for the [Ocean Best Practices System](#), focusing on data publication and AI-driven analytics.