Taverna Introduction

Aleksandra Pawlik
myGrid Team
University of Manchester

VLIZ, 2014-10-06 / 2014-10-08
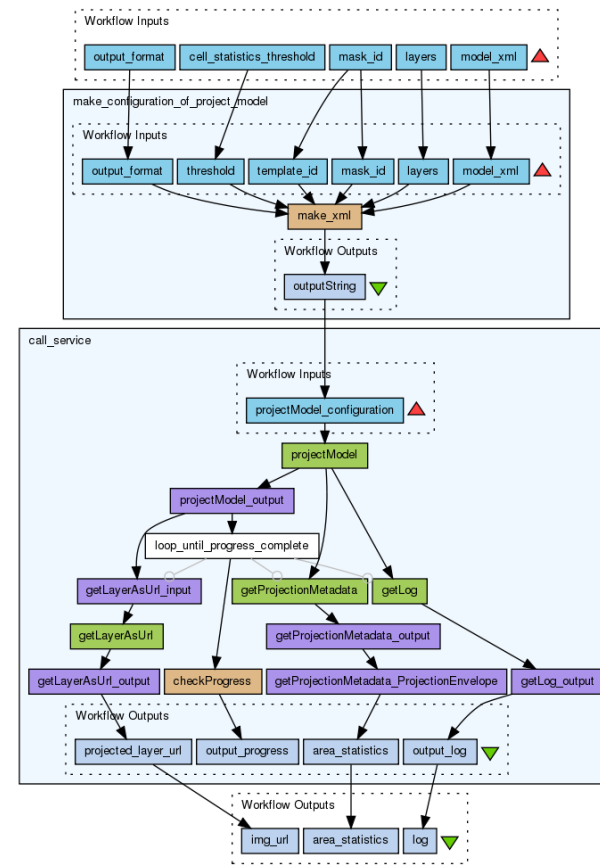http://www.taverna.org.uk/

Workflows

- Sophisticated analysis pipelines

- A set of services to analyse or manage data (either local or remote)

- Data flow through services

- Control of service invocation

- Iteration

- Automation

- Access to intermediate results

http://www.myexperiment.org/workflows/3966

Workflows in Taverna

- This tutorial will give you a basic introduction to designing, and reusing workflows in Taverna and some of its main features.

- Workflows in this practical use small data-sets and are designed to run in a few minutes. In the real world, you would be using larger data sets and workflows would typically run for longer

Taverna: Download and installation

- Taverna for biodiversity can be downloaded from

  http://www.taverna.org.uk/download/workbench/2-5/biodiversity/

- Follow the instructions on the website to install Taverna for your operating system.

  The following page shows a screenshot of Taverna and the different panels that make up the workbench

Design perspective

- The **Design perspective** is where you create and edit workflows

Workflow Diagram

The workflow diagram is the visual representation of the workflow, it:

- Shows inputs, outputs, services and data flows

- Allows editing of the workflow by dragging and dropping and connecting services together

- Enables saving of workflow diagrams for publishing and sharing
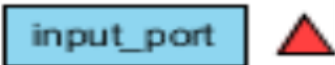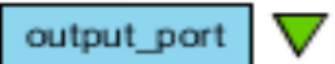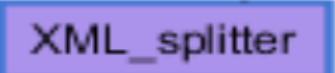
Workflow Diagram - Service types

Type	Description
BeanShell	A user editable scripting operation using Beanshell script
Nested workflow	A sub-workflow exposed as a single operation
RShell	A user editable script in R
Spreadsheet import	An operation for loading spreadsheet data in CSV and Excel format
Text constant	A single constant string
XPath	Extraction of data from XML data

Type	Description
Local service	A standard Beanshell
WSDL web service	A processor accessing a standard SOAP service
REST web service	Accessing RESTful web services
Tool	Calling a local or remote tool/script
Interaction	Interaction in a web browser
Component	Re-usable component workflow

Workflow diagram – non services

- Taverna workflow diagrams also have:

Element	Description
input_port ▲	Input port
output_port ▼	Output port
(merge circle)	"Merge" - make a list out of elements (often sub-lists)
XML_splitter	XML Splitter
(data flow arrow)	Data flow
(control flow symbol)	Control flow (run after)

Workflow Explorer

- The Workflow Explorer shows the detailed view of your workflow. It shows default values and descriptions for service inputs and outputs and it shows where remote services are located. It also shows configuration details, such as iteration and looping

- Workflow validation details can also be found here. Before a workflow is run, Taverna checks to see if it is connected correctly and if its services are available.

Available Services Panel

Lists services available by default in Taverna

- Local java services
- WSDL Web Service – secure and public
- RESTful Services
- R Processor services (for statistical analyses)
- Beanshell scripts
- Xpath scripts
- Spreadsheet import service

The services panel also allows you to add new services or workflows from the web or from file systems – there are loads more available!

Results perspective

- The **Results perspective** shows
  - your current and previous workflow runs,
  - the data associated with the runs

- You can view the run's input and output values and also the data that flowed between services

- The data can be saved

myExperiment perspective

- The **myExperiment perspective** gives you access to some of the capabilities of the myExperiment website

- The **Starter Pack** provides useful examples for you to try

- You can also **Search** for workflows to **Open** and run
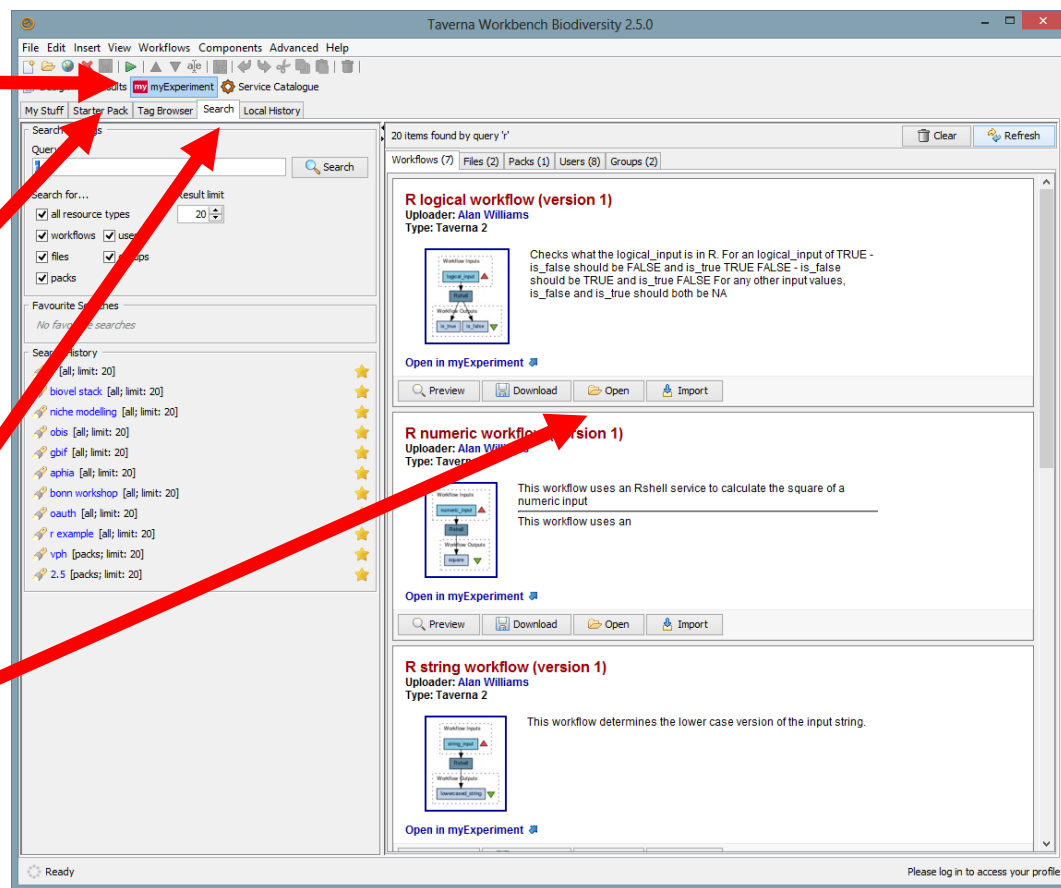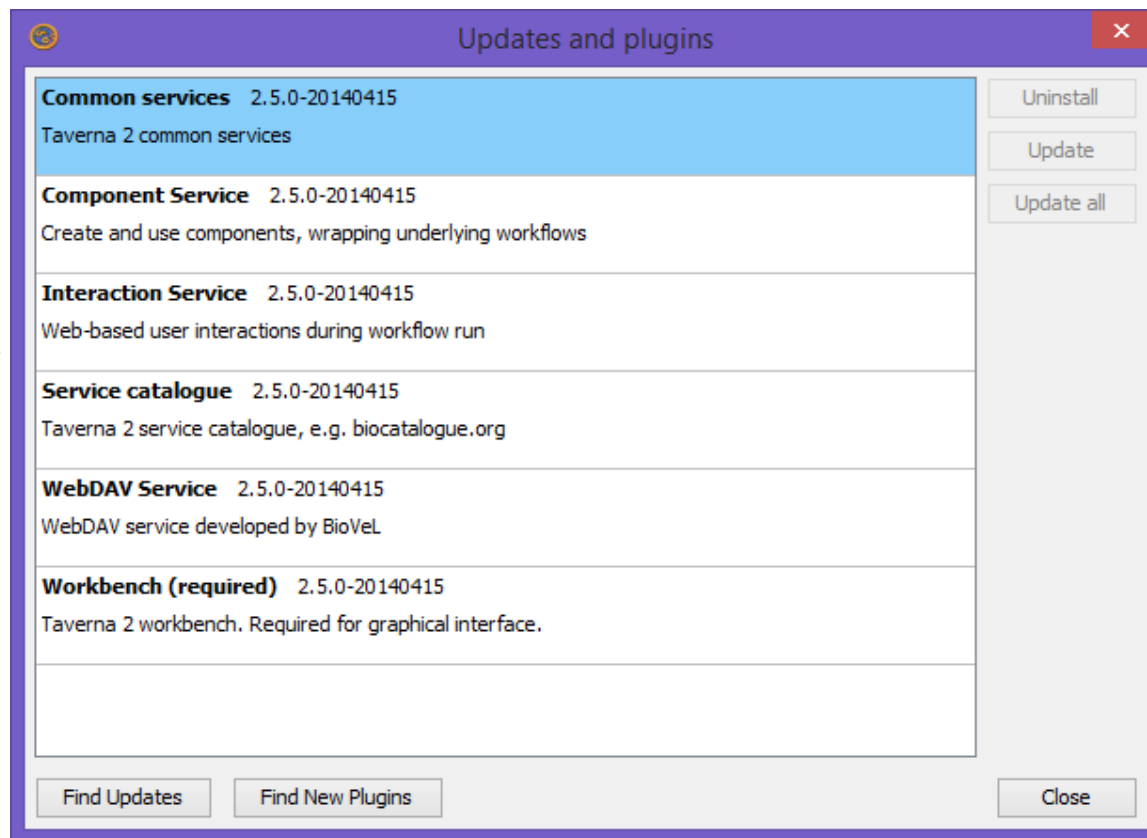
Service Catalogue perspective

- The **Service Catalogue perspective** gives you access to some of the capabilities of the biodiversityCatalogue website

- You can search for services

- Check their status

- Add them to the workflow or the service panel – available by right-clicking the service
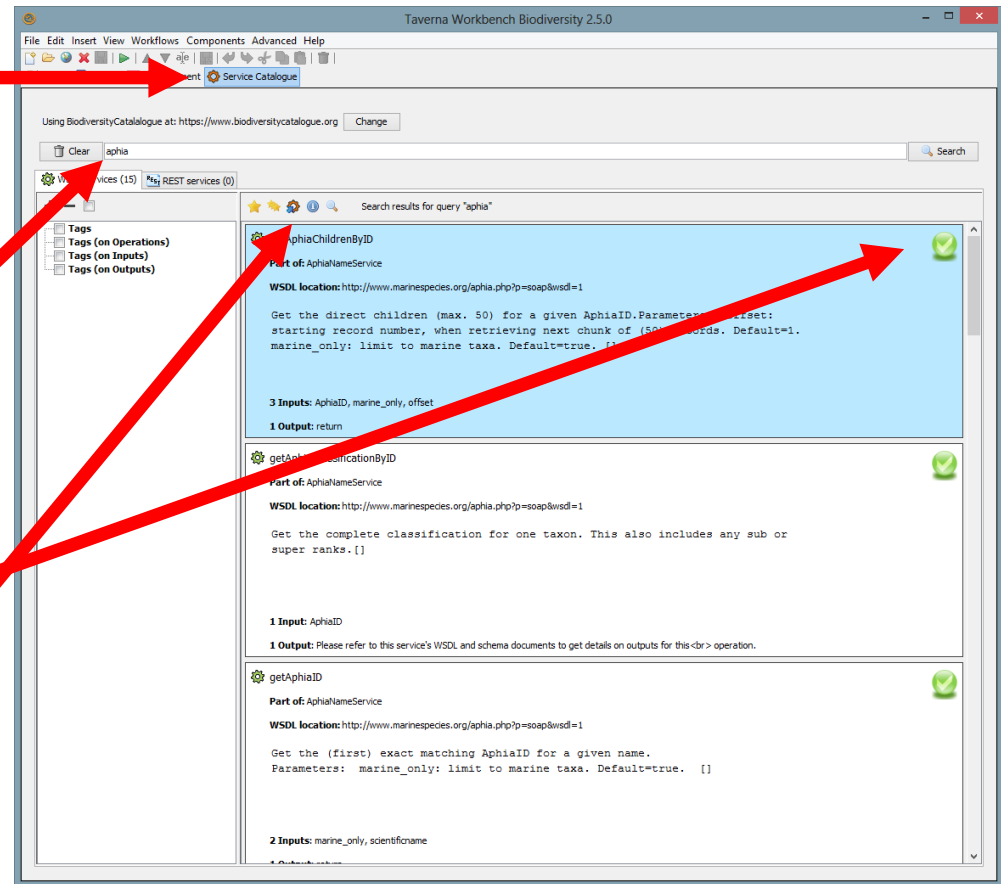
Updates and Plugin Installation

- Taverna updates are issued on the regular basis

- There is also a number of plugins which are developed for Taverna

- To get the updates and plugins select "Advanced -> Updates and plugins"

Updates and plugins

Common services   2.5.0-20140415
Taverna 2 common services

Component Service   2.5.0-20140415
Create and use components, wrapping underlying workflows

Interaction Service   2.5.0-20140415
Web-based user interactions during workflow run

Service catalogue   2.5.0-20140415
Taverna 2 service catalogue, e.g. biocatalogue.org

WebDAV Service   2.5.0-20140415
WebDAV service developed by BioVeL

Workbench (required)   2.5.0-20140415
Taverna 2 workbench. Required for graphical interface.

Uninstall
Update
Update all

Find Updates   Find New Plugins   Close

Taverna Workflow Management
Open extensibility

- Plug-in framework
  - Command line tool
  - Data Services: VOTables for AstroTaverna
  - Optimisations: E.g. Holl. model parameter sweeps
  - Infrastructures: Grid, HPC, Web Services
  - Domains: CDK, BioMart, VOTable
  - Commodities: Excel Spreadsheets, Open Refine, R

- Plug into other frameworks & platforms
  - Portals: Scratchpads
  - Interactive platforms: iPython Notebook
  - Wfms: KNIME Node, Galaxy tool, Kepler Actor

- Third party applications
  - Taverna Online
  - XworX
  - OGC chainer