Enacting Taverna Workflows in Galaxy

Dr. Konstantinos Karasavvas

Netherlands Bioinformatics Centre

Leiden University Medical Center

Introduction

- Bioinformatics tools
  - large number and diversity
  - more than one tool for the job
  - different access mechanisms
  - non-interoperable

- Experiments need multiple tools

- Need to aggregate tools together
  - specify execution ordering
  - handle tools I/Os
  - pipelines, workflows, …

Netherlands
Bioinformatics
Centre

Some Background: Galaxy (1)

- Web portal and framework for bioinformatics
  - data sources (UCSC, BioMart, …) and analysis tools
  - new tools
  - unified intuitive interface
  - other: pipelining, provenance, sharing, open source

Netherlands
Bioinformatics
Centre

Some Background: Galaxy (2)

Some Background: Taverna (1)

- Workflow management system
    - for bioinformatics resources (WS) and other domains
    - flexible and expressive workflow language
    - several extensibility points (e.g. shims and plugins)
    - other: provenance, open source

Some Background: Taverna Workbench

Some Background: myExperiment (1)

- Scientific social networking web site
  - sharing and re-use of scientific experiments
  - workflow repository for sharing and reuse

Some Background: myExperiment (2)

Taverna Workflows in Galaxy: motivation

- Galaxy
  - easy to use for biologists
  - e.g. NGS analysis, ...
- Taverna
  - powerful expressive workflows
  - e.g. text mining, ...
- active communities that add new functionality
  - some overlap but added value if combined
- approaches
  - incorporating Galaxy tools in Taverna
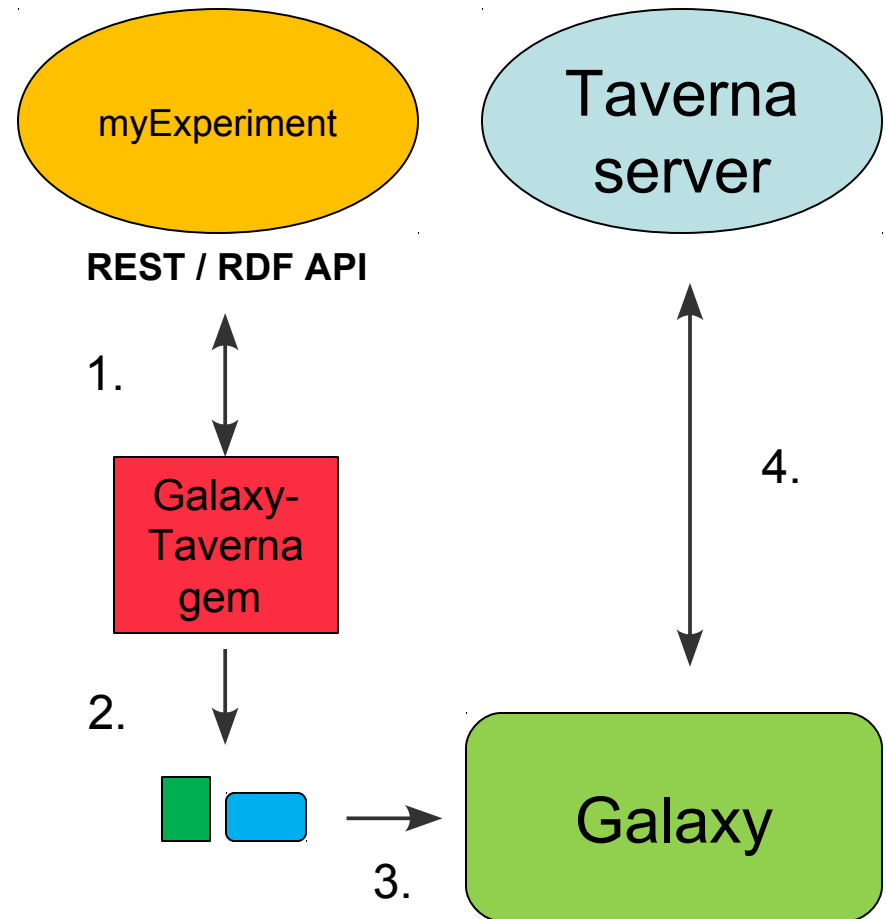 - incorporating Taverna workflows in Galaxy

Taverna Workflows in Galaxy: requirements

- Taverna
  - a server (access)
- myExperiment
  - browse workflows
- Galaxy
  - a server (admin)
  - tool to connect to Taverna
- Galaxy-Taverna component
  - GUI plus config file
  - program (binary or script)

Taverna

myExperiment

?
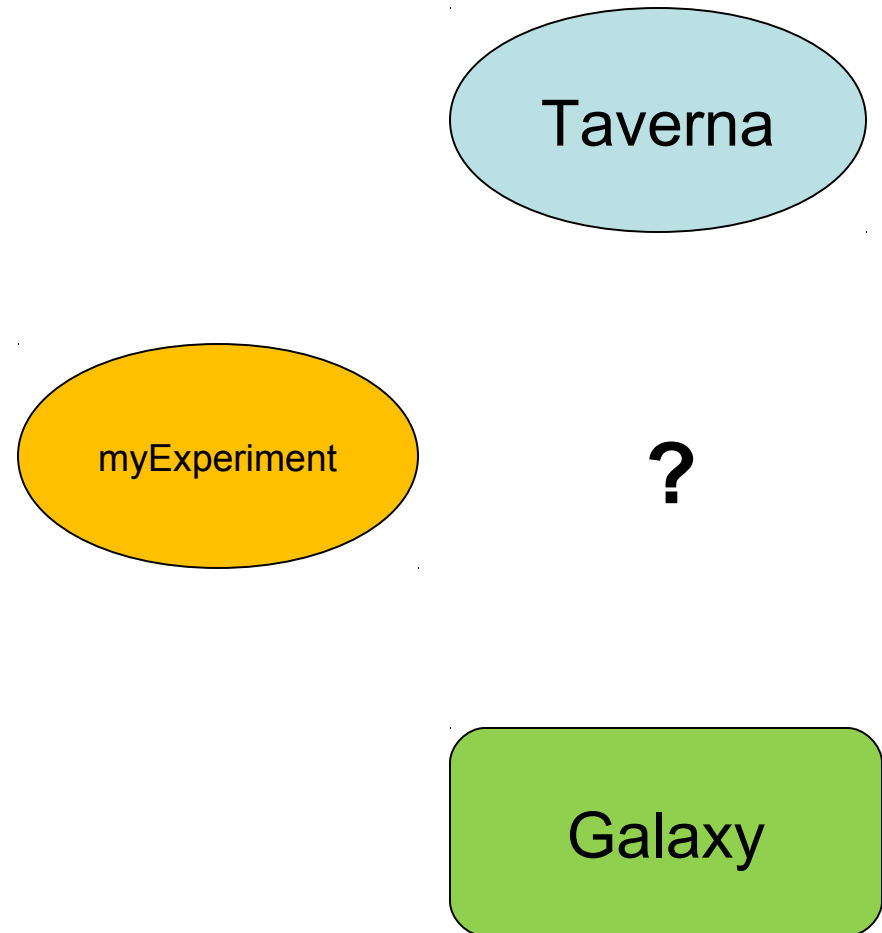
Galaxy

Taverna Workflows in Galaxy: phase one

- Galaxy-Taverna component
  - ruby gem
    - behind the scenes
  - generates a Galaxy tool
  - requires a workflow description
- Workflow description
  - myExperiment
  - why?
- Galaxy
  - tool needs to be manually installed

myExperiment

Taverna server

REST / RDF API

1.

Galaxy-Taverna gem

4.

2.

3.

Galaxy

LU MC

nbic Netherlands Bioinformatics Centre

myExperiment

Download Workflow as a Galaxy Tool

Taverna Workflows in Galaxy: requirements

- Galaxy server
  - admin access
  - Ruby environment and two gems
    - $ sudo apt-get install ruby rubygems ruby1.8-dev libxml2-dev
    - $ sudo gem install t2-server --version 0.5.3
    - $ sudo gem install rubyzip
- Taverna server
  - for testing, a taverna server is provided
- Please test (www.myexperiment.org)
  - trivial if you already have Galaxy installed
  - useful taverna workflows for your pipelines
    - I would love to know

LU
MC

nbic
Netherlands
Bioinformatics
Centre

Taverna Workflows in Galaxy: phase two

- Future work

- Galaxy
  - integrate myExperiment as an external display app.

- Galaxy new functionality
  - dynamic loading of new tools
  - selected workflows will be generated and installed in one step

- Further investigation is needed

- More information
  - https://trac.nbic.nl/elabfactory/wiki/eGalaxy
  - http://galaxy.psu.edu/
  - http://www.taverna.org.uk/
  - http://www.myexperiment.org/
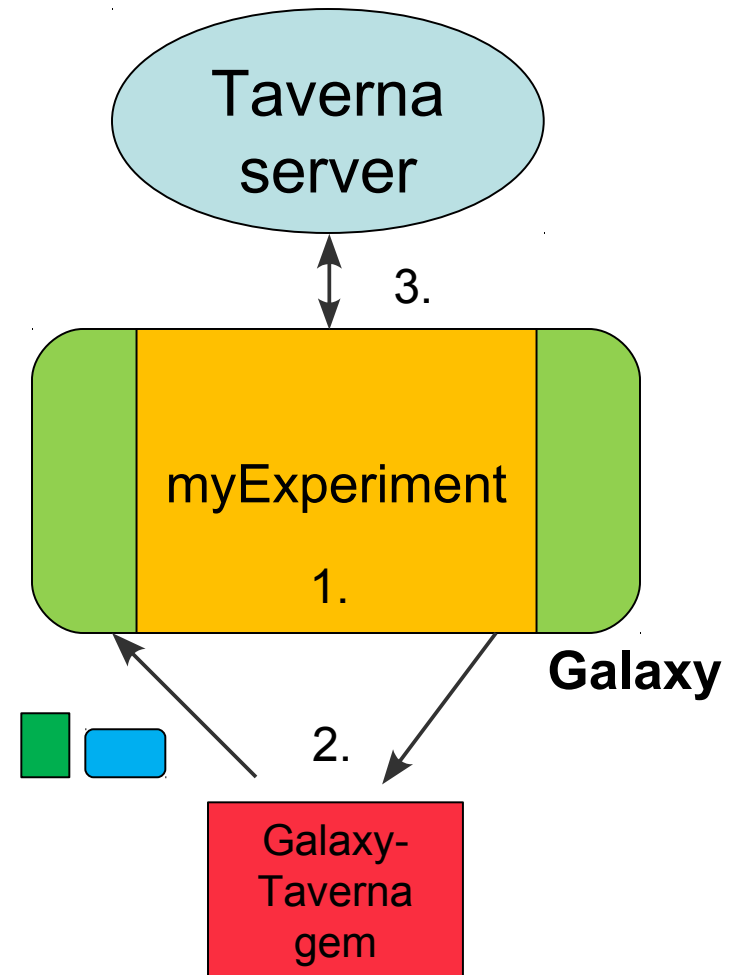
- Questions?
  - kostas.karasavvas@nbic.nl

nbic
Netherlands
Bioinformatics
Centre