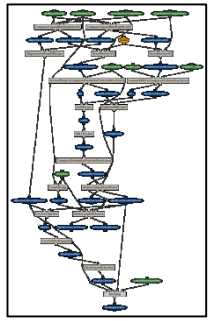
Scientific workflow publication and mining

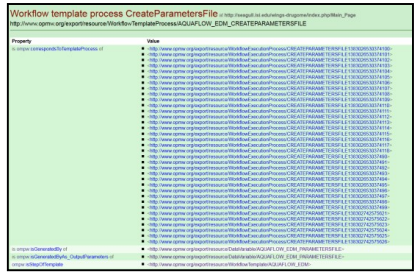
Daniel Garijo, with Yolanda Gil and Oscar Corcho (UPM)

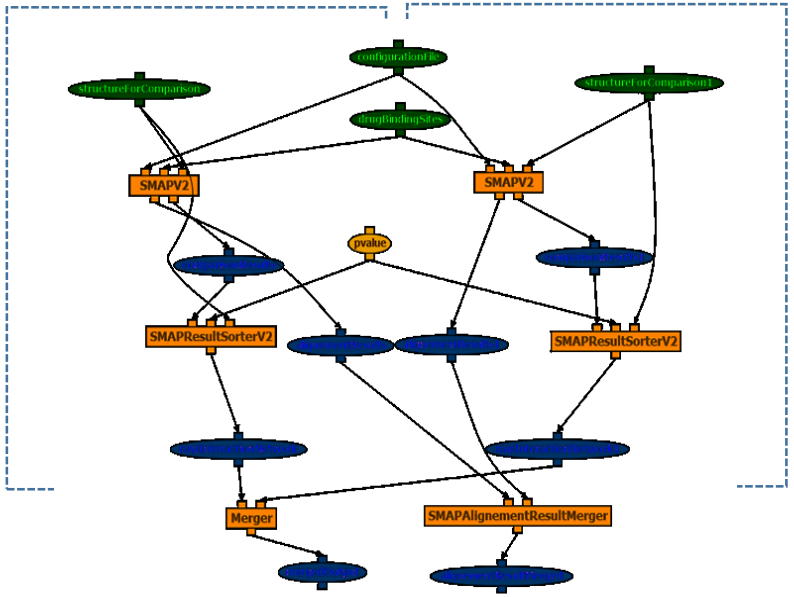


Workflow

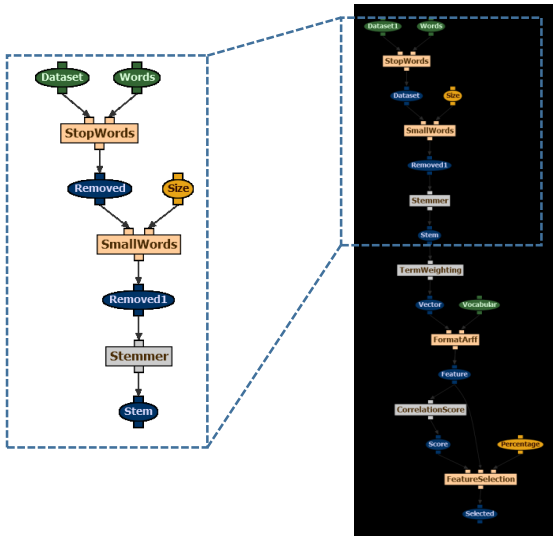
Specification 1 → Modeling 2 → Generation 3 → Publication 4 → Exploitation 5

Linked Data





Repeated workflow fragments in a workflow



Finding parts of a workflow in other workflows

Data Narratives

Daniel Garijo, with Yolanda Gil

Using the information produced by workflows and their executions to automatically write how the results were produced.

Data Narrative 1: Execution view **Hide**

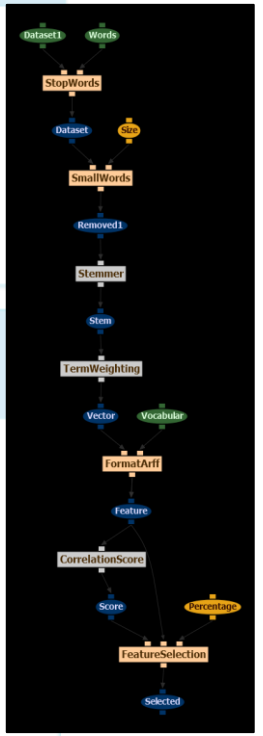
See more | See less

Parameters used in the execution

The Detect Topics method was run on dataset stop W.txt (input WordsToRemove) and dataset File.txt (input DocumentToFilter), with NUMBER1476159251570 set to 10 and ITERATIONS1476159251570 set to 5. The Image1476159251570 results are stored online.

Text automatically generated by DANA.

Links to results



Data Narrative 6: Software view **Hide**

See more | See less

Software used to produce the results

The method Trending Words Visualization has 5 steps. First, the input data is analyzed by Remove Markup, followed by Remove Stop Words, Word Count, and Quick Sort. The final results are produced by Tag Cloud. The steps use the following software:

Tag Cloud uses a bash script and a Java program (see the project website) to perform its functionality.The software is licensed under a GPLv2 license.
Remove Markup uses a bash script program to perform its functionality.The software is licensed under a CreativeCommons license.
Word Count uses a bash script program to perform its functionality.
Quick Sort uses a bash script and a C program (see the project website) to perform its functionality.The software is licensed under a BSD license.
Remove Stop Words uses a bash script and a Javascript program (see the project website) to perform its functionality.The software is licensed under a GPLv3 license.

Text automatically generated by DANA.