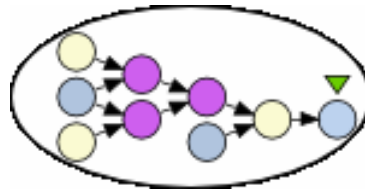# A Workflow Enactment Portal for Bioinformatics

Paolo Romano
Bioinformatics and Structural Proteomics
National Cancer Research Institute of Genoa, Italy
paolo.romano@istge.it

Rationale and goals

Moving from an interactive to an automated approach for data integration requires new technologies and tools.

Some starting assumptions
- XML schemas for the creation of the models of the information,
- XML based languages for data representation and exchange,
- Web Services for the interoperability of software
- Computerised workflows for the definition and execution of analysis processes

Many workflow management software for bioinformatics applications are being proposed, but they all require knowledge of the systems and skills and time for development of the workflows.

Features (i)

We designed biowep, a worlflow enactment portal that:

- allows for the carrying out of predefined workflows
- supports workflows annotation by using a simple ontology for bioinformatics processors (domain, task, i/o)
- implements search and selection of workflows on the basis of their annotation
- supports retrieval of workflows on the basis of users' profiling
- allows storing and retrieval of workflows' executions and related results
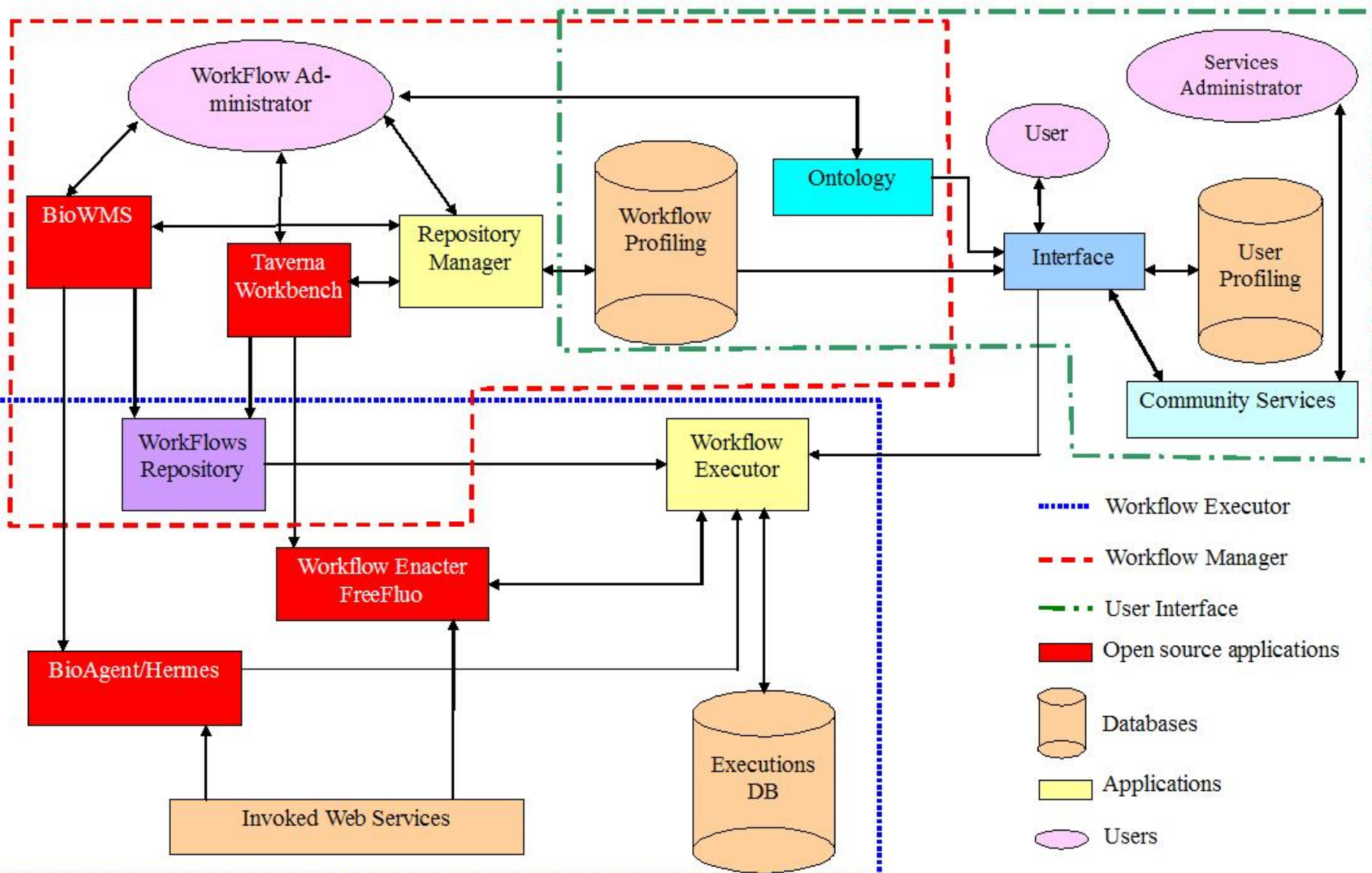
LITBIO

Features (ii)

The system:

- is partially based on open source (Taverna WB, FreeFluo e mySQL)
- makes access to data and analysis that are made available through Web Services
- archives workflow either in XScufl o XPDL formats
- is available under LGPL license
- requirements: Java SDK + Tomcat + mySQL

Available on-line:

http://bioinformatics.istge.it/biowep/ (reference site)
http://bioinformatics.istge.it:8080/biowep/ (portal)

LITBIO

Overall architecture

Workflows

Workflows are:

- created by the admin by using Taverna or BioWMS

- archived in the related format in the repository

- updated, when needed (workflow vs version)

- can be submitted by users or service providers through the reference site

LITBIO

Workflows' annotation

Workflows and their main steps are both annotated by using a simple ontology of bioinformatics data and tasks including

- application domains
- main task
- input and output data

Users' profiling

Users are classified on the basis of:

- their role in the organization
  - computer scientist / physician / researcher / patient / officer / …
- domains of interest

Workflows are listed according to user profiling and previous executions

Planned activities

- End users support (two implementations)
  - manuals,
  - support through biowep@istge.it
- Bug fixing (curation of children diseases)
- Revision of applications:
  - administration (end of summer / begin of autumn)
  - visualization of results (end of 2006 / begin of 2007)
- Inclusion of further WMS
  - analysis and selection undergoing

LITBIO