

Open Science Near Real Time Data: an example of the application of FAIRness in an oceanographic context

Alexia Cociancich, Sebastian Plehan, Elena Partescano, Alessandra Giorgetti
National Institute of Oceanography and Applied Geophysics - OGS

Abstract. . Open Science is central to European research policy, promoting collaboration, quality, and accessibility. The FAIR Data Principles are essential for effective data management. The National Institute of Oceanography and Applied Geophysics applies these principles to oceanographic data from the Adriatic Sea, addressing challenges in data acquisition and transmission. Since raw data lacks metadata, the National Oceanographic Data Centre reconstructs metadata to ensure usability. ERDDAP facilitates standardized data access, supporting interoperability across European research infrastructures like EMSO ERIC. A geoportal enhances accessibility with intuitive visualizations. Achieving full FAIRness remains a challenge, but ongoing efforts align with European Open Science goals.

Keywords. FAIR, Open Science, Oceanography, Near RealTime

Introduction

Open Science is at the heart of European research policy. It aims to facilitate the exchange, collaboration and improvement of the quality of research and to place science at the centre of human and social development (https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/open-science_en).

Open science policy and practice is changing the way research is conceived, managed, shared and evaluated.

Today, it is no longer enough to simply publish the results of publicly funded research without making the data as open and accessible as possible.

In this context, the FAIR Data Principles (Wilkinson et al. 2016) developed in 2016 recognize that good data management across the entire lifecycle is critical to the success of Open Science. Researchers who follow these principles produce data that are Findable, Accessible, Interoperable and Reusable (FAIR)

1. Open Science at OGS and the Challenge of Oceanographic Data

Open Science is one of the research and innovation missions that the National Institute of Oceanography and Applied Geophysics - OGS has identified in order to expand the user community of scientific data.

For more than 10 years, the OGS has been managing a network of oceanographic instruments in the Adriatic Sea that can transmit data in near real time (Partescano et al. 2017). The difficulties that the environment poses to data acquisition and transmission significantly increase the challenges of subsequent data management. The heterogeneity of the

parameters recorded and the instruments used requires a considerable effort in managing the data obtained.

Firstly, it should be noted that the transmitted raw data does not contain any metadata for reasons of efficiency. Although the latter must be associated after transmission to optimize throughput, without metadata to contextualise the data, it would be difficult to analyze, which is why the National Oceanographic Data Centre is committed to reconstructing the information content through the metadata process, because distributing data in near real time without respecting the basic principles of the FAIR approach would be of little use.

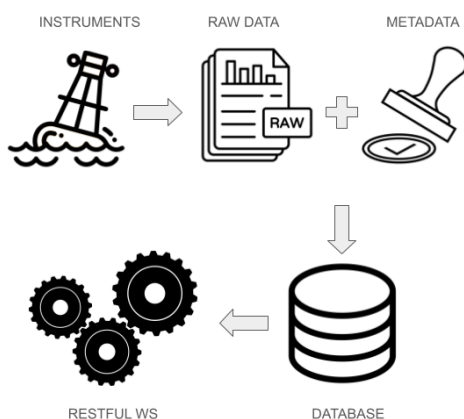
2. Ensuring FAIRness through Metadata and Data Services

To guarantee the FAIRness of the data, the correct assignment of metadata is crucial for its findability and the use of standard vocabularies is equally important to ensure a high level of interoperability. In this sense, the services provided by BODC (Thomas et al. 2015) are a fundamental support for the oceanographic research community.

To ensure data accessibility in a standardized and harmonized way, ERDDAP (<https://nodc.ogs.it/erddap>) was chosen as the preferred solution. As one of the most widely used open source software in the scientific field related to environmental data, it allows data managers to combine different sources and publish datasets in a powerful RESTful web service (Mendelssohn & Simons 2008) as well as a simple and intuitive web interface. This enables users to filter and export data in several formats, improving data accessibility and interoperability.

Fig. 1

Data flow
schema



3. Harmonizing Data and Enhancing Accessibility for All Users

A further strength is the possibility to use ERDDAP in a federated way. For example, this has enabled the EMSO ERIC infrastructure (<https://emso.eu>) to make the datasets of different research centres distributed across the European territory via a single endpoint. The harmonization example carried out within the EMSO ERIC community (Martínez et

al. 2024), of which the OGS is a member, addressed the challenge of overcoming the complexity and heterogeneity of different data and metadata formats. The use of semantic artefacts such as vocabularies and standards together with the use of ERDDAP and common metadata standards are elements that have allowed to achieve a high level of FAIRness. Finally, a geoportal (<https://nodc.ogs.it/geoporta>) has been developed to provide the most important information on the stations of the monitoring network in the Adriatic Sea in near real time to ensure an immediate and more intuitive graphical representation even for non-experts.

4. Conclusions

In conclusion, improving the FAIRness of data (Tanhua et al. 2019) is a crucial goal for all those involved in data management in order to meet the expectations of the European Commission. To obtain this, communities and data management infrastructures have faced numerous efforts and overcome challenges, but the path that leads to open science is not over yet, there are numerous initiatives to achieve a complete FAIRness of data, which is even more challenging in the context of the management of data acquired in near real-time.

5. Acknowledgements

The authors would like to thank the Technological development and support for acquisitions infrastructure (TEC) of the OGS and EMSO community as well as the financial contribution of ITINERIS project, the Italian Integrated Environmental Research Infrastructures System funded by the EU - Next Generation and PNRR Funds for the realisation of an integrated system of research and innovation infrastructures.

Bibliographic references

- Martínez, E., Libes, M., Fratianni, C., Partescano, E., Cociancich, A., Pensieri, S., Paladini de Mendoza, F., Rodero, I., (2024) EMSO Eric Metadata Harmonization Efforts, International Conference on Marine Data and Information Systems - Proceedings Volume, p 105-107 <https://dx.doi.org/10.13127/MISC/80/36>
- Mendelssohn, R. & Simons, R. (2008) ERDDAP - An Easier Way for Diverse Clients to Access Scientific Data From Diverse Sources, AGU Fall Meeting Abstracts. https://www.researchgate.net/publication/252989046_ERDDAP_-_An_Easier_Way_for_Diverse_Clients_to_Access_Scientific_Data_From_Diverse_Sources
- Partescano, E., Brosich, A., Lipizer, M. et al. (2017) From heterogeneous marine sensors to sensor web: (near) real-time open data access adopting OGC sensor web enablement standards, Open geospatial data, softw. stand. 2, 22 <https://doi.org/10.1186/s40965-017-0035-2>
- Tanhua, T.; Pouliquen, S.; Hausman, J.; O'Brien, K. M.; Bricher, P.; Bruin, T.; Buck, J. J.; Burger, E. F.; Carval, T.; Casey, K. S.; Diggs, S.; Giorgetti, A.; Graves, H.; Harscoat, V.; Kinkade, D.; Muelbert, J. H.; Novellino, A.; Pfeil, B. G.; Pulsifer, P.; Van de Putte, A. P.; Robinson, E.; Shaap, D.; Smirnov, A.; Smith, N.; Snowden, D. P.; Spears, T.; Stall, S.; Tacoma, M.;

Thijsse, P.; Tronstad, S.; Vandenberghe, T.; Wengren, M.; Wyborn, L.; Zhao, Z., (2019) Ocean FAIR Data Services, *Frontiers in Marine Science* <https://dx.doi.org/10.3389/fmars.2019.00440>

Thomas, R., Lowry, R. K., and Kokkinaki, A., (2015) Moving Controlled Vocabularies into the Semantic Web,

American Geophysical Union vol. 2015, Art. no. IN21D-1714, <https://ui.adsabs.harvard.edu/abs/2015AGUFMIN21D1714T/abstract>

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship, *Scientific Data*. <https://doi.org/10.1038/sdata.2016.18>

Authors

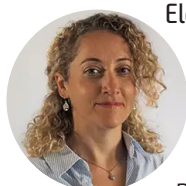
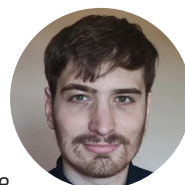


Alexia Cociancich acociancich@ogs.it

Alexia Cociancich has a master's degree in computer engineering and has worked as an IT consultant for 8 years in the field of industrial automation. Since 2019, she has been working at the National Institute of Oceanography and Experimental Geophysics (OGS) and is involved in database design and development for the National Oceanographic Data Centre (NODC).

Sebastian Plehan splehan@ogs.it

Sebastian Plehan is an experienced programmer with expertise in software development, database administration and IT support. He currently works at the OGS within the National Oceanographic Data Centre, where he contributes to the maintenance of services, and the development of new solutions for scientific data management projects. Passionate about new technologies, he is also studying at university to deepen his knowledge.



Elena Partescano epartescano@ogs.it

Elena Partescano is a data manager with expertise in managing, validating, analyzing, and organizing oceanographic data within relational databases. She supports real-time and delayed monitoring systems and works with SQL on Unix/Linux, metadata via XML and Mikado, XSLT, and GIS platforms. Elena also designs data entry interfaces, manages XML DB and XQuery, configures RESTful services, and ensures compliance with OGC standards for metadata catalogs and observations.

Alessandra Giorgetti agiorgetti@ogs.it

Alessandra Giorgetti is an expert in oceanographic data management. Senior technologist at OGS and deputy director of its Oceanography Section, she has led the Italian National Oceanographic Data Center since 2006, within UNESCO's IODE network. Giorgetti coordinates EMODnet Chemistry, promotes open science, and represents Italy in major European marine observation programs. In 2024, she was appointed to an EU expert group for harmonizing ocean observation standards.

