

Mobyle: a new full web bioinformatics framework

Bertrand Néron¹, Hervé Ménager¹*, Corinne Maufrais¹, Nicolas Joly¹, Pierre Tuffery², Catherine Letondal¹

¹: Groupe Logiciels et Banques de Données, Institut Pasteur, 28, rue du Dr Roux, 75724 PARIS Cedex, France
{bneron,hmenager,maufrais,njoly,letondal}@pasteur.fr; *: presenting author

²: UMR-S 726, RPBS, Université Paris Diderot-Paris 7 75205 PARIS CEDEX 13, France

pierre.tuffery@univ-paris-diderot.fr

website: <http://bioweb2.pasteur.fr/projects/mobyle/>,

downloads: <ftp://ftp.pasteur.fr/pub/gensoft/projects/mobyle/>

Mobyle is the successor of Pise, a system that provides a web environment to define and execute bioinformatics analysis software. The design of the new system is based on the need to extend functionalities towards service interoperability and data integration, while optimizing the usability of the interface for biologists. Based on extensive user studies, we developed the end-user interface as a Web Portal that provides a **global and integrated view** of all the elements needed to perform analyses. The system itself is based on a set of program descriptions that automate their use, from the construction of the web interface up to the software execution. These descriptions also include a syntactic and semantic description of the processed biological data, which is the technical basis for the advanced features such as:

- **Data reusability**: the preservation of user data and parameters facilitates the reuse of input values or results between different programs.
- **Automatic data validation and format conversion**: the description of the expected data and their format allows the system to verify and convert input values if necessary.
- **Service discovery and workflow authoring assistance**: services are provided through a searchable menu; furthermore, data type compatibility mechanisms between results and potential program inputs let users either interactively pipe tasks or build complete workflows to run them.

These various features have been implemented within an architecture that consists of:

- **The Mobyle server**, based on a set of python modules, handles the various aspects of job, data and session managements. The most critical element, job submission, includes: various parameter validations, command line building, and the interaction with an execution system.
- **The Web portal** provides a unified access to the system. We use Ajax to coordinate the different types of information that need to be available and to enable the user to explore many functionalities in a single page.
- **The Storage System**, this includes job submissions, as well as the most important data of a user's workspace. The details of the job execution ensure its traceability, while the semantic tagging of the data improve their reusability.

Future work is planned to provide additional functionalities:

- A unified access to multiple mobyle servers, from one portal: **the Mobyle Grid**. This will facilitate the interoperation (chaining, data reuse) of distributed resources,
- The possibility to build and execute **Workflows** composed of Mobyle jobs from the Web Portal,
- The automated publication of mobyle programs as **Web Services** (BioMoby, SOAP). Some work has already been done in that direction with the PlayMOBY project¹.

Written in python, Mobyle should be compatible with any Unix-like system, requires an Apache web server, and has been tested with various job queueing systems (local batch, Torque/PBS, SGE). It is distributed under the terms of the GNU GPLv2. Two servers already provide access to the system: <http://mobyle.pasteur.fr> and <http://bioserv.rpbs.jussieu.fr/cgi-bin/MobylePortal/portal.py>.

¹PlayMOBY is an independent project from the INRA LIPM bioinformatics team. You can find more information about it on <http://lipm-bioinfo.toulouse.inra.fr/biomoby/playmoby/>