Renewing bioinformatics workflow system by using a Web 2.0 approach

Colella R[1], Quinto V[2], Vaccina A[1], Scioscia G[1], Leo P[1]

Motivation

In this work a Web 2.0 technology and approach is adopted to enhance a key field of bioinformatic platform research: manage and automate analysis workflows. The approach considered the adoption of recent Web 2.0 technologies, such as mash-up platforms, that enable the rapid creation, sharing, and discovery of reusable application building blocks (widgets, feeds, mashups), known also as consumables, as an alternative environment to support bioinformatic workflows design and execution. The usage of Mashup is expanding in the business environment. Business Mashup, for instance, is adopted for integrating business and data services by providing the ability to develop new integrated services quickly. Typically, Mushup provides organizations a much more flexible modality to combine internal with external services and then creating new services that are usually accessed through user-friendly Web browser interfaces. We applied Mushup principles to the bioinformatics workflow context with the final aim to collect insights to develop a new kind of bioinformatic workflow systems. We did our experiments by prototyping a number of widgets as well as bioinformatics consumables that can be mushed-up in a typical Mushup environment and built some bioinformatic pipelines to validate the effectiveness of our approach.

Methods

Consumables (widget and services) have been developed by using the Lotus Widget Factory, an Eclipse plug-in that provides an easy-to-use development environment that enables developers of all skill levels to create dynamic widgets rapidly almost without writing code (except for a bit of Java and Javascript). Lotus Widget Factory is a component of IBM® Mashup® CenterTM, which also includes:
• the mashup builder, useful to configure and wire widgets on a mashup page; pages can be published to the catalogue and shared with other users. • MashupHub, a catalogue for mashup objects: feeds, data mashups, mashup pages, REST services, spaces, and widgets. The catalog includes community features for sharing information with other users; objects that are stored in the catalogue can be tagged, rated and commented.

[1] IBM Italia S.p.A., Bari [2] Exhicon S.r.l., Bari

Results

A first set of widgets have been developed to perform basic data manipulation operations such as: uploading Flat files, i.e. files containing biosequences and information associated to them, stored in specific formats, from local file system, selection of sequences to be input to the workflow, execution of a REST (Representational State Transfer) service call and some others. The core widget receives data (as XML) from one or more widgets, invokes a generic WSDL-described web service and sends (as XML) the results to all the widgets that the user has wired to it, iterating the execution of the web service call with respect to the input data set dimension and to a user-chosen parameter. Some prototypal workflows have been assembled and tested with a number of these basic widgets, making use of some of the two hundred algorithms of EMBOSS suite, exposed as web services.

Contact e-mail

roberto_colella@it.ibm.com