Contribution ID: **9** Type: **Oral**

A programmable web platform for distributed data access, analysis and visualization

Tuesday, 6 July 2021 17:00 (10 minutes)

We introduce a novel client-server Web platform for data access, processing, analysis, and visualization. The platform was designed to simultaneously meet a set of capabilities not available in a joint way in similar software systems. The platform: (a) provides secure access to large amounts of data hosted in institutional data servers; (b) allows users to operate in any modern device (computers, tablets, even smartphones); (c) is intuitive to use and provides on-line help, making it easy to use, even for sophisticated data analysis; (d) runs in a distributed environment, profiting from remote hosting and computing power; and (e) allows integration of heterogeneous data and user provided data analysis codes in different programming languages. These requirements were inspired in the needs of users for the analysis of the massive databases of nuclear fusion devices, in particular, the ITER database.

The client runs in any HTML enabled device, under virtually any operating system, and allows the user to interactively specify a desired data flow through a series of analysis or visualization routines. This flow takes the form of a graph of nodes (or modules), each one consisting of an icon with customizable properties. Each icon encapsulates a given data processing algorithm and the whole set of icons form a ready to use library of standard data handling, analysis, and visualization routines. Users conducting the analysis use their field of expertise to design the desired combination of routines, appropriately customize their properties and, optionally, add new icons with their own code. Each icon can encapsulate routines written in one of several accepted programming languages, including Fortran, C, Matlab, Python and R.

For the execution of graphs, there is a specific server that receives a request for traversing the specified graph. Then, each graph node is executed in sequence and connects its output with the input of the next ones (the output can be used as input to several modules), as indicated by the graph. The system transparently takes care of accessing data in remote servers, checking access permissions, transferring data if necessary, running code written in different programming languages in possibly different computing facilities (potentially in parallel ways), passing data to and from the inputs and outputs of the nodes, and creating the required final output data or visualization plots. The result, also in HTML form, is returned to the client for user inspection. It should be noted that the platform has been prepared not only for the interactive execution of codes but also for batch processing.

The presentation describes the platform, its architectural design and the implementation decisions. We provide an example of use that implements the development of an adaptive disruption predictor from scratch.

Member State or IGO

Spain

Speaker's Affiliation

Universidad de Murcia, Murcia

Primary authors: Prof. ESQUEMBRE, Francisco (Universidad de Murcia); Dr CHACÓN, Jesús (Universidad Complutense); Dr SAENZ, Jacobo (UNED); Dr FABREGAS, Ernesto (UNED); Dr FARIAS, Gonzalo (Universidad Católica de Valparaiso); Dr VEGA, Jesús (Laboratorio Nacional de Fusión, CIEMAT); Prof. DORMIDO-CANTO, Sebastián (UNED)

Presenter: Prof. ESQUEMBRE, Francisco (Universidad de Murcia)

Session Classification: Remote Experimentation and virtual lab 1

Track Classification: Remote Experiments and Virtual Laboratory