



Delivering marine data from the cloud - Recent developments on the SeaDataCloud Discovery and Access service

Peter Thijsse (1), Dick M.A Schaap (1), and Michele Fichaut (2)

(1) MARIS, Voorburg, Netherlands, (2) IFREMER, France

SeaDataNet is an operational pan-European infrastructure for managing marine and ocean data and its core partners are National Oceanographic Data Centres (NODC's) and oceanographic data focal points from 34 coastal states in Europe. Currently SeaDataNet gives discovery and access to more than 2.3 million data sets for physical oceanography, chemistry, geology, geophysics, bathymetry and biology from more than 650 data originators. The population has increased considerably in cooperation with and involvement in many associated EU projects and initiatives such as EMODnet.

The SeaDataNet infrastructure has been set up in a series of projects in last two decades. Currently the SeaDataNet core services and marine data management standards are upgraded in the EU HORIZON 2020 'SeaDataCloud' project that runs for 4 years from 1st November 2016. This upgrade includes a movement "to the cloud" via a strategic and technical cooperation of the SeaDataNet consortium with the EUDAT consortium of e-infrastructure service providers. This is an important step into the EOSC domain.

SeaDataNet maintains and provides users pan-European Directories of organisations, cruises, monitoring systems and projects, all supported by extensive vocabulary services for the marine domain. One of the main components however is the CDI service that provides users discovery and access to the large federated pool of marine data resources which are managed and maintained by more than 110 connected data centres. Even though the wealth of harmonised data is appreciated and supporting many dataproducts and services, the federated approach and user interface did no longer meet the advanced needs of users. Under SeaDataCloud the CDI Data Discovery and Access service is now being upgraded by introducing a central data buffer in the cloud which will be continuously synchronised by replication from the data centres. The buffer itself will be hosted and horizontally synchronised between 5 EUDAT e-data centres. During the replication additional quality control will take place on the central metadata and associated data in the buffer. The upgraded user and machine-to-machine interfaces will improve the overall quality, performance and ease-of-use of the CDI service towards human users and machine processes.

The presentation will provide more technical background on the upgrading of the CDI Data Discovery and Access service, adopting the cloud. It will report on the current status, the experiences of developing services in the cloud, and demonstrate the functionality of the 1st public release, currently expected in April 2019.