Marine Science with Open Science

Gianpaolo Coro*

*Istituto di Scienza e Tecnologie dell'Informazione 'A. Faedo'
National Research Council of Italy (CNR)
Via Moruzzi 1, 50124 Pisa, Italy
e-mail: coro@isti.cnr.it, web page: https://www.isti.cnr.it/about/people.php?id=383

ABSTRACT

Modern approaches to Science are being changed by the need to extract knowledge from the large amount of data produced by Web and mobile applications, and by scientific and industrial experiments [1]. These so called "big data" are characterised by large volume and high production rate, heterogeneous complexity and representations, and unreliability of their content, whilst their purpose and longevity are dependent on the possibility to combine and process them within models and processes. In the last decade, novel and interdisciplinary paradigms have been proposed to face the challenges of big data collection and processing, and to promote collaborative Science. Among these paradigms, the Open Science (OS) provides technological and methodological requirements to ensure longevity of both data and processes, based on the three "R"s of the scientific method: Reproducibility, Repeatability, and Re-usability [2]. Key elements of OS are collaborative systems for data processing and the open publication of processes and results. E-Infrastructures (eIs) are computer science systems that follow the OS paradigm. An eI is a network of hardware and software resources that allow users or scientists residing at remote sites to collaborate and exchange information in a context of data-intensive Science [3]. An eI allows integrating processes from several domains, transforming them into services, and connecting data from several sources to make them available to groups of scientists working together. To this aim, an eI usually includes (i) distributed storage and parallel/distributed processing systems, (ii) secure multi-policy data access and harmonisation services, (iii) data and models integration services and related catalogues, and (iv) data sharing and social networking facilities. The effectiveness of eIs is correlated with the representation of data and processes though standards, the interoperability between data-access, processing, and sharing services, and the tracking of experimental input, output, and parameters (*provenance*).

In this presentation, I will show examples of eI-based OS experiments in Marine Science that are compliant with the EU Blue Growth strategy. The aim is to explain how big data produced by hydrodynamic and environmental models and vessels monitoring systems can be combined with big data of marine species observations, through Artificial Intelligence models developed in other domains (e.g. material science and telecommunications), with the aim to produce new knowledge. For example, this combination allows to predict invasions of marine areas by dangerous species and to assess species habitat suitability in ocean areas [5, 4]. Projections of environmental parameters in time under different greenhouse gases emissions - calculated by several data providers - can be combined together to understand the response of marine areas to climate change and its consequences on fisheries, biodiversity, and food provisioning [6, 7]. Vessel data and annual catch statistics can feed complex machine learning and signal processing models to assess fisheries stocks status [10], predict future exploitation locations [9, 8], and advise policy makers to prevent disasters in the fisheries industry [11]. All these OS experiments are exactly repeatable through Web services and Web interfaces and thus allow for third-party validation, which enhances their reliability. Further, they can be applied to other similar cases (reproducibility), and all data and processes are represented through Open Geospatial Consortium standards and are reusable in other experiments. Thanks to these properties, several of them have been adopted by policy makers and currently have concrete impact on society [12, 13].

REFERENCES

[1] Hey, T., Tansley, S. and Tolle. K.M. *The fourth paradigm: data-intensive scientific discovery*. 1st Edition, Microsoft Research, 2009.

[2] EU Commission. Definition of Open Science. *Open Access Web publication: https://ec.europa.eu/programmes/horizon2020/en/h2020-section/open-science-open-access*, (2016).

[3] Coro, G., Panichi, G., Scarponi, P. and Pagano, P. Cloud computing in a distributed einfrastructure using the web processing service standard. *Concurrency and Computation: Practice and Experience*, Vol. **29 (18)**, pp. e4219, (2017).

[4] Coro, G., Vilas, L. G., Magliozzi, C., Ellenbroek, A., Scarponi, P. and Pagano, P. Forecasting the ongoing invasion of Lagocephalus sceleratus in the Mediterranean Sea. *Ecological Modelling*, Vol. **371**, pp. 37-49, (2018).

[5] Coro, G., Magliozzi, C., Ellenbroek, A. and Pagano, P. Improving data quality to build a robust distribution model for Architeuthis dux. *Ecological modelling*, Vol. **305**, pp. 29-39, (2015).

[6] Coro, G., Magliozzi, C., Ellenbroek, A., Kaschner, K. and Pagano, P. Automatic classification of climate change effects on marine species distributions in 2050 using the AquaMaps model. *Environmental and ecological statistics*, Vol. **23 (1)**, pp. 155-180, (2016).

[7] Coro, G., Pagano, P. and Ellenbroek, A. Detecting patterns of climate change in long-term forecasts of marine environmental parameters. *International Journal of Digital Earth*, Vol. **13** pp. 1-19, (2018).

[8] Coro, G., Large, S., Magliozzi, C. and Pagano, P. Analysing and forecasting fisheries time series: purse seine in Indian Ocean as a case study. *ICES Journal of Marine Science*, Vol. **73 (10)**, pp. 2552-2571, (2016).

[9] Coro, G., Fortunati, L. and Pagano, P. Deriving fishing monthly effort and caught species from vessel trajectories. *In OCEANS-Bergen, 2013 MTS/IEEE*, pp. 1-5, (2013).

[10] Froese, R., Demirel, N., Coro, G., Kleisner, K. M. and Winker, H. Estimating fisheries reference points from catch and resilience. *Fish and Fisheries*, Vol **18**, pp. 506–526, (2017).

[11] Froese, R., Winker, H., Coro, G., Demirel, N., Tsikliras, A. C., Dimarchopoulou, D., Scarcella, G., Quaas, M., and Matz-Lück, N. Status and rebuilding of European fisheries. *Marine Policy*, Vol. **93**, pp. 159-170, (2018).

[12] Oceana. Report of the presentation of the CMSY EU stocks assessment at the EU Parliament. *Open Access Web publication: https://eu.oceana.org/en/our-work/froese-report/overview*, (2017).

[13] Food and Agriculture Organization of the United Nations. Expert Meeting on Climate Change Implications for Mediterranean and Black Sea Fisheries and advices for Mediterranean countries. *Open Access Web publication: http://www.fao.org/family-farming/detail/en/c/1142401/*, (2017).