WF4Ever: PDL services integration in Taverna

Julian Garrido
Instituto de Astrofísica de Andalucía – CSIC
jgarrido@iaa.es
on behalf of the Wf4Ever Team

INTEROP, October 2012

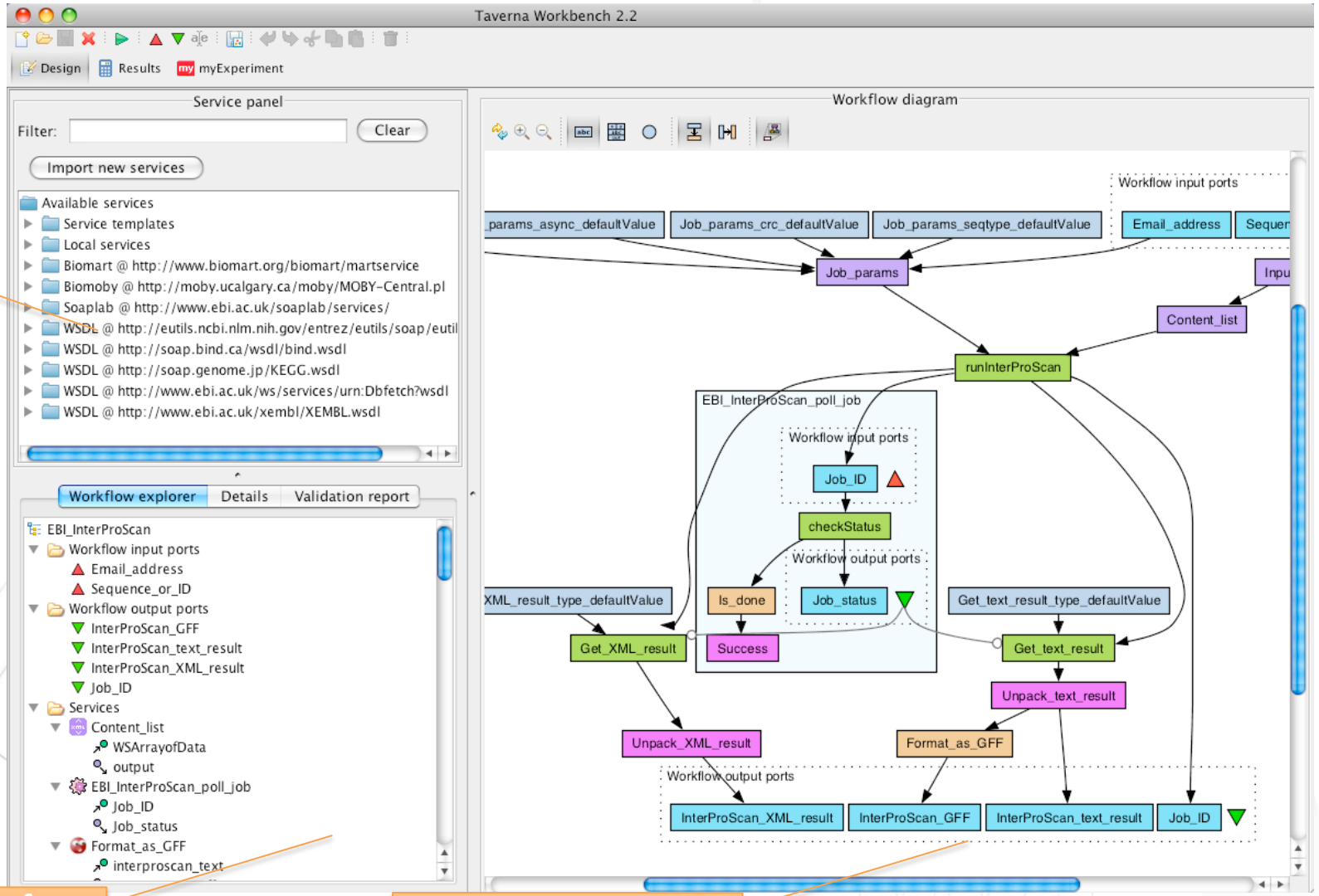
Wf4Ever - Preservation of scientific workflows in data-intensive science

EU funded FP7 STREP Project
December 2010 – December 2013

1. **Intelligent Software Components (ISOCO, Spain)**
2. University of Manchester (UNIMAN, UK)
3. Universidad Politécnica de Madrid (UPM, Spain)
4. Poznan Supercomputing and Networking Centre (PSNC, Poland)
5. University of Oxford (OXF, UK)
6. Instituto de Astrofísica de Andalucía (IAA, Spain)
7. Leiden University Medical Centre (LUMC, NL)

- Taverna is an open source domain independent Workflow Management System – a suite of tools used to design and execute scientific workflows.
- Available as a command line tool and a server.
- Allows for the **automation of experimental methods** through the use of a number of different (local or remote) services from a very diverse set of domains.
- Allows a scientist with limited computing background and limited technical resources and support to **construct highly complex analyses**.
- Handles parallelism, threading, monitoring, service discovery, provenance tracking.

Taverna workbench



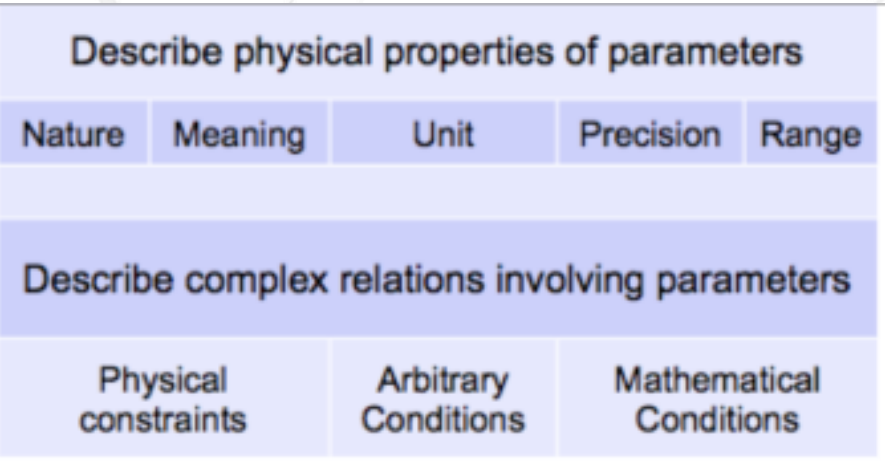
Available services

Tree view of workflow structure

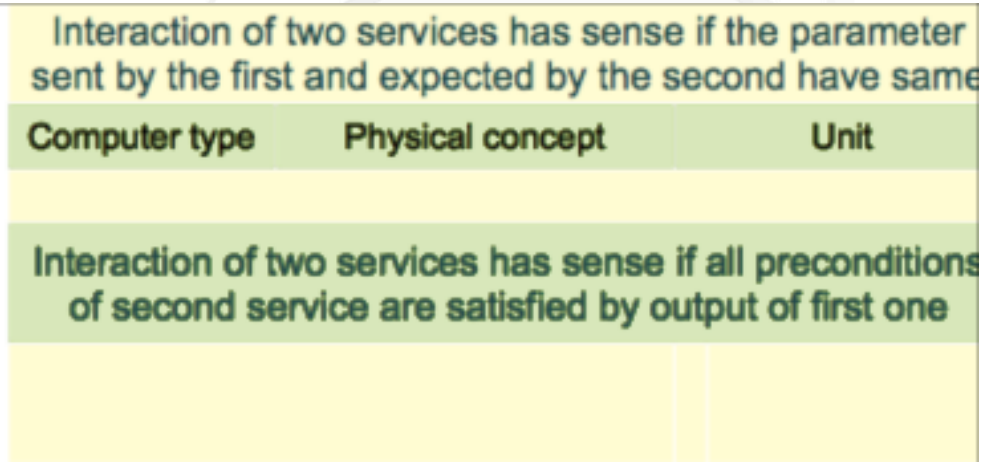
Workflow

- Parameter Description Language: A generic grammar for parameters (and relates constraints) description. Presented in the INTEROP as a working draft. (Carlo Maria Zwölf, Franck le Petit, Paul Harrison)
- PDL aim is to answer to two major issues in scientific services, (description and interoperability needs.

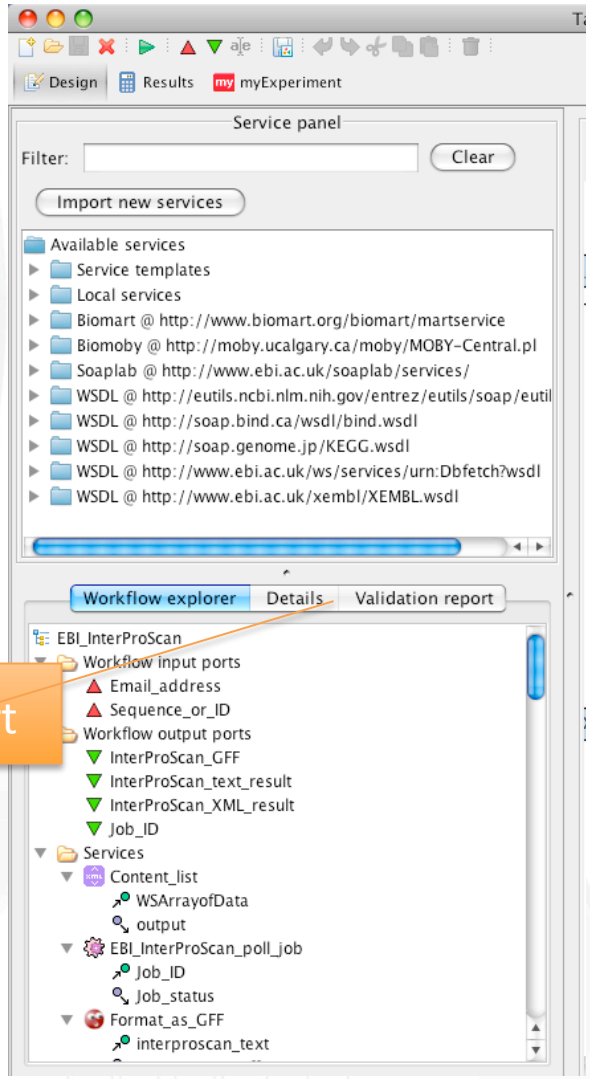
Description needs



Interoperability needs



- Running time:
  - PDL Validation client. Do input/output data fit the restrictions?
- Design time:
  - The output of a service is the input of another.
  - Computing interoperability (precision, SKOS concept, Unit, UCD, Utype)



- Missing points in Taverna:
  - Include validation capabilities. Running without errors doesn't mean the result is correct.
  - Improve service discovery and recommendations in Taverna: Which services can I connect to the output of another?

Questions