

Taverna Tutorial

Building a simple workflow

Stian Soiland-Reyes and Christian Brenninkmeijer
University of Manchester
materials by Katy Wolstencroft and Aleksandra Pawlik

<http://orcid.org/0000-0001-9842-9718>

<http://orcid.org/0000-0002-2937-7819>

<http://orcid.org/0000-0002-1279-5133>

<http://orcid.org/0000-0001-8418-6735>



*This work is licensed under a
[Creative Commons Attribution 3.0 Unported License](http://creativecommons.org/licenses/by/3.0/)*

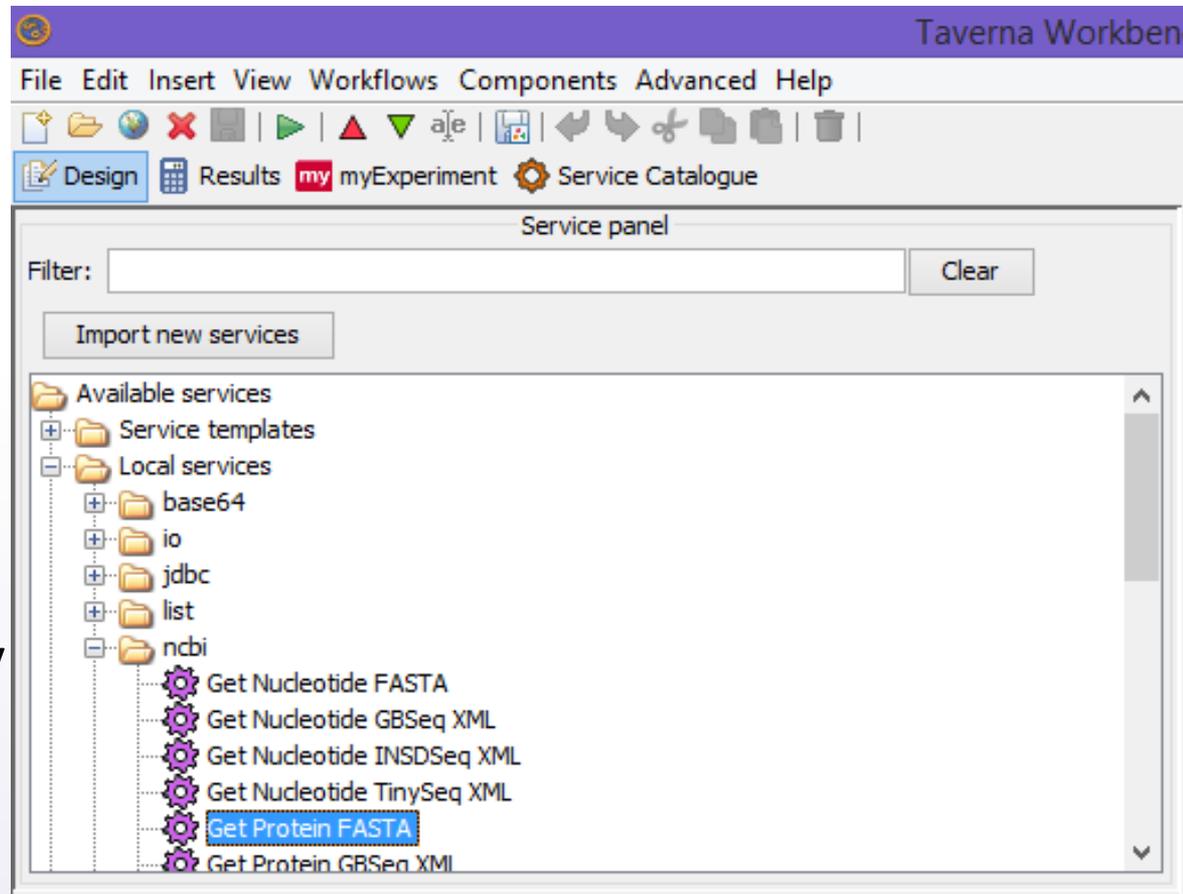
Bonn University, 2014-09-01
<http://www.taverna.org.uk/>



Available Service

We will start with something easy - retrieving a protein sequence from a remote database and identifying functional motifs

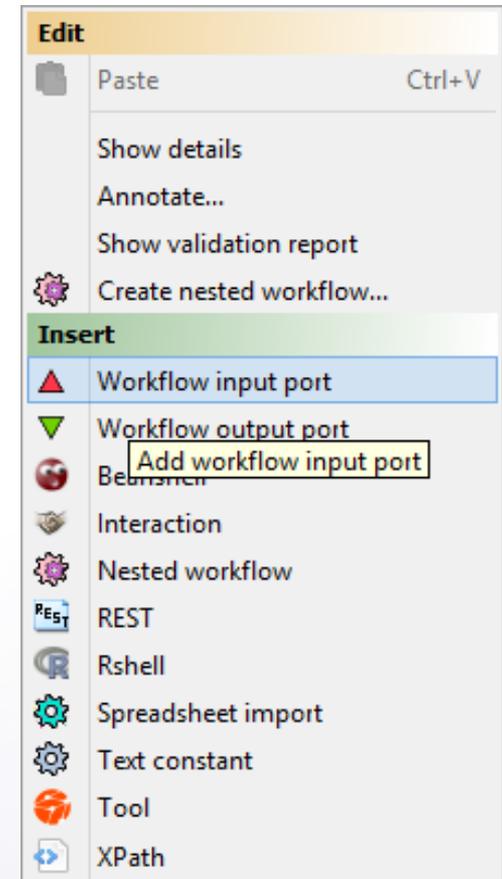
- Expand the list
- Available Services
 - Local Services
 - ncbi
- Select 'Get Protein FASTA' and drag-and-drop it into the empty workflow diagram





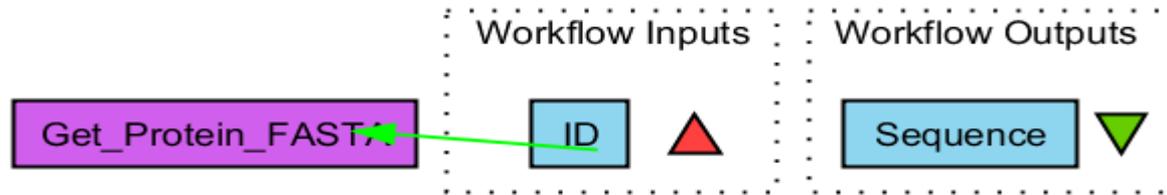
Workflow input/output ports

- In a blank space in the workflow diagram, **Right-click** and select **Workflow input port** from the *Insert* section
- Type in a name for this input (e.g. ID) and click **ok**
- Do the same to create a new **Workflow Output Port**. Call this output Sequence





Connecting ports

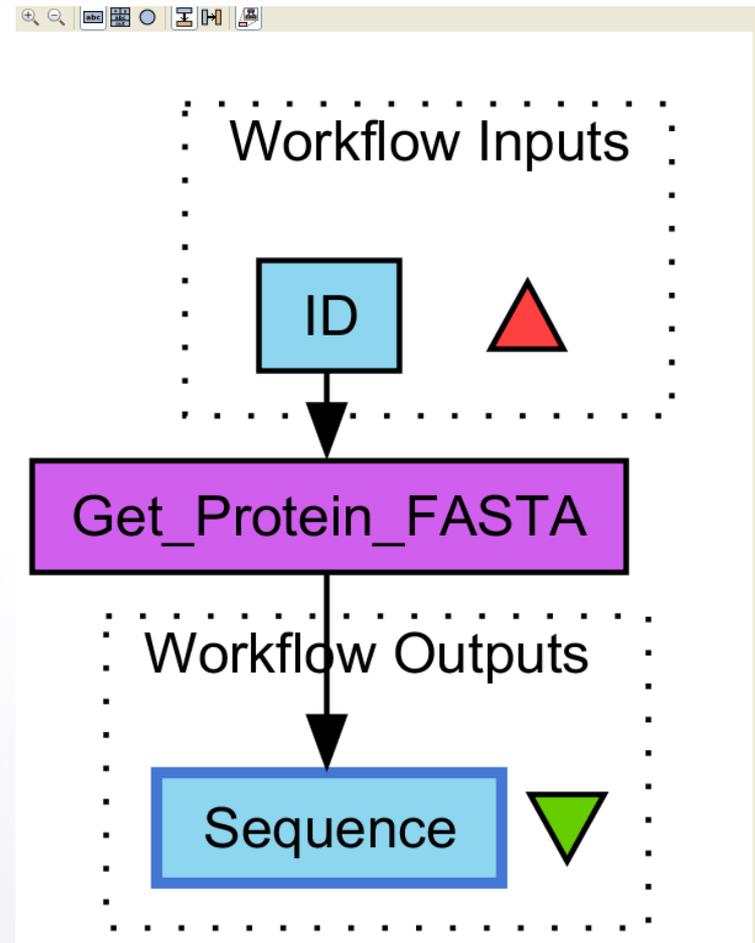


- You now have 3 boxes in the diagram and we need to connect them up to build our workflow
- Click on the input box *ID* and drag towards *Get_Protein_Fasta* and let go. An arrow will connect the two boxes.



Your first workflow

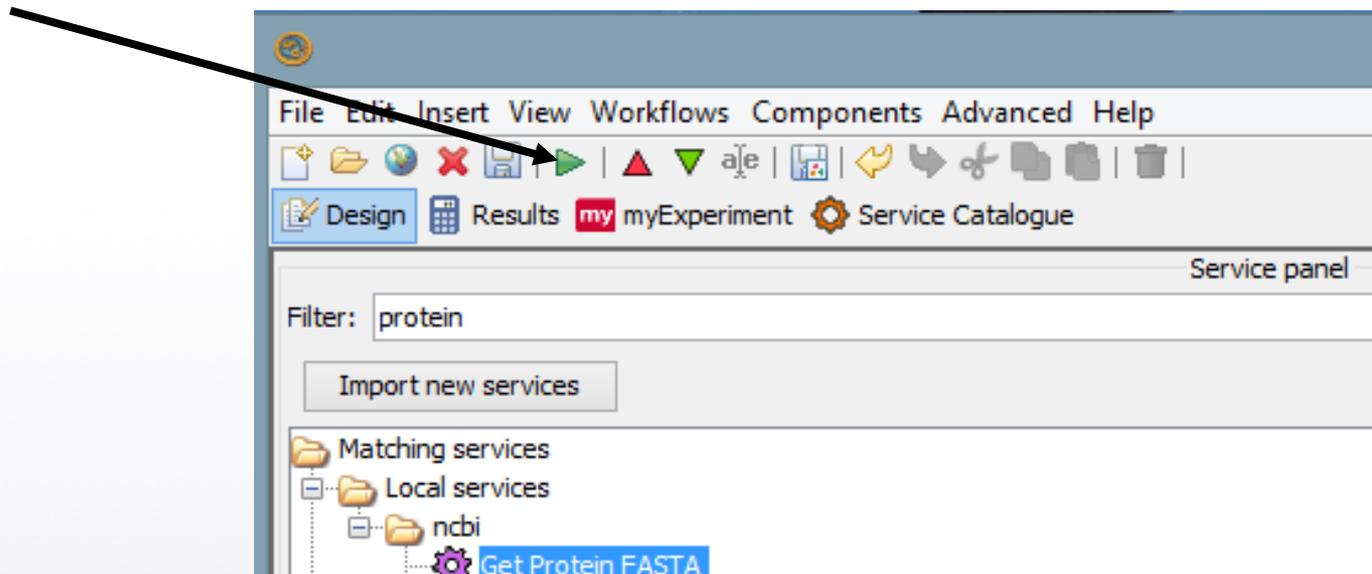
- Click on the output box *Sequence*, drag towards “*Get_protein_fasta*”, and let go. An arrow will connect the two boxes.
- You have now built your first workflow!





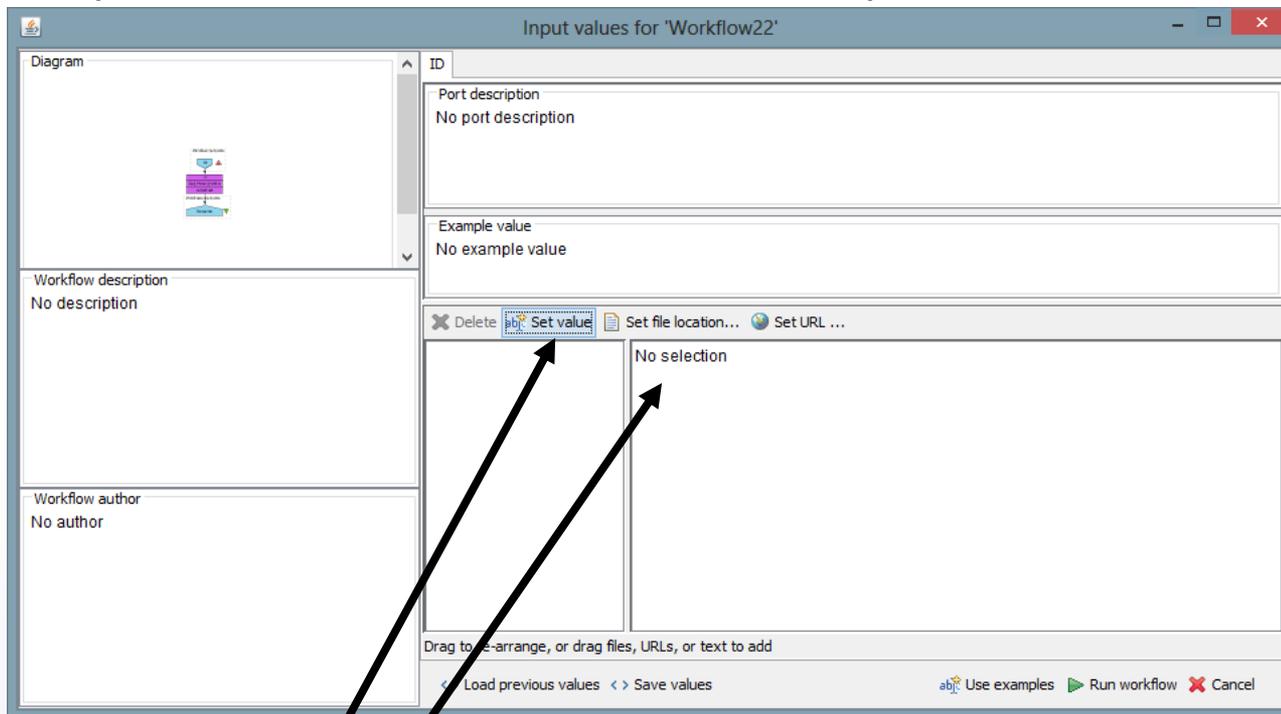
Running workflow

- In the menu, select “**File -> Run workflow**”, or click on the green play button at the top of the workbench



Providing workflow inputs

An input window will appear. We have not yet added a description of the workflow or the input.



Click on '**Set Value**' in the input window and add a Uniprot protein identifier (e.g. P15409) where it says "*some input data goes here*"



Workflow results

- Click **Run workflow**
- The workbench changes to the **Results** perspective
- In the bottom left, click on **Value 1**
- You will now see a protein sequence from Uniprot

The screenshot shows the 'Workflow results' window. At the top right, there are three buttons: 'Finished' (with a green checkmark), 'Pause' (with a pause icon), and 'Cancel' (with an 'X' icon). Below these is a tree view with a single node labeled 'Value 1'. To the right of the tree view, there is a 'Value type' dropdown menu set to 'Text', a 'Refresh' button, and a 'Wrap text' checkbox. The main area of the window displays a protein sequence in a monospaced font:

```
>gi|341941210|sp|P15409.2|OPSD_MOUSE RecName: Full=Rhodopsin
MNGTEGPNFYVFPFSNVIGVVRSPFEQPQYYLAEPWQFSLAAYMFLLIIVLGFPINFLT
PLNYIILLNLAVADLFMVFGGFTTTLTYTSLHG YFVFGPTGCNLEGGFFATLGGEIALWSLV
KPM SNFRFGENHAIMGVVFTWIMALACAAPPLVGWSRYIPEGMQCSCGIDYYTLKPEVN
HFTIPMIVIFFCYGQLVFTVKEAAAQQQESATTQKAEKEVTRMVIIMVIFFLICWLPYA
SNFGPIFMTLPFAFAKSSSIYNPVIYIMLNKQFRNCMLTTLCCGKNPLGDDDDASATASK
```



Validate your Workflow

- Taverna can check to see that everything is connected properly and that all the services in your workflow are available
- Go to the workflow explorer (“**Design**” button) and click on ‘**Validation report**’ tab
- See if Taverna has found any problems with the workflow. Errors will be displayed in red, warnings in yellow. Workflows with warnings often still run.
- If there are problems, follow the instructions to resolve them by clicking on the ‘**Solution**’ tab
- Are you able to create a workflow that gives warnings or errors?

Tip: Try deleting the data link to the workflow output port