ercim-news.ercim.eu Number 123 October 2020

ERCIM NEWS

Special theme:

Blue Growth

Also in this issue

Research and Innovation"

Evaluation of Synthetic Data
for Privacy-Preserving Machine Learning

Understanding and Managing Ocean Sustainability: The Blue-Cloud Project

by Leonardo Candela and Pasquale Pagano (ISTI-CNR)

The Blue-Cloud flagship project of the Directorate-General (DG) for Research and Innovation Unit of the European Commission is establishing a thematic marine cloud serving the blue economy, marine environment and marine knowledge agendas and the European Open Science Cloud. The project links the horizontal e-infrastructures supported by DG CONNECT and DG GROW, long-term marine data initiatives supported by DG MARE, research infrastructures supported by DG for Research and Innovation and other recently funded thematic clouds.

The oceans, seas, and coasts are home to diverse marine ecosystems, provide a wealth of resources, influence climate, and offer many economic opportunities. However, marine ecosystems are sensitive to long-term global change and combinations of local pressures can affect ecosystems in unpredictable ways. A thorough understanding of the natural dynamics and effects is needed in order to develop sustainable management strategies and conservation of these ecosystems. To support the required research, we need a platform that offers seamless access to marine data as well as computing and required services. The Blue-Cloud project [L1] is addressing this goal through a set of five compelling demonstrators that address societal challenges:

- *Zooplankton and Phytoplankton Essential Ocean Variables Products*, led by the Marine Flanders Institute (VLIZ), is building upon a range of oceanographic data from multiple streams to produce and share unique 3D and 4D synergistic zooplankton and phytoplankton products. The products will contribute knowledge (they represent a class of ocean variables to be measured considered by the scientific community as crucial for ocean observation in a global context) helping to quantitatively reduce uncertainty about the present state of the marine plankton ecosystems and their response to climate change.
- *Plankton Genomics*, led by the European Molecular Biology Laboratory (EMBL), is building an environment enacting a deep assessment of plankton distribution, dynamics and fine-grained diversity to molecular resolution, focusing on discovery of species and functions and exploring genetic and morphological markers of plankton diversity and abundance.
- *Marine Environmental Indicators*, led by the Euro-Mediterranean Center on Climate Change (CMCC), is developing an online service with associated cloud based analytical computing framework and dedicated web interface to provide and display indicators and information on the environmental quality of the ocean.
- *Fish, a matter of scales*, coordinated by the Food and Agriculture Organisation of the UN (FAO), is improving data management and analytic capabilities for fisheries by building a global vertically integrated toolset to manage public fisheries' statistical data from ingestion, and harmonisation through to publication.
- *Aquaculture Monitor*, coordinated by the Food and Agriculture Organisation of the UN (FAO), is using Copernicus data, and combining AI with in-situ datasets to obtain regional-level inventories of aquaculture activities. This will enable the development of a robust and replicable environment for monitoring aquaculture in marine cages and coastal areas.

The Blue-Cloud platform is neither built from scratch nor operating in a vacuum; it is designed to leverage existing and forthcoming data sources, infrastructures, and services operated by diverse providers under different settings. It will offer a unified, yet evolving, working environment that provides access to and use of the aggregated assets. Moreover, it uses and complements the resources of the European Open Science Cloud (EOSC) [L2], i.e. the platform the European Commission is supporting to help scientists find and access data and services across a plethora of providers. The Blue-Cloud platform will contribute data and services to expand on EOSC's existing system.

The Blue-Cloud platform aggregates assets from major domain specific data providers (Figure 1), including SeaDataNet (marine environment), EMODnet Bathymetry (bathymetry), EMODnet Chemistry (chemistry), EurOBIS –EMODnet Biology (marine biodiversity), Euro-Argo and Argo GDAC (ocean physics and marine biogeochemistry), ELIXIR-ENA (biogenomics), EuroBioImaging (microscopy), EcoTaxa (bio images), WekEO (CMEMS ocean analysis and forecasting and C3S climate analysis and forecasting), and ICOS-Marine (carbon). In addition to this content-centric approach, the platform offers a rich array of facilities to support the entire lifecycle of a research workflow with

Figure 1: Blue-Cloud Federation of European Infrastructures.

dedicated working environments built on the D4Science infrastructure [1].

D4Science, which has been operating since 2006, is an infrastructure that facilitates the development and operation of virtual research environments (VREs) for several "communities of practice", from domains including agri-food, earth science, marine science, social sciences, and humanities. Each VRE is a dedicated working environment designed to serve the needs of its community. D4Science-based VREs are web-based, community-oriented, collaborative, user-friendly, open-science-enabler working environments for scientists and practitioners who are collaborating on research tasks.

From an end-user perspective, each VRE manifests as a unifying web application (and a set of application programming interfaces (APIs)) that: (i) comprises several components and (ii) runs in a plain web browser. Each VRE component transparently provides the user with services and facilities possibly operated from diverse providers. In fact, every VRE acts as a gateway that provides seamless access to relevant datasets and services while hiding their diverse origins. Each VRE offers some basic components that help users collaborate [1], namely: (i) a workspace to organise and share digital artefacts; (ii) social networking to communicate with co-workers using posts and replies; (iii) data analytics to share and execute analytics methods using a transparent distributed and heterogeneous computing infrastructure; (iv) a catalogue to document and publish any digital artefacts. To help users integrate their assets into the VRE, three integration patterns are supported (besides implementing completely new services) [2]: (i) integration of existing applications; (ii) integration of analytics methods and workflows; and (iii) integration of datasets and other resources for discovery and access.

The five demonstrators are building specific working environments, services, tools, and datasets and making all of this available by a dedicated gateway [L3].

In addition to technical work, the project is developing a pragmatic, policy-oriented roadmap to 2030, undertaking extensive stakeholder consultations with the wider marine and ICT communities. The goal is to co-design the future strategic development of the wealth and diverse infrastructures and initiatives for collecting, managing and providing access to marine data.

In summary, Blue-Cloud showcases the development of a platform for marine scientists and stakeholders by intertwining a set of interoperable infrastructures and provides insights on how to further develop the resulting platform in the future.

Links:

[L1]: https://www.blue-cloud.org/
[L2]: https://kwz.me/h1Z
[L3]: https://blue-cloud.d4science.org/

References:

[1] M. Assante et al.: "Enacting open science by D4Science", Future Generation Computer Systems, Volume 101, 2019, Pages 555-563, https://doi.org/10.1016/j.future.2019.05.063

[2] M. Assante et al.: "Co-creation of Virtual Research Environments: the D4Science Experience". 12th International Workshop on Science Gateway (IWSG2020), 2020

Please contact:

Pasquale Pagano
ISTI-CNR, Italy
pasquale.pagano@isti.cnr.it

Reinforcing Fisheries Management through Semantic Data Integration

by Yannis Tzitzikas and Yannis Marketakis (FORTH-ICS)

The European project BlueCloud is developing pilot demonstrator applications with the goal of establishing a marine-themed European Open Science Cloud (EOSC) for the blue economy and marine environment. "Fish, a matter of scales" is one of these demonstrators that aims to improve data management and analytical capabilities of fisheries.

Fisheries management is a laborious task that relies on data analysis using complex models and fine-grained software over several sources of information in order to deduce certain facts with the overall aim of improving the sustainability of fisheries. It includes the usually manual process of identifying and combining different parts of information, which is an extremely time-consuming and error-prone process. The key indicators for efficient fisheries management are stocks and fisheries. Stocks refer to groups or individuals of a species occupying a well-defined spatial range (e.g. swordfish in the Mediterranean Sea). Fisheries describe the activities leading to the harvesting of the fish within a particular area, using a particular method or equipment and purpose of activity (e.g. the Atlantic cod fishery in the area of East and South Greenland). The knowledge of the status and the trends of stocks and fisheries at regional, national and local levels is the key factor for highly reliable fisheries management.

To this end, the Global Record of Stocks and Fisheries (GRSF) [1], developed within the context of the EU H2020 project BlueBRIDGE (GA no: 675680, 2015-2018), has collated stocks and fisheries information from three distinct data sources: FIRMS from the Food and Agriculture Organization of the United Nations (FAO), RAM Legacy Stock Assessment database and FishSource from the Sustainable Fisheries Partnership. These sources were chosen because they contain complementary information both conceptually and geographically. By collating these sources, the reporting coverage of any of the single entities is increased. To achieve