Building an online and interactive scientific data explorer for LifeWatch observatory data

Dekeyzer Stefanie, Deneudt Klaas, Wydooghe Benny, Mortelmans Jonas, Debusschere Elisabeth, Reubens Jan and Hernandez Francisco

Flanders Marine Institute (VLIZ), InnovOcean site, Wandelaarkaai 7, 8400 Ostend, Belgium
E-mail: stefanie.dekeyzer@vliz.be

The marine observatory that is being established as a Flemish contribution to LifeWatch has become almost fully operational: the spatial surveys with the RV Simon Stevin that cover the Belgian part of the North Sea are running with a monthly frequency; the sensor networks with biosensors such as bird GPS tags, acoustic fish tag receivers, C-PODs, and acoustic bat recorders are growing; and data is being generated constantly. In order to provide access to the observatory data and to support scientific validation of the generated data, a scientific data explorer was developed. This online and interactive tool allows users to perform exploratory data analysis and create advanced data visualizations, without any data visualization experience needed.

In the data explorer application, users can simply define a selection of data they are interested in. Consequently, the application displays a series of visualizations based on the default settings: choropleth point maps, box plots, times series line charts and X-Y plots. Further modifying or optimizing the visualizations is straightforward through an easy accessible side bar. The user can request additional data fields such e.g. as tides, moon illumination, and sun position. The data can be displayed in tabular form and downloaded as tab delimited text files to the users' local drive for further analysis.

From a technological perspective, the tool has been built using the R Shiny framework. R Shiny is a web application framework for R, suitable for both desktop and mobile devices. The choice for using this R based approach is supported by the fact that R has well-developed, powerful, high level functions for data processing and visualization. Since R is broadly accepted as open source programming language for data analysis in biodiversity and ecosystem studies, the developed scripts can be made available and re-used by scientists. A dedicated and performant virtual Linux machine has been set up at VLIZ running the Shiny server. Furthermore, the application uses Leaflet, ggplot2, Plotly, Dygraph and DataTables. The system is able to query MSSQL, PostgreSQL, Geoserver (WFS) and MongoDB servers.

The application is made available through the LifeWatch.be regional portal: http://lifewatch.be/en/lifewatch-data-explorer. Six thematic portals have been set up: Underway Data Explorer, Station Data Explorer, Zooplankton Data Explorer, GPS Bird Tracking Data Explorer, Fish Telemetry Data Explorer, and Bat Recorder Data Explorer.

During the VLIZ Marine Science Day, the LifeWatch Data Explorer will be demonstrated during an interactive game.

Keywords: observatory data; data exploration; data analysis; data visualization; scientific validation