



INSITUDE: web-based collaborative platform for centralized oceanographic data reception, management, exploration, and analysis

Dmitry Khvorostyanov, Victor Champonnois, Alain Laupin-Vinatier, Jacqueline Boutin, Gilles Reverdin, Nathalie Lefèvre, Antonio Lourenco, Alban Lazar, Jean-Benoit Charrassin, and Frédéric Vivier

CNRS, LOCEAN/IPSL, France (dimitry.khvorostyanov@locean.ipsl.fr)

LOCEAN laboratory of the Pierre and Simon Laplace Institute (IPSL) is in charge of a number of scientific projects and measurement campaigns that result in a large flow of heterogeneous oceanographic data managed at LOCEAN. The data are of various origins and include *in situ* data from buoys, ships, moorings, marine mammals and satellite missions for salinity, altimetry, ocean color, temperature. LOCEAN also has an instrumental development team that designs and deploys buoys in various parts of the global ocean, with a need to receive and track the data in the near-real time. The data PIs can be involved in different research groups and projects, and while focusing on providing their data, they might need to collaborate with other teams providing complementary datasets.

To address these needs, the INSITUDE platform is developed at LOCEAN with these goals in mind: (1) receive, manage, track in the near-real time, and explore diverse data; (2) assist scientific experts in the data quality control; (3) facilitate cross-uses of *in situ* and satellite data available at LOCEAN.

The software consists of four components: (1) Django application for the meta-data management; (2) Data processing software (Python); (3) Flask application for server-side interactions with the database; (4) Interactive data exploration/validation front-end.

The basic workflow involves the following steps:

(1) The user specifies the relevant meta-data using the web interface of the Django application; the meta-data database is thus updated;

(2) The processing core is launched automatically at regular times during a day: it reads the meta-data from the database, queries the mailboxes and/or external web services for the data requested, receives, decodes and processes the data, and fills the measurements database. It also generates ASCII data files for selected datasets, which can be downloadable via dedicated web pages or can be used for processing with external user programs (e.g. matlab or python scripts);

(3) The data stored in the measurements database can be interactively explored using DataViewer applications, allowing zoomable views of time series, vertical profiles, and trajectories shown on the virtual globe. Data from different campaigns and for different variables can be viewed together. The quality control assistant allows experts to seamlessly validate the data by assigning quality flags to selected data points or regions, optionally after computing relevant statistics. The validated data can then be visualized and saved based on desired quality flag values.

The INSITUDE platform facilitates data sharing across multiple teams and collaborations between data providers and data experts, researchers and engineers, enabling research projects focused on cross-exploration of various datasets, studies of processes involving both in situ and satellite data, and interpretation of in situ data in a larger-scale context owing to the satellite data. The system offers centralized intuitive acquisition control and access to the data received, along with the related meta-data (projects, campaigns, buoys, people, etc.), facilitates data quality control/validation.

The INSITUDE platform is currently used at LOCEAN and can be deployed in data centers of national data infrastructures, such as the French ODATIS/DATA TERRA.