

IODA: a new federated web platform for collaboration and sharing of data analysis resources in Fusion Data Research

Wednesday, 14 June 2023 16:30 (30 minutes)

The analysis of data from fusion devices is a common and important task in Fusion Data Research (FDR). One which poses several practical challenges to scientists.

First, experimental programs generate an enormous amount of data which is hosted by dedicated institutions. Accessing this raw data requires authorization and either a very fast direct connection or enormous local storage capabilities. Second, the analysis of this data involves applying many standard data analysis and visualization libraries, dedicated domain-specific routines, and programming new, experimental code. Scientists constantly need to remain fluent in a great deal of applicable software routines, libraries, and platforms, which are often under development, sometimes not-yet-well-documented, and perhaps implemented in different programming languages. Finally, the complexity of the computations and the size of the data may imply long processing times unless high computational power is used. This may include parallel computation, highly specialized software, or dedicated hardware (such as GPU or FPGA) ...the installation and use of which are far from simple.

These three peculiarities make FDR work laborious, difficult to communicate, and more difficult to reproduce. The FDR community needs to set a common platform for efficient work in the discipline. One that, ideally, establishes an open, federated way for scientist to share data, validated analysis software, and computing equipment, while respecting scientists' freedom to choose from whatever exists or add newly developed tools. Such a platform would not only facilitate FDR daily work, but would also make communication, reproducibility, and replicability of results easier.

This paper presents IODA (acronym for Input-Output Data Analysis), a new client-server Web platform that aims to provide a viable solution for the cited problems. IODA clients run on any Web-enabled device (PCs, laptops, and tablets) allowing scientists to interactively design a directed graph representing a given access and analysis of remote, distributed data, and the visualisation of the results. The client can then send the graph to the main server for execution, which a cloud of federated computing servers cooperatively run, returning the result of the computations back to the main server for the client to analyse.

The joint, transparent to the user, work of this ecosystem effectively provide scientists with i) simplified, secure access to distributed data, ii) verified software routines for analysis and visualization, iii) access to network-available specific computing hardware, and iv) the capability to introduce the user's own code in the analysis.

We will show the user interface of the platform's client, describe its server-side architecture (the real heart of the federated data analysis capabilities), and list current and future platform components that help address the cited FDR community needs.

Speaker's Affiliation

Universidad de Murcia

Member State or IGO/NGO

SPAIN

Primary authors: ESQUEMBRE, Francisco (Universidad de Murcia); Dr SAENZ, Jacobo (UNED); VEGA, Jesús (CIEMAT); Dr CHACÓN, Jesús (Universidad Complutense de Madrid); DORMIDO-CANTO, Sebastián (UNED)

Presenter: ESQUEMBRE, Francisco (Universidad de Murcia)

Session Classification: DB/2 Information retrieval, dimensionality reduction and visualisation in fusion databases

Track Classification: Information retrieval, dimensionality reduction and visualisation in fusion databases