

myExperiment: social software for sharing workflows

¹David De Roure, ²Carole Goble, ²Jiten Bhagat, ¹Don Cruickshank, ³Marco Roos

University of Southampton, UK; University of Manchester, UK; ³University of Amsterdam, The Netherlands

The ^{my}Grid project (<http://www.mygrid.org.uk>) develops the popular Taverna workflow management system, now used throughout the world by 350+ organisations for a whole range of Life Science problems. An open source initiative from the outset, Taverna's core development activity is supported by the Open Middleware Infrastructure Institute UK (<http://www.omii.ac.uk>), so that e-scientists can rely upon it as part of their regular collection of tools.

Workflow design is challenging and labour-intensive. Reusing a body of prior designs through catalogues is highly desirable. Reuse is a particular challenge when scientists are outside a predefined Virtual Organisation or enterprise. These are individuals or small groups, decoupled from each other and acting independently, who are seeking workflows that cover processes outside their expertise from a common pool of components, especially when workflows are shared across discipline boundaries and when inexperienced scientists need to leverage the expertise of others. As Taverna's popularity increased we observed a workflow exchange activity emerging amongst our users and a strong desire for a place where workflows could be collected and opened to peer review.

myExperiment (<http://myexperiment.org>) is an initiative from the ^{my}Grid project to create a Virtual Research Environment which makes it easier for workflow workers to gossip about and exchange workflows, regardless of the workflow system – Taverna, Kepler, Triana, BPEL etc. Scientists rarely care about the workflow engine they use: they typically care about the workflow itself, its function and the services it uses. Rather than just making workflows available for sharing, myExperiment actually facilitates and encourages that – it takes a step beyond existing workflow repositories by crossing project, community and product boundaries, emphasising social networking around the workflows, providing gateways to other environments and forming the foundation of a personal or laboratory workbench.

myExperiment enables scientists to *discover, reuse and repurpose workflows*, and to *enact them* from a web page using a remote enactment service. In contrast to social websites, it has detailed support for ownership, attribution, licensing, sharing and permissions – meeting the particular needs of the scientist. More generally, the myExperiment concept is about sharing digital objects which include data, results, provenance information, tags, associated documentation, etc. These individual items can be collected together to form *research objects*, for example to record an experiment. myExperiment *aids reuse* because workflows can be discovered not just by what they do but based on how they are used by the community, with tags and reviews adding to the 'collective intelligence'.

myExperiment is an *open source codebase* (released under BSD licence), and individuals and laboratories are free to install their own myExperiment instances; they can then link them up using a federation model if and as they wish. The system is implemented using the Ruby on Rails open source Web application framework using an agile development method. The public site (www.myexperiment.org) was made available in November 2007 and over the period January-April 2008 the site gained over 750 registered users, with many others accessing public content without needing an account; visits came from 78 countries. The site has been well received by the community, carrying over a third of all publicly available Taverna workflows.

While the public site provides a service for those who do not already have sharing and collaboration mechanisms in place, we also expose myExperiment functionality through simple RESTful APIs so that it can be accessed through existing interfaces, including wikis and web pages. This also enables the creation of other interfaces such as Google Gadgets, myExperiment add-ons for sites such as Facebook and functionality mashups over myExperiment. Much of our planned effort is in linking up with other web tools and platforms.

Although myExperiment focused initially on workflows, its underlying architecture is neutral about the objects it manages. Web services and workflows are inter-twined; in Taverna, workflows commonly call remote web services. In the same way that workflows can be catalogued, tagged and curated by the community, so should the growing band of web services in bioinformatics. With this in mind, the ^{my}Grid project, along with the EBI, has just launched (May 2008) the BioCatalogue project. This project builds on myExperiment to create a fully curated catalogue of Web services covering their functional, operational, usage and provenance metadata. By analysing the designs and patterns of services in workflows we can also improve the curation of services.