Science Demonstrator 1: Support EISCAT_3D Users to Reprocess Data Using User's Algorithms (Use Case IC_3)

Please provide your feedback on this Science Demonstrator using the questionnaire at https://survey2.icos-cp.eu/ENVRIplus-evaluator!

Overview

Often the data products from a RI, derived from lower level/raw data, are pre-cooked. That means they are produced by applying some default processing algorithms and parameters, like spatial and temporal resolution, the use of model parameters, model and process selections. This demonstrator is an implementation results of IC_3 use case that describes a model that enables scientific researchers to re-process data by defining other selections of these inputs from other sources.

Scientific Objectives

EISCAT_3D is an environmental research infrastructure on the European Strategy Forum on Research Infrastructures (ESFRI) roadmap. Once assembled, it will be a world-leading international research infrastructure to study the high atmosphere in the Fenno-Scandinavian Arctic and to investigate how the Earth's atmosphere is coupled to space.

In general, EISCAT data products (including future EISCAT_3D data) are not raw data as sampled in the receivers, but have already undergone several steps of processing (along the chain voltages -> filtering -> time averaged spectral data -> inversion of physical parameters) by standard analysis algorithms in the EISCAT ICT system. This also implies that data are processed with a predefined set of parameters, e.g. spatial and temporal resolution, inversion model selection, model parameters and allowed parameter ranges.

This use case addressed a requirement of the EISCAT user community, namely to allow individual scientists to process the original experimental data using their own algorithms. The challenge in this use case is how to make use of ENVRIplus services to demonstrate that EISCAT scientists can re-process data by defining other selections of parameters and algorithms.

Description

A typical usage scenario is that users select data together with an algorithm and invoke a workflow with tuned parameters. The data to be used in the test case is from the present EISCAT facilities, and the processing software is provided by EISCAT and originally written in Matlab. In this use case, we chose a Matlab processing package that could be converted into open source software. The process is to generate a graphical visualisation of the experimental data. Figure 1 shows the EISCAT Real Time Graph (RTG), which is primarily developed to follow the radar run in real-time and to produce a display of basic parameters, such as returned power spectra.

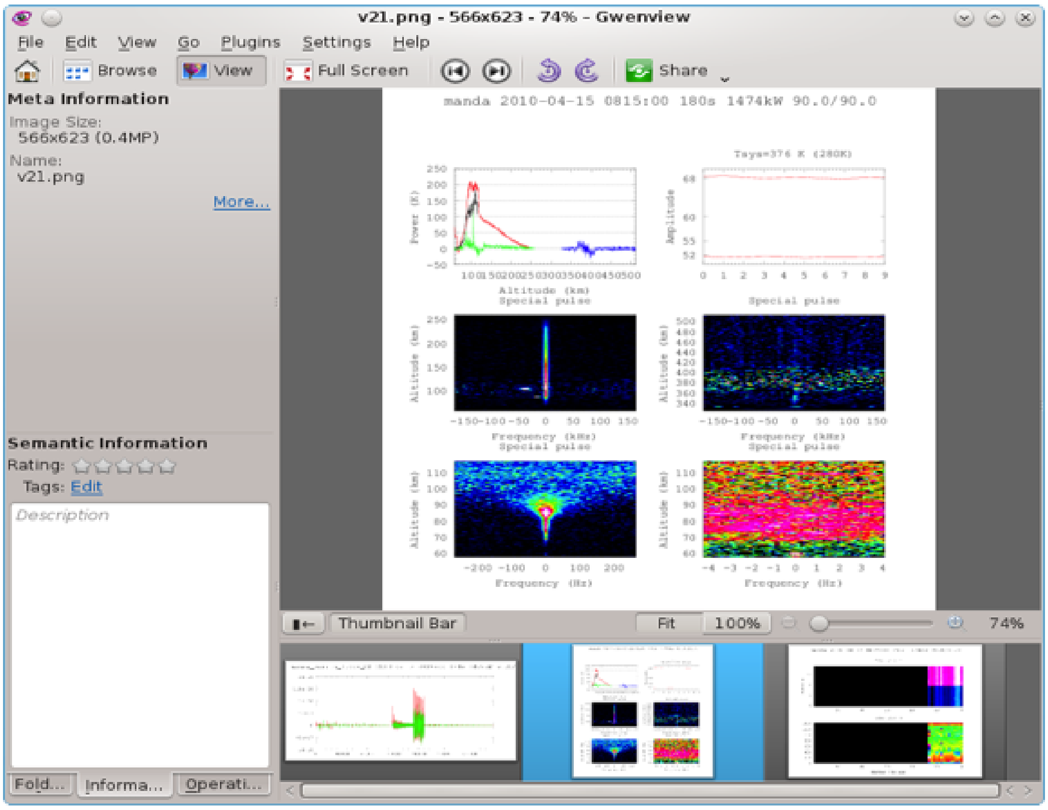

Figure 1: EISCAT Real Time Graph produces visualisation of the processing of raw EISCAT radar data with basic parameters.

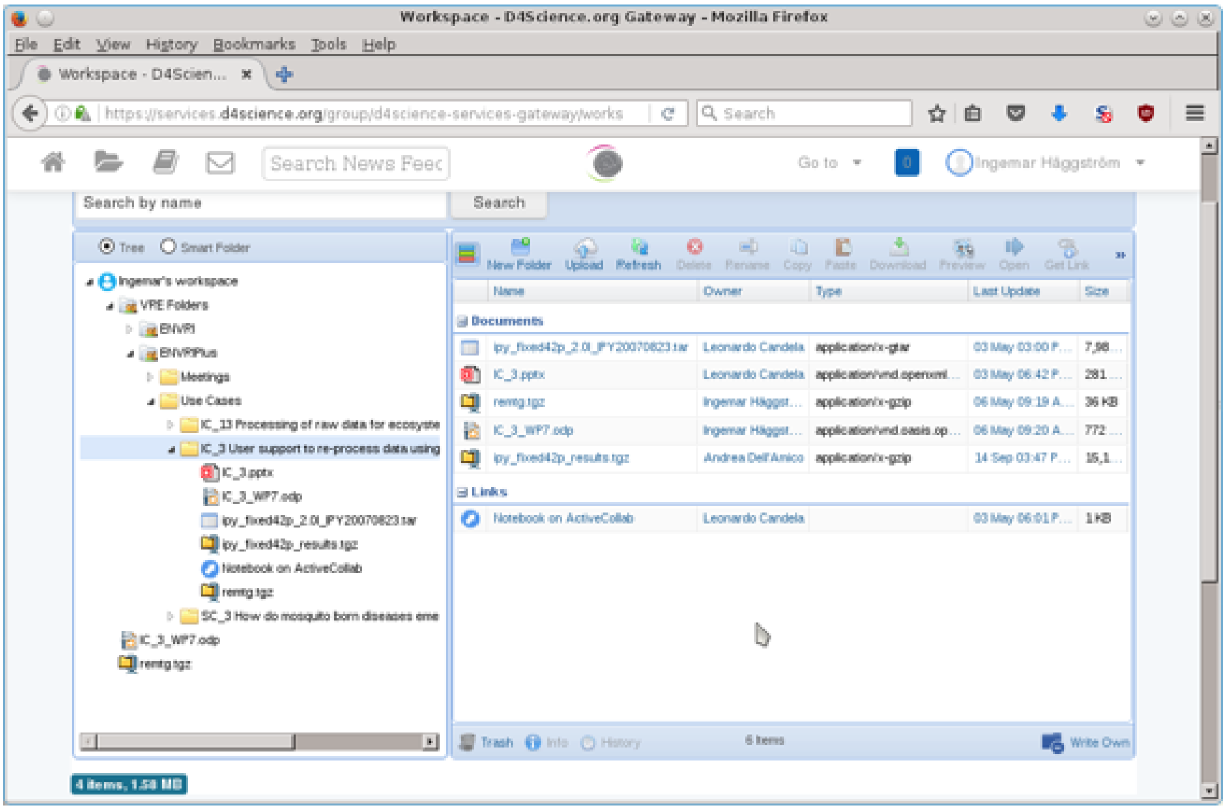

Figure 2: In gCube, EISCAT users can create processing algorithms and share them with others.

The RTG was modified to run under GNU Octave, and to produce the plots in batch mode with some selected input parameters. The whole setup was installed into the D4Science gCube Virtual Research Environment, provided by ENVRIplus WP7. The D4Science gCube Data Analytics is executed via a web dashboard based GUI. The ticketing system of D4science has been effectively used for tracking the implementation and communicating with WP7 developers. Some initial results are shown in Figure 2, which demonstrates that in the gCube platform, the EISCAT RTG algorithms are translated into Octave, and compiled and published in an ENVRIplus community space. In the same way, an algorithm created by a single EISCAT user can be shared and reused by others.

Advantages

For many ENVRIplus RIs, the ultimate goal is to provide quality-checked observation data to their user communities. The ease of user access to data and analysis resources has a direct impact on the users' ability to deliver new research results and is thus directly linked with the impact and value of the RIs. This issue is therefore given increased emphasis when RIs design their ICT systems. In recent years, Virtual Research Environments (VREs) have emerged as an important approach to provide web-based systems helping researchers to collaborate. WP7 (T7.1) has set up a VRE for the ENVRIplus community using the D4Science platform. D4Sciencesupports a flexible and agile application development model based on the notion of Platform as a Service (PaaS), in which components may be bound instantly at the time they are needed. In this way, it enables user communities to define their own research environments by selecting the constituents (the services, the data collections, the machines) among the pools of resources made available through the D4Science e-Infrastructure.

Access to a VRE for user-driven data analysis is a common request not only in the environmental sciences community. The implementations from this use cases can be easily extended to support other domain applications' needs, for example, it would benefit space and solar-terrestrial physics, solar system physics (meteors and asteroids), and astronomy. The pilot experiences can be easily shared with the whole EISCAT community, which covers member countries including, Finland, France, Norway, Sweden, USA and UK, and its global collaborations reach China, Japan, Korea.

Link to the Demonstrator

Youtube video is at: https://youtu.be/YEEMUvnSHUM

The service prototype can be accessed at: https://services.d4science.org/group/rprototypinglab/data-miner .

Simply request a use account, you will be able to login to the VRE. There, on 'Executes an experiment', selects ENVRI and Webtg V2.

Contributors

- Dr Ingemar Häggström, EISCAT Scientific Association, ingemar.haggstrom@eiscat.se