Using BiodiversityCatalogue

Aleksandra Pawlik
myGrid Team
University of Manchester

VLIZ, 2014-10-06 / 2014-10-08
http://www.taverna.org.uk/
Building a biodiversity Workflow

• We will start with something easy – we will use a GBIF service to retrieve information about the occurrences of a species which name we will provide

• Go to the www.biodiversitycatalogue.org and search for “gbif”
Service selection

- From the results select *GBIF Occurrence Web Service*
Service description

• Have a look at the service description

<table>
<thead>
<tr>
<th>Overview</th>
<th>REST Endpoints (6)</th>
<th>Examples</th>
<th>Monitoring</th>
<th>History</th>
</tr>
</thead>
</table>

Provider: data.gbif.org

Location: not available

Submitter / Source: Robert Haines (Admin) (5 days ago)

Base URL: http://data.gbif.org/ws/rest/occurrence

Documentation URL(s):
http://data.gbif.org/ws/

by Robert Haines (Admin) (5 days ago)

Description(s):
This service provides a range of filters for selecting occurrence records. The currently supported response formats include TDWG Darwin Core records and KML (for use with Google Earth).
Examples

• Select the **Examples** tab and see how the service can be used
Using a REST template

- We want the service to return the results for a species which name we will provide in darwin format, that have coordinates included and we want to limit the number of results to 100 so our REST configuration will look like this:

  http://data.gbif.org/ws/rest/occurrence/list?scientificname={scientificname}&format=darwin&coordinatestatus=true&maxresults=100

- The `{scientificname}` means the REST service has a parameter called `scientificname`. Within Taverna, the parameter value can be passed into the service call.

- In Taverna Workbench go to the Services Panel
- From the Available Services select Services Template and REST
- Right-click on it as select Add to workflow (see the next slide)
REST template selection

- Beanshell – A service that allows Beanshell scripts, with dependencies on libraries
- Interaction
- Nested workflow – A service that allows you to have one workflow nested within another
- REST Service – A generic REST service that can handle all HTTP methods
- Rshell – A service that allows the calling of R scripts on an R server
- SpreadsheetImport – A service that imports data from spreadsheets
- Text constant – A string value that you can set
- Tool – A service that allows tools to be used as services
- XPath Service – Service for point-and-click creation of XPath expressions for XML data

Workflow explorer | Details | Validation report
• Enter the following into the URL template field:

http://data.gbif.org/ws/rest/occurrence/list?scientificname={scientificname}&format=darwin&coordinatestatus=true&maxresults=100

• Click **Apply** and **Close** to save the configuration
Service renaming

- Let’s change the name of the service to: gbifLocatedOccurrenceInDarwin
At the top of the workflow diagram panel, change the view to show all ports by clicking on the icon shown below:

- This view allows you to see any data input/output or parameter value options for your chosen service.
- The REST service should have an input port called `scientificname`.
Workflow port creation

• 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. sciName) and click “ok”

• Do the same to create a new workflow output. Call this output “locatedOccurences”
Workflow connection

• Connect the input and output ports

• Your workflow should look like this
Running the workflow

- Run the workflow by selecting “file -> run workflow”, or by clicking on the play button at the top of the workbench.
Setting input values

• You’ll get a pop up window where you can enter the data for the workflow. Select “Set value”
Setting input values

- Click “Select value” and enter “Marmota marmota” and then at the bottom of the window “Run workflow”
You should see the workflow running
• Once the workflow finished running click on “Value” to see the results
Saving a workflow

• Let’s save the workflow now as “Species_Occurrence”
Service ports

- Most of the time, you don’t need to connect all ports. Some are optional and some already have default values set.
- Service documentation should tell you this. You can use the BiodiversityCatalogue to find documentation and user descriptions.
- Change the orientation of the port names to fit them on the screen more easily by clicking on the icon shown below.
Adding a Workflow Description

• Right-click on a blank part of the workflow diagram and select “Annotate”

• Add some details about the workflow e.g. who is the author, what does it do

• You can also add examples and descriptions for the workflow inputs by selecting them and selecting “Annotate”

• Add an example for the species “Marmota maromta”

• Save the workflow by going to “File -> save workflow”

• Run the workflow again and look at the results