



PARALLEL TRACKS – TRACK 4 – NEW SERVICES

During the preparation phase of the EUDAT proposal a number of scientific communities were engaged with and were involved in discussing the first EUDAT core service set. From this the well-known short list of 4 core services was extracted, the requirements were defined and basically much focus was put to get these 4 services running. Due to the ongoing interactions it became obvious that one of these services has 4 different flavors for different types of data providers. In parallel EUDAT defined two more services which are currently being worked on also a result of the EUDAT interactions (see the note on community engagement on the web: <http://www.eudat.eu/published-articles>).

Now it is time to begin further intensive discussions about what the candidates for future common services within a Collaborative Data Infrastructure could be. Already back in March 2013 at the EUDAT user forum in London some candidates were proposed. Currently EUDAT has launched a questionnaire covering many communities and individuals to understand their preferences and ideas. Furthermore, for some areas (dynamic/realtime data, workflow support, semantic services, access issues) EUDAT is organising special workshops in Mid-September to understand better what services could be useful as solutions for these areas. The EUDAT 2nd conference will be a further step, the presentation of an evaluation of all suggestions received to date and opportunity to openly discuss the outcome with all community experts and individuals that are interested. This will help EUDAT to find proper service cases for the coming phases that are based on the opinions of multiple communities together with the 5 core ones.

PARALLEL TRACKS – TRACK 4 – NEW SERVICES

PARALLEL SESSION: TRACK 4 – NEW SERVICES - 4.3 WORKFLOWS

CHAIR: CHRISTIAN PAGÉ, CERFACS

DATE & TIME: WEDNESDAY 30TH OCTOBER – 09:00 - 10:30

ROOM: VERDI + MASCAGNI

OVERVIEW:

Well-described and documented scientific workflows that can be executed to achieve new results are becoming more and more important in all scientific disciplines to cope with the increasing amount of data in appropriate ways and to increase the reproducibility of scientific results. This is true both for raw data generated by sensors and software systems and processed in regular ways, and also for many areas of derived data - the long-tail data. As we move towards “data fabric solutions”, workflow support for manipulating data will be essential. Data infrastructure initiatives such as EUDAT and DataONE are already working on workflow systems and building up expertise, while large institutions such as LANL and SDSC are also looking into such workflow systems to offer services for scientists. It is not yet fully clear which environment will be offered in these cases or exactly what types of services is data infrastructures should offer. EUDAT will continue to work with community experts to test service concepts that allow users to execute workflows on data stored in the EUDAT data domain.



AGENDA

Workflows Chairs: Morris Riedel, Christian Page
Morris Riedel: Workshop Results
Christian Page: Climate Use Case
Erhard Hinrichs: Linguistics Use Case
Other contributions and General discussion on Workflow Issues in EUDAT
Update New Services
Final Round on New Services