The EVER-EST Virtual Research Environment (VRE): outcomes and solutions for Earth Science

Grande V.1, Castellan G.1, De Leo F.1, Foglini F.1, Ferraresi S.2, Gómez J. M. 3, Mantovani S.2, Marelli F.4, Palma R.3

1 Institute of Marine Sciences - National Research Council (ISMAR-CNR), Via Gobetti, 101 40129 Bologna, Italy
2 Meteorological and Environmental Earth Observation (MEEO), Viale Volaro 195/A Ir. 21 44123 Ferrara, Italy
3 Pianrat Supercomputing and Networking Center (PISC), Iana Paws 0 20, 62-139 Poznań, Poland
4 Transaer Sr. Via Giovanni Amendola, 46 00185 Rome, Italy

The EVER-EST Project

The EVER-EST project aimed to develop a generic Virtual Research Environment (VRE) tailored to the needs and validated by the Earth Science domain. Real use cases, taken from pre-selected communities covering different Earth Science research scenarios led the implementation of the system's services and capabilities. ISMAR-CNR represented the Sea Monitoring Virtual Research Community (VRC) to do with: information fragmentation, lack of formalized methodologies, lack of a unique platform where the community can work together and with powerful tools.

Central to this approach is the concept of Research Object (RO): a semantically rich aggregation of resources that bring together data, methods and people in scientific investigations. ROs enable the creation of digital artifacts that can encapsulate scientific knowledge and provide a mechanism for sharing and discovering assets of reusable research (Fig. 1).

The EVER-EST VRE (Fig. 6) allows users to create, manage and share Earth Science ROs, to discover, access and process data, to visualize data relying on OGC standards (OpenSearch, WCS, WPS, WMS), to execute remote workflows. The VRE enables scientists to collaborate in different modes and to communicate throughout the research lifecycle (e-collaboration functionalities): the portal provides access to training materials and online courses (e-learning functionalities) to guide future data scientists to overcome the daily challenges and support their ability to curate the information.

The VRE integrates different tools:

- Collaboration spheres, a user-oriented interface for the visualization of correlation between similar objects based on collaborative filtering and versatile keyword content-based recommendations (Fig. 4);
- Jupyter Notebook, for capturing the whole computation process, developing, documenting, and executing code, as well as communicating the results;
- ROHub, a digital library system supporting the storage, lifecycle management, sharing and preservation of research findings via ROs (Fig. 5);
- SeaFile, a repository for the personal data storage that guarantees file synchronization, version control, public link sharing, desktop client and web API;
- Virtual machines, a powerful shared environment for create workflows and process data.

Conclusion

The Sea Monitoring user had to change his mentality about different aspects of the scientific daily life. The most relevant work concerned in formalize the scientific lifecycle, moving from a conceptual to an executable workflow, in fact, the VRC spent more time in implementing and orchestrating different codes and scripts re-executable. We organized the scientific researches to make them entirely sharable and open to other scientists. The EVER-EST approach allows the users to:

- saving time in data processing;
- having more storage space and higher powerful computer capabilities;
- easily exchange data and processing methodologies;
- better and quick communicate;
- access remotely data, software, research results, and documentation;
- organize a scientific workflow in a single digital object, findable and reusable;
- increase the number of publications and visibility from public;
- document scientific work;
- maintaining attribution through DOI placement;
- publish grey literature;
- ensure long term preservation of research work;
- re-use, preserve and share both data and scientific processes.