Research activities and innovative developments in European research institutes

The D4Science Research-Oriented Social Networking Facilities

by Massimiliano Assante, Leonardo Candela, Donatella Castelli and Pasquale Pagano

Modern science calls for innovative practices to facilitate research collaborations spanning institutions, disciplines, and countries. Paradigms such as cloud computing and social computing represent a new opportunity for individuals with scant resources, to participate in science. The D4Science.org Hybrid Data Infrastructure combines these two paradigms with Virtual Research Environments in order to offer a large array of collaboration-oriented facilities as-a-Service.

Scientists are expected to produce enhanced forms of scientific communication based on publication of "comprehensive scientific theories" – including the data and algorithms on which they are based –to make it possible for "others to identify errors, to support, reject or refine theories and to reuse data for further understanding and knowledge" [1]. This is a pressing requirement not only in the context of "big sciences" (e.g. physics, astronomy, earth observation) but also in the long tail of science, ie the large number of relatively small laboratories and individual researchers that have the potential to produce a bulk of scientific knowledge but have no access to large-scale dedicated IT.

To effectively serve such scenarios, D4Science.org is operating a Hybrid Data Infrastructure (HDI), ie an IT infrastructure built as a "system of systems" that integrates other infrastructures including grid and cloud, services and information systems. The HDI can thus offer its users a disparate set of technologies, computing and storage resources made dynamically available via the elastic acquisition model offered by Cloud technologies. It provides Virtual Research Environments (VREs) "as-a-Service", ie, web-based working environments where groups of scientists can transparently and seamlessly access shared sets of resources (data, tools and computing capabilities). The VRE services include a social networking area that promotes innovative scientific collaboration patterns inspired by social computing and supported by the underlying infrastructure facilities.

This social networking facility provides an environment to foster large scale collaborations, ie scenarios where many - potentially geographically distributed - co-workers can access and process large amounts of data. It offers: (a) a continuously updated list of events/news produced by users and applications (Home Social) that becomes an easy-to-use trigger of a continuously updated timeline of user-centric facts, (b) a folder-based file system to manage complex information objects, including files, datasets, workflows and maps, in a seamless way (Workspace), (c) an email-like facility for exchanging "large" messages with co-workers, ie messages with attachments in the form of the complex information objects described above (Messages), (d) a list of events organized by date, e.g. publication of, or comments on
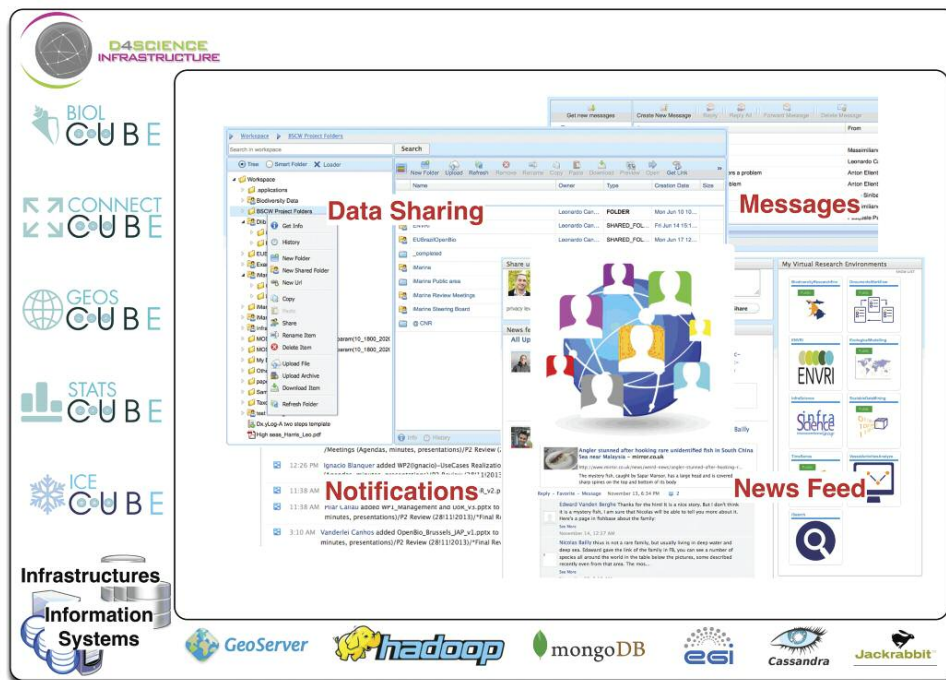
Figure 1: D4 Science infrastructure.

a research product (Notifications), (e) a settings area where the user can configure various aspects of the system including his/her data and notification preferences (Personalization).

The Home Social consists of two facilities. The News Feed makes users' and applications' updates available to every user according to his/her preferences and enables users to comment, subscribe or re-share these updates. These updates are "actionable", e.g. contain a link to the actual product or facility. The Share Updates enables users to post updates or interesting links to others and applications to post possible updates of a new product or facility.

The Workspace resembles a folder-based file system, where the added value is represented by the type of information objects it can manage in a seamless way. It supports items ranging from binary files to information objects representing tabular data, workflows, species distribution maps, time series, and comprehensive research products. Through it, data sharing is fostered, making results, workflows, annotations and documents etc immediately available.

The Messages realise a common email environment as-a-Service whose distinguishing feature is its integration with the rest, e.g. it is possible to send as an attachment any information object residing in the workspace, however big and complex, without consuming bandwidth.

The Notifications alert users on an as-it-happens basis. These notifications offer a sense of anticipation and create a productivity boost. Users receive an alert (through a priori selected channels, e.g. email, web portal, Twitter) notifying them when something of interest has happened in their VRE(s).

The Personalization provides users with facilities to customize the overall behaviour of the "social area". It enables information to be specified, including biographic data, interests and skills, and notification settings.

The D4Science social networking facilities will reshape the modern approach to communication, largely implemented by LinkedIn, Twitter, and Facebook, by porting it to research communities. Sharing of large datasets, quite common in the big data era, and workflows, have become as easy as sharing an image. Scientific collaboration and communication have become immediate and smart by sending a short message or posting a tweet. The virtualization of the working environments through the on-demand and timely creation of dedicated VREs, the virtualization of the resources offered as-a-Service through the HDI, and now the support for collaboration and communication make D4Science a unique service for the effective production of comprehensive scientific theories.

Link:
D4Science website: www.d4science.org

Reference:
[1] G. Boulton et al.: "Science as an open enterprise," The Royal Society, Final report, 2012, http://royalsociety.org/policy/projects/science-public-enterprise/report/

Please contact:
Massimiliano Assante
ISTI-CNR, Italy
E-mail: massimiliano.assante@isti.cnr.it