Number 109 April 2017

ERCIM NEWS

www.ercim.eu

Special theme:

Autonomous Vehicles

Also in this issue:

Research and Society:

Virtual Research Environments

Making the Development and Deployment of Virtual Research Environments Easy and Effective

by Leonardo Candela, Donatella Castelli and Pasquale Pagano (ISTI-CNR)

Virtual research environments are emerging as an invaluable tool for scientists, enabling professionals in different fields to collaboratively and seamlessly access and use resources (computing, datasets, services) spread across several providers. This solution is particularly relevant in long-tail science contexts, i.e., when researchers and practitioner communities lack dedicated resources to perform their research. Implementing such a solution requires an approach that is open, flexible, and can easily evolve.

Virtual Research Environments (VREs) are web-based, community-oriented, collaborative, user-friendly, open-science-compliant working environments for scientists and practitioners working together on a research task [1]. They share commonalities with Science Gateways (SG) and Virtual Laboratories (VL). The overall goal is to provide scientists and researchers with integrated and user friendly access to data, computing and other services that are usually spread across multiple diverse data and computing infrastructures. Furthermore, they are designed to enact and promote collaboration among their members.

To develop and operate this type of working environment, it is necessary to: (a) develop a set of interoperability solutions that can interface the VRE services with the subset of existing resources that are relevant to a particular research project and offered by 'third-party' providers, (b) develop a set of basic user-friendly services promoting collaborative and open-science-friendly interaction among the VRE members, (c) consider short and long term provisioning of both the working environment and the products resulting from it.

Owing to these characteristics, the development and operation of such environments is incompatible with 'from scratch' and 'isolated' approaches. Developing and maintaining a 'mediator' to interface with a given e-infrastructure requires significant investment in time and effort that can be better afforded by applying economies of scale and scope (namely, by using it in the context of many VREs).

To meet this need, we have designed and developed a technology (actually, a software system) named gCube [2] capable of creating and operating an e-infrastructure offering VREs with the 'as-a-service' paradigm.

gCube has been progressively endowed with: (i) a rich array of 'mediators' for interfacing with existing 'systems' and their enabling technologies including distributed computing infrastructures (e.g., EGI) and data providers (e.g., by relying on standards like OAI-PMH, SDMX, OGC W*S) as well as for making it possible for third-party service providers to easily exploit gCube facilities (e.g., OAuth, OGC W*S, REST APIs); (ii) a set of basic services including a shared workspace where the objects used and resulting from VRE activity (beyond simple files) can be stored, organised and accessed as if they were in a 'standard' file-manager; a social networking area where members of each VRE can have discussions, share news and other material of interest, rate each item of a discussion, classify the discussion items by hashtags, refer to people or groups thus to call for actions from them, etc.; a user management area where authorised people are allowed to manage VRE membership, to create groups, assign members to groups, assign roles to members, invite new members, etc.; an open, customisable and extensible set of facilities made available for the needs of the specific community. These include a project management and issue-tracking system with a wiki, a rich and extensible data analytics platform, a flexible 'products' catalogue where any (research) artefact produced in the VRE that is worth being published can be easily made available by equipping it with rich metadata including licence and provenance, a rich array of domain data management facilities. VREs are created by using a wizard-based approach where a VRE designer is simply requested to select (among the existing ones) the facilities and resources he/she is willing to have in the VRE, and then upon approval the VRE is automatically provisioned and made available by a web-based portal.

This technology is currently enacting the D4Science e-Infrastructure] and exploited to create and operate more than 70 diverse VREs [L1]. Overall, these VREs are serving more than 3,100 (returning) scientists in 44 countries across a rich array of diverse communities usually associated with international initiatives and projects, e.g., i-Marine (fisheries and marine biodiversity scientists), BlueBRIDGE (fisheries and aquaculture scientists, educators & SMEs), SoBigData.eu (social mining scientists), ENVRI+ (environmental scientists), AGINFRA+ (agriculture scientists), PARTHENOS (cultural heritage practitioners), EGIP (geothermal scientists), OpenAIRE-Connect (multidisciplinary community dealing with scholarly communication and open science, EDISON (data science educators).

Link:

[L1] services.d4science.org/explore

References:

[1] L. Candela, D. Castelli, P. Pagano: "Virtual Research Environments: An Overview and a Research Agenda" Data Science Journal. 12, pp. GRDI75–GRDI81, 2013. DOI: http://doi.org/10.2481/dsj.GRDI-013

[2] M. Assante, et al.: "Virtual research environments as-a-service by gCube", PeerJ Preprints 4:e2511v1 2016, https://doi.org/10.7287/peerj.preprints.2511v1

Please contact:

Leonardo Candela
ISTI-CNR, Italy
leonardo.candela@isti.cnr.it