Enhancing SeaDataNet e-infrastructure for ocean and marine data, new opportunities and challenges to foster data re-use

Leda Pecci (1), Michele Fichaut (2), Dick Schaap (3) and the SeaDataCloud consortium

Nowadays, there is a general awareness regarding the potential value of ocean data sharing and the benefits that an e-infrastructure allowing access to harmonised multi-disciplinary ocean data can bring to the society. A huge amount of data is produced every day by ocean observations all around Europe. The ability to leverage this valuable potential depends on the capacity of the already established European (EU) ocean data infrastructures to support new needs in the field of ocean data management and to adopt the emerging technologies.

SeaDataNet (SDN) is a digital ocean and marine data infrastructure to support ocean research and societal needs

One of the most relevant e-infrastructures in Europe managing marine and ocean data is SeaDataNet (SDN) that started in early 2000 and evolved over the years by means of different European funded projects. The infrastructure provides online unified access to many harmonised marine and ocean datasets, alongside data products and metadata services. Its main purpose is to provide easy access to heterogeneous marine data-sets which are managed by the data centres, geographically distributed in all European seas. To integrate the data from observations into a coherent and standard managing system has been a challenge over the years.

SeaDataNet challenges

New technologies are dynamic and in the past few years have changed significantly with a growing impact in data and applications management. For this reason SeaDataNet is upgrading its Information Technology (IT) architecture thanks to the SeaDataCloud project, funded by the European Commission. It has the aim to renew the electronic infrastructure, allowing it to offer new services, based on the cloud and High Performance Computing (HPC) technologies. The partnership with EUDAT, a consortium of high performance centres, has allowed it to improve the sharentability and the availability of the data.

SeaDataNet achievements

SERVICES IMPROVEMENTS

Enhanced graphical user interface (GUI)
Faster downloads
Improved reliability of the system
More agile data ingestion system for data providers
Improved machine to machine interfaces
Expansion with new data types: Gliders, HF-radar, Flow cytometry

NEW SERVICES

DOI Digital Object Identifier minting service assigns a unique and persistent identification to ocean data for purposes of data citation and reproducibility of research results.

MySeaDataNet service provides information to end users about ownership status and history and a tool to save and share searches.

Data ingestion service is a tool accessible in the user space of data centres to deliver metadata in the central repository. A complete picture of the received orders divided into standing and approved is available in online reports, it also includes names, e-mail addresses and affiliations of the end-users that have ordered data.

PROTOTYPES

A VRE (Virtual Research Environment) is available for partners to test and give feedback, it provides different tools for oceanographers with the aim of using the computational capacity and the storage of servers in the cloud. Sensor Web, the on line version of ODV software, for data analysis, allows to aggregate, extract, manage and qualify large dataset. DIVAand, an improved version of DIVA (Data-Interpolating Variational Analysis) which is a software designed to perform spatial interpolation of oceanographic observations, has been implemented in the cloud. Furthermore, the VRE offers advanced visualisation services to gain more knowledge. It allows to deal in an efficient way problems involving multidimensional matrices for the creation of climatology and other gridded products. Sensor Web Enablement (SWE) demonstrator is for publishing data using Sensor Web standard, it makes possible to share real-time sensor data streams.

Towards FAIRness in data sharing some activities are ongoing to improve the data, accessibility, interoperability and reuse of the available data.

Acknowledgement

SeaDataCloud project has received funding from the European Union’s Horizon 2020 Research Infrastructure programme, under the grant agreement No 730960”. The authors would like to thank the SeaDataCloud partners for their contributions to the SeaDataCloud infrastructure.