

EGU2020-8073

<https://doi.org/10.5194/egusphere-egu2020-8073>

EGU General Assembly 2020

© Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.



Delivering marine data from the cloud using the SeaDataCloud Discovery and Access service

Peter Thijssse¹, Dick Schaap¹, and Michele Fichaut²

¹MARIS, Voorburg, Netherlands

²IFREMER, Brest, 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. The upgraded services include 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.

One of the main components of SeaDataNet is the CDI Data Discovery and Access service that provides users access to marine data from 100 connected data centres. The previous version of the CDI service was appreciated for harmonising the dataset, but also had some flaws towards usability of the interface and performance. Under SeaDataCloud the CDI Data Discovery and Access service has now been upgraded by introducing a central data buffer in the cloud that continuously synchronises by replication from the data centres. The "datacache" itself is being hosted and horizontally synchronised between 5 EUDAT e-data centres. During the implementation of the replication process additional quality control mechanisms have been included on the central metadata and associated data in the buffer.

In October 2019 the actual public launch took place of the operational production version of the upgraded CDI Data Discovery and Access service. The user interface has been completely redeveloped, upgraded, reviewed and optimised, offering a very efficient query and shopping experience with great performance. Also, the import process for new and updated CDI metadata and associated data sets has been innovated, introducing successfully cloud technology.

The upgraded user interface has been developed and tested in close cooperation with the users. It

now also includes the “MySeaDataCloud” concept in which various services are offered to meet the latest demands of users: e.g. save searches, sharing datasearches and eventually even pushing data in the SDC VRE. The user interface and machine-to-machine interfaces have improved 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, and adopting the cloud. It will report on the current release (<https://cdi.seadatanet.org>), demonstrate the wealth of data, present the experiences of developing services in the cloud, and demonstrate the advantages of this system for the scientific community.