



XPath service

Aleksandra Pawlik
myGrid Team
University of Manchester

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



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





The Basics of XML

- XML – eXtensible Markup Language
- Designed for the storage and transport of data
 - This includes passing data between services or retrieving data from a Web page
- Provides a machine readable dataset
- Many service providers export data in XML



XML Example

```
<?xml version="1.0"?>
```

```
<note>
```

```
  <to>Katy</to>
```

```
  <from>Helen</from>
```

```
  <heading>Reminder</heading>
```

```
  <body>Don't forget about Bonn Trip!</body>
```

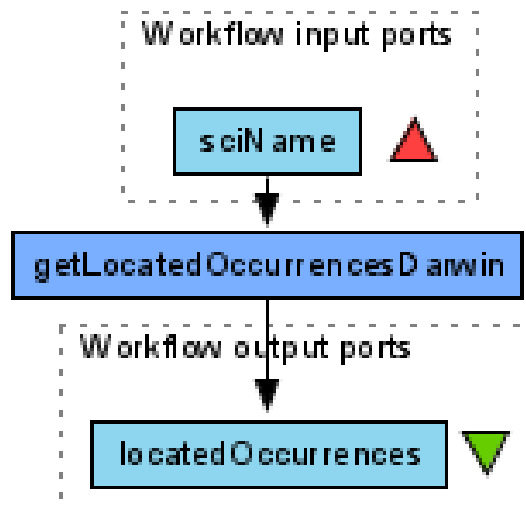
```
</note>
```

- The following website has lots of information about XML and tutorials: <http://www.w3schools.com/xml>



Extracting data from XML

- Open the *Species_Occurrence* workflow
 - See: <http://www.myexperiment.org/workflows/4487>





Extracting data from XML

- As the input value enter *Marmota marmota*
- Run the workflow
- Look at the output – it is XML data
- Save the output as occurrences.xml



Extracting XML Elements

- Often, we need to extract particular elements of the XML to feed into the next services in the workflow
- Taverna has a dedicated service to help you do this
- Go to the **Service panel** and find **XPath Service** in the **Service Templates** section.
- We will use this service to explore the results of the *getLocatedOccurrencesDarwin* service



Extracting data from XML

- Now add an **XPath Service** to the workflow
- In the pop-up window, click **Load XML from file** and browse to the occurrences file and click **Open**
- Click the green arrow to load it into the service
- By expanding the + icons, you can explore the file
- Expand the XML under *gbif:dataProviders*, until you reach *gbif:name*
- Click *gbif:name*



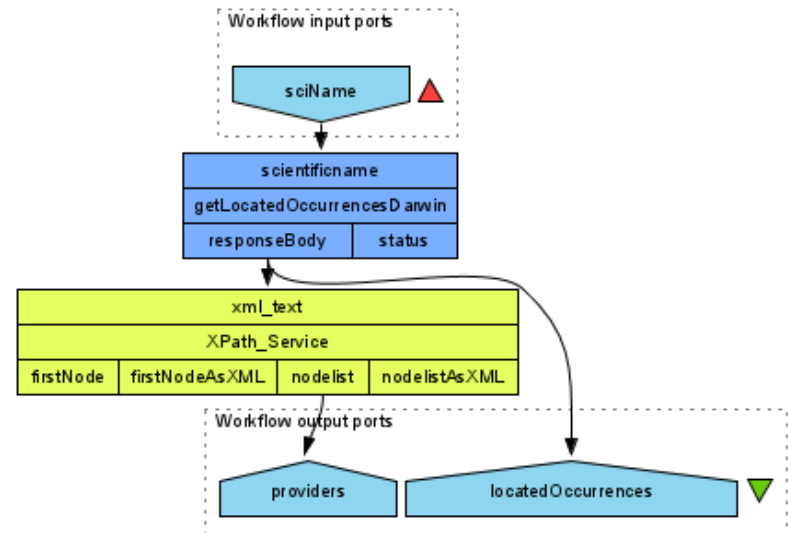
Extracting data from XML

- Click on **Generate Xpath expression**
- You can check and edit the XPath expression if you want
- Then click on **Run Xpath**
- Look at the predicted results
- Click **Apply** and **Close**
- This service will now pull out all the names of each provider



Extracting data from XML

- Connect the input of the *XPath Service* to the *responseBody* output of *getLocatedOccurrencesDarwin*
- Add an output port to the workflow
- Connect the workflow output port to the *odelist* port of the *XPath Service*
- Save and run the workflow again





Summary

- You can now configure XPath services from example XML
- You can select parts of an XML to extract