Preparing for Taverna using SOAP

Graphical Interface to SOAP based Web Services

Francisco Simões Roque
Center for Biological Sequence Analysis
Technical University of Denmark

EMBRACE Workshop on Bioinformatics of Immunology
January 25th 2007
What is Taverna?

- A workflow development and enactment environment
- Part of myGrid project
- Collectively aimed at facilitating standard scientific procedures in silico, especially in workflow systems
Other approaches

- Cut and paste, perl, shell scripting, excel
- Time intensive
- Manual process, hard to scale
- Hard to document and reproduce
- Steep learning curve
Taverna’s approach

- Capture the scientific method as a formal process model
- Users can construct models from libraries of available components in a graphical editing environment
- Publish process definitions as scientific methods, enact and automatically scale to large data sets, multiple runs
Taverna Workbench
Taverna Workbench

- Powerful and at least partially user friendly tool for building or editing workflows.
- Allows editing and saving workflows for publishing and sharing.
- New services can be gathered from anywhere in the web.
- The processes for running and invoking a single service are the basics for any workflow.
- Generation of results are the same however complicated a workflow becomes.
Service Discovery

![Image of service discovery interface with available processors and options]

- Add new API Consumer...
- Add new Biomart service...
- Add new Biomoby scavenger...
- Add new Soaplab scavenger...
- Add new Styx Grid Service...
- Add new Talisman scavenger...
- Add new WSDL scavenger...
- Collect scavengers from model
- Expand all
- Collapse all
Taverna Features

- Taverna enacts workflows and adds further value over this functionality.
  - Implicit iteration support
  - Result browsing and data encapsulation
  - Provenance recording
  - Fault tolerance features
Implicit Iteration

• Allows services to consume collections of items without service modification
• Equivalent to higher order map functions
• Graphical configuration
• Intuitively understood by the users
Data Encapsulation

- Data can be tagged with terms from ontologies, free text descriptions and MIME types.
- Metadata tags can be used to locate and launch pluggable view components.
- Contextualization of data in life science types.
Example Result Browser
Provenance

• Providing computation access to services creates new challenges, workflow technology amplifies them further.
• Potentially complex result data.
• Scientists need to be able to show how a given result in the data is arrived at.
• Metadata in the results is as important as the result values.
Intermediate Results
Fault Tolerance

• In an open service world, we have no control over the majority of analysis services.
• Such services may fail, become inaccessible or their APIs change with no notice.
• Taverna allows configurable failure handling including dynamically rescheduling processors with alternate implementations.
Time for demonstration...
Summary - Taverna and SOAP

• Standard workflow language for accessing Web Services.
• Integration with virtually any SOAP based service.
• Results can be saved for universal access.
• Result context explained and guidance for further investigations is given.
Exercise Workflow

• The workflows for the RNAmmer example can be downloaded from: