

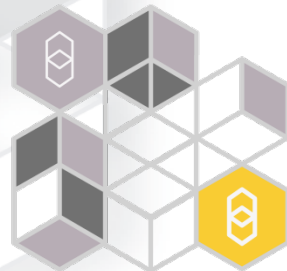


<http://taverna.org.uk/>

Stian Soiland-Reyes & Robert Haines
myGrid, School of Computer Science
University of Manchester, UK

TAVERNA WORKFLOW MANAGEMENT SYSTEM

myGrid



ITER IM workshop
Château de Cadarache, 2011-06-08

What is myGrid?

- ⦿ An e-Science Collaboration Since 2001
- ⦿ Not a grid!
- ⦿ Numerous partners involved:
 - University of Manchester
 - University of Southampton
 - University of Oxford
 - EMBL-EBI
- ⦿ Provides sustainable and production quality software
 - Supported by OMII-UK, EPSRC and BBSRC
- ⦿ Mixture of developers, bioinformaticians and researchers

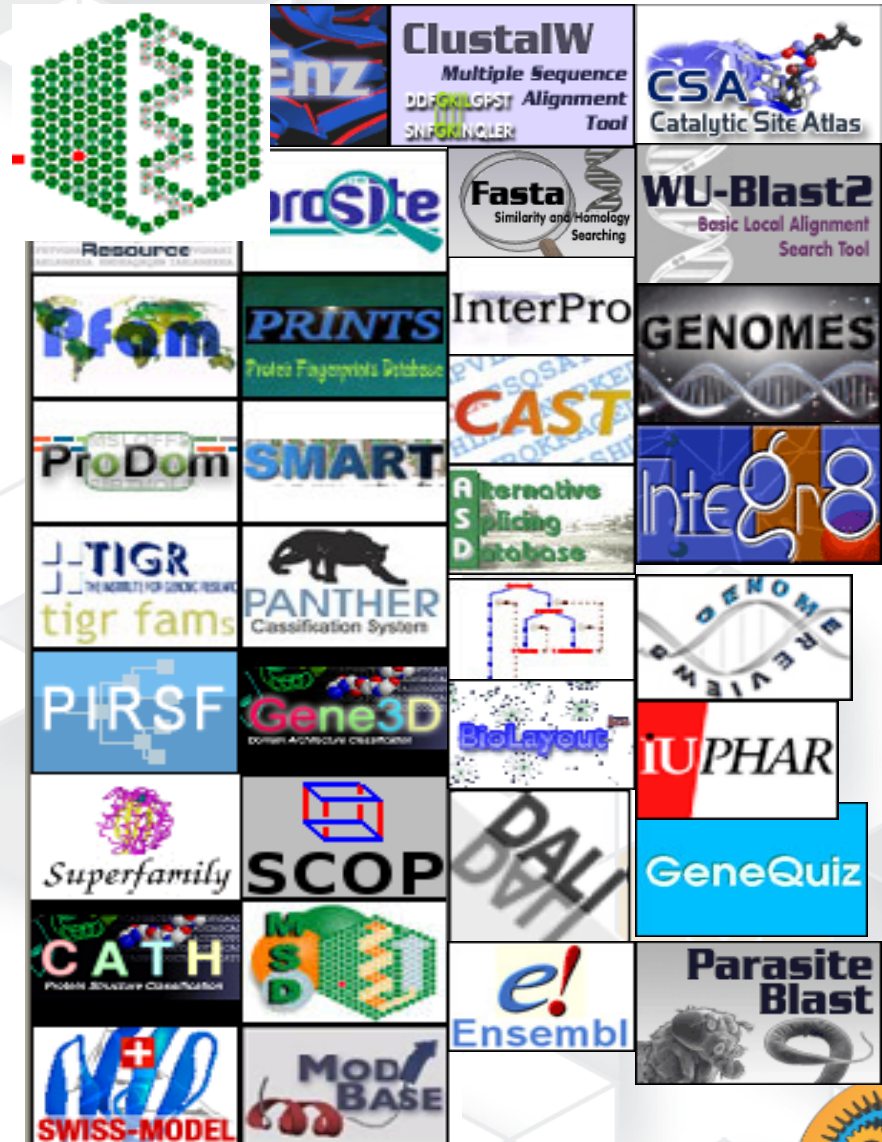
Software | Services | Content | Skills | Community



Motivation: Bioinformatics

Challenge:

- Large amounts of data
- Many open questions
- Numerous freely available public datasets and analysis tools



Manual approach

- ⦿ **Search** using public web sites and databases
 - Pubmed
 - Uniprot
 - EBI BioMart
- ⦿ **Copy and paste** to web tools for analysis
 - NCBI Blast
 - EBI InterPro
- ⦿ Further processing **locally**
 - R
 - Perl
 - Python



Manual: disadvantages

- **Scale** of analysis task overwhelms researchers – lots of data
- User **bias** and premature filtering of datasets – *cherry picking*
- **Hypothesis-Driven** approach to data analysis
- Constant **changes in data** - problems with re-analysis of data
- **Implicit methodologies** (hyper-linking through web pages)
- **Error proliferation** from any of the listed issues – notably human error

Web services and workflows

◉ Web services

- Technology and standards for exposing code and data resources that can be **programmatically consumed** by a remote third party
- Description on how to interact with the service, parameters, documentation

◉ Workflows

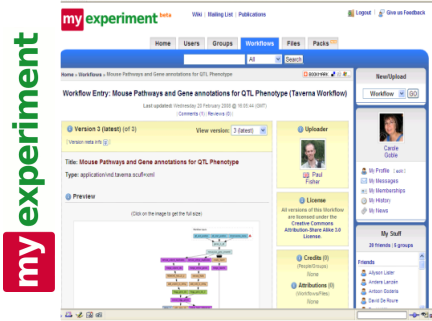
- General technique for describing and executing a process
- Describe **what** you want to do running **which services**



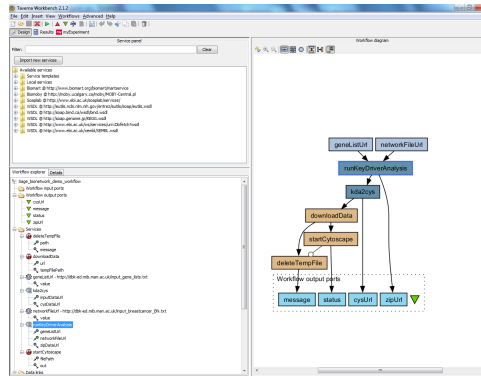
The Taverna Open Source Suite of Tools

Web Portals

Workflow Repository



GUI Workbench

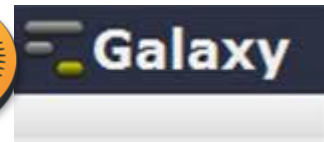


Client User Interfaces



Virtual Machine

Third Party Tools



Service Catalogue

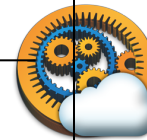


Workflow Engine



Provenance Store

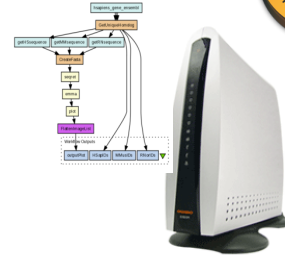
Workflow Server



Activity and Service Plug-in Manager



Open Provenance Model

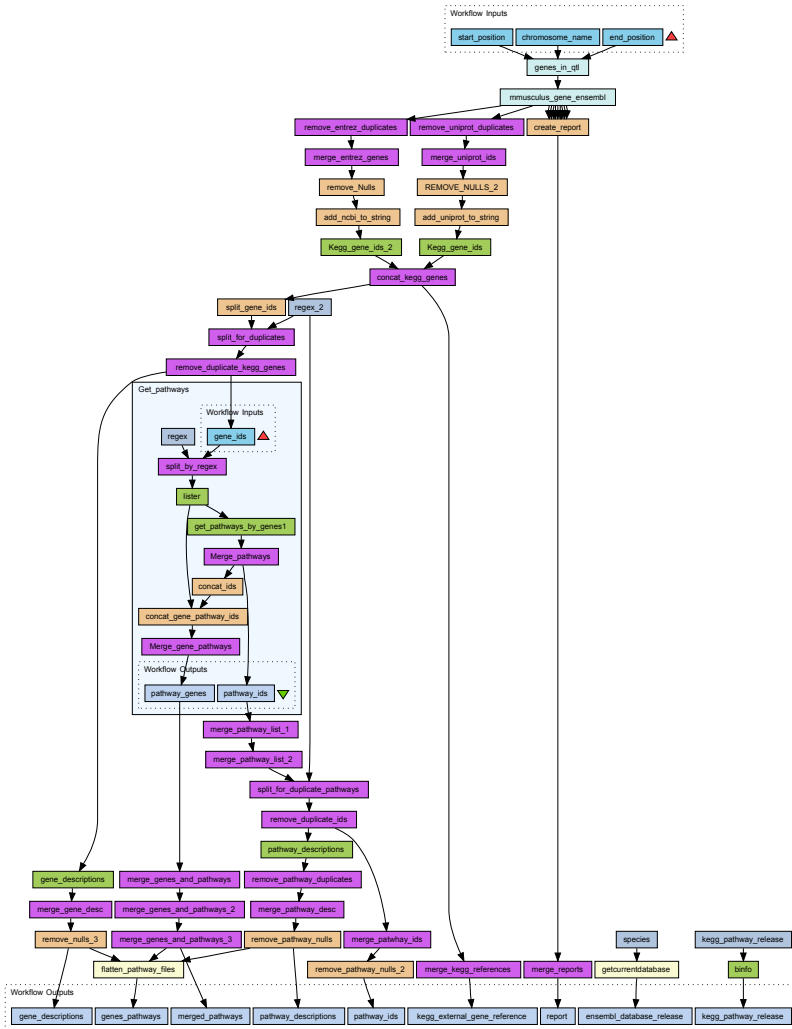


Secure Service Access



Programming and APIs

Taverna workflows



- A set of (local and remote) **services** to analyze or manage data
- **Nested** workflows are also services
- **Data-links** connects services
 - i.e. output from service A is input to service B and C
 - Describes the **desired dataflow** instead of process coordination
- Automatic iterations
- Can customize **list handling** and control links



What types of services and data?

- ⦿ WSDL/SOAP web services
 - Secured invocation with HTTPS/SSL/WS-Security
- ⦿ RESTful web services
 - Secured invocation with HTTPS/Basic Auth
- ⦿ Spreadsheet import
- ⦿ Command line tools (local, SSH)
- ⦿ Inline scripts (Beanshell, R)
- ⦿ Excel/CSV spreadsheets
- ⦿ Java APIs
- ⦿ Customizations:
 - BioMart, BioMoby / SADI
 - Soaplab
 - Grid services (EGEE gLite, caGrid, PBS, UNICORE)
 - ... **your** tool (Plugin tutorial in wiki)



Service limitations

- ⦿ Web service creation involves **wrapping** existing tools or writing WS code
- ⦿ Web services can go down
 - → can use redundant services in workflow
 - → Service **monitoring**
- ⦿ Transferring data up/down to WS slow
 - → Support **references** in WS interface
- ⦿ Executing command line tools directly requires execution access
 - Trickier to share workflows, require either SSH/grid **credentials** or installing tools locally



Which services?

- ⦿ Taverna is **general**, can connect to standard **web services** and **command line** tools for any domain
- ⦿ in bioinformatics..
 - From professional third-party organisations providing robust & open data/analysis services
 - ..to under-the-desk web services for one particular purpose, ran by PhD students
 - → <http://biocatalogue.org/> - 2000+ services from 140+ providers – crowd sourced and quality monitored



Top 20 tags on BioCatalogue [See All Tags]

- alignment | bioinformatics | biology | **BioMoby** | database | DDBJ | DOBJ_record | DNA | edit | EMBOSS |
 - EMBRACE | fasta_format | KEGG | KEGG_record_id | locus | protein sequence | retrieving | **retrieving** | soaplab |
 - StructuralStudies
- more

Filtering

Current Filters Applied
none

Select filters from below...

Enable tag filters

- Service Types (2)**
- SOAP (1657)
 - REST (73)

- Service Categories (60)**
- Sequence Analysis (184)
 - Text Mining (22)
 - Ontology (22)
 - Phylogeny (4)
 - Microarrays (42)
 - Data Retrieval (154)
 - Genomics (18)
 - Proteomics (35)
 - Systems Biology (22)
 - Bioinformatics (14)
 - Chemoinformatics (11)

Displaying services 1 - 10 of 1730 in total | Sort by: Newest

« Previous 1 2 3 4 5 6 7 8 9 ... 172 173 Next »

Switch to Simple View

MutalyzerService SOAP

aka Mutalyzer Name Checker | aka Mutalyzer Syntax Checker | Annotations: 55 | 42 | 13 | 0

aka Mutalyzer Position Converter

Categories: Sequence Analysis | Functional Genomics

Mutalyzer is a tool to check if descriptions of sequence variants comply to the standard human sequence variant nomenclature of the Human Genome Sequence Variation Society (HGVS) (For an overview, visit http://www.hgvs.org/mutnomen/). Mutalyzer aims to encourage the proper use of nomenclature in publications and reduce redundancy in sequence variation databases (See http://www.mutalyzer.nl/2.0/help for detailed information). In principle, Mutalyzer can check descriptions of sequence variants detected in other organisms, provided that the standard HGVS nomenclature is applied. The webservice provide programmatic access to functionality underlying the following browser interfaces: Mut...

by Peter Taschner (4 days ago)

Provider: www.mutalyzer.nl | Submitter / Source: Peter Taschner (4 days ago)

Tags on this service: mutation | nomenclature | position conversion | variant

WSDL Location: http://www.mutalyzer.nl/2.0/service.wsdl | 11 SOAP Operations

dbfetch REST

aka WSDbfetch (REST) | Annotations: 9 | 0 | 9 | 0

Categories: Data Retrieval

Dbfetch is an acronym for "database fetch". Dbfetch provides an easy way to retrieve entries from various databases at the EBI in a consistent manner. It can be used from any browser as well as well as within a web-aware scripting tool that uses wget, lynx or similar.

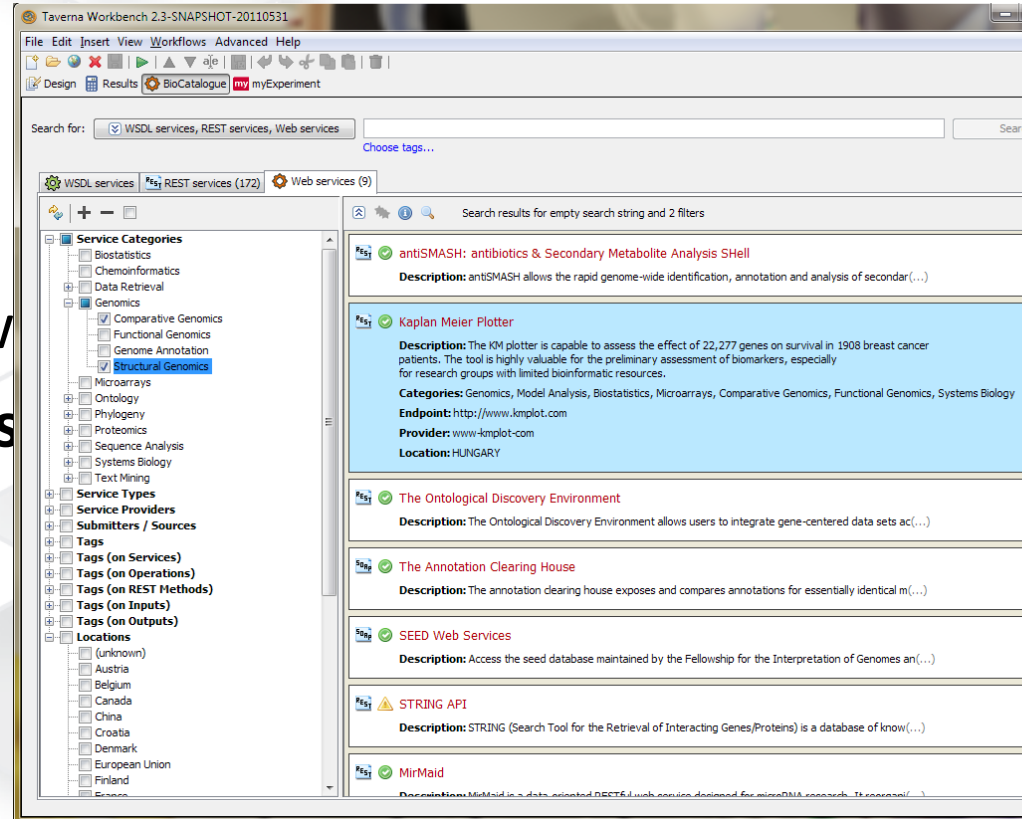
by Hamish McWilliam (about 1 month ago)

Provider: European Bioinformatics Institute (EBI) | Submitter / Source: Hamish McWilliam (about 1 month ago)



BioCatalogue integration

- Search services from workbench
- Add services to workflow
- View service descriptions and **uptime** status from within workflow



BioCatalogue



"The Life Science Web Services Registry"

<http://www.taverna.org.uk/>



<http://www.mygrid.org.uk/>

Taverna workbench

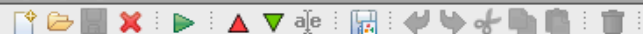
- Graphical desktop tool
- No server installation required
- Drag-and-drop services into diagram
- Connect services, run, reconnect, rerun
- Integrates diverse set of tools

The top screenshot shows the 'Progress report' window with the following table:

Name	Status	Queued	Iterations done	Iterations w/...	Average time...	First Iteration...	Last Iteration...
%: bits_iterate	Running	0	1	0	23 ms	14:15:49	14:15:50
boo - _l	Finished	0	1	0	139 ms	14:15:54	14:15:54
Concat	Running	180	39	0	31 ms	14:15:54	14:15:54
Concat	Running	0	39	0	38 ms	14:15:54	14:15:54
BiomartAndEMBOSSAnalysis	Running	2	36	0	38 ms	14:15:54	14:15:54
BiomartAndEMBOSSAnalysis	Running	2	36	0	38 ms	14:15:54	14:15:54

The middle screenshot shows the 'Workflow diagram' window with a flowchart of services. The bottom screenshot shows the 'Service panel' and 'Workflow diagram' windows.





Service panel

Filter: protein

Clear

Import new services

Matching services

Local services

ncbi

- Get Protein FASTA
- Get Protein GBSeg XML
- Get Protein INSDSeq XML
- Get Protein TinySeq XML

Biomart @ <http://www.biomart.org/biomart/martservice>

INTERPRO (EBI UK)

- protein - Protein Matches
- uniprot - UniProt Protein Matches

Workflow explorer

Details

Validation report

BiomartAndEMBOSSAnalysis

Workflow input ports

Workflow output ports

- HSapIDs
- MMusIDs
- outputPlot
- RNorIDs

Services

CreateFasta

- hsSeq
- mmSeq
- rnSeq
- fasta

emma

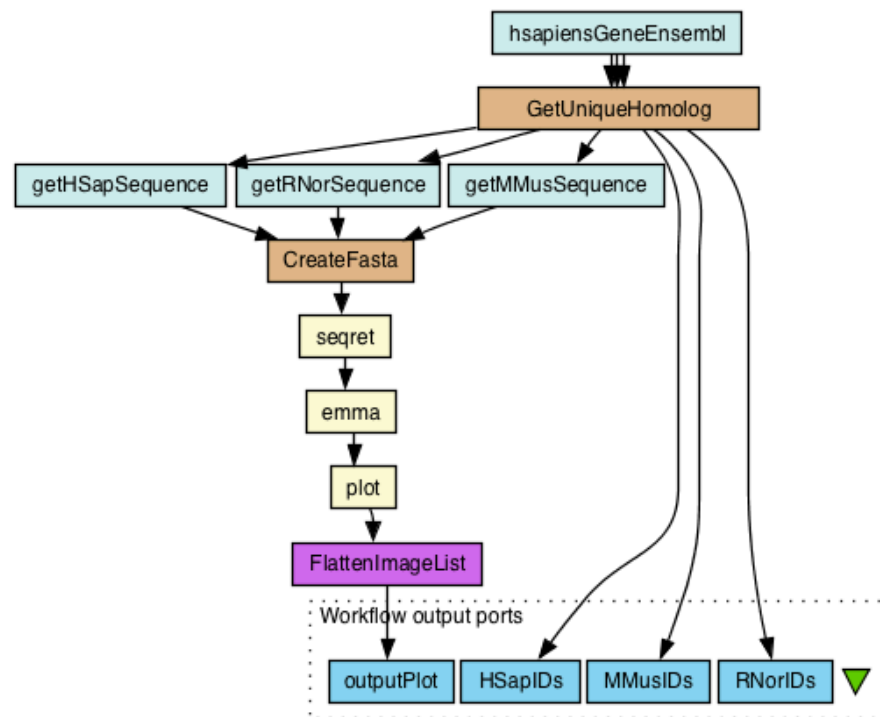
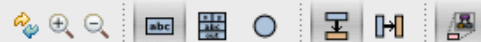
FlattenImageList

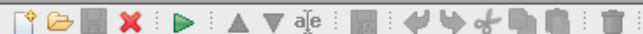
- inputlist
- outputlist

getHSapSequence

- hsapiens_gene_ensembl.ensembl_gene_id_filter

Workflow diagram





Workflow runs

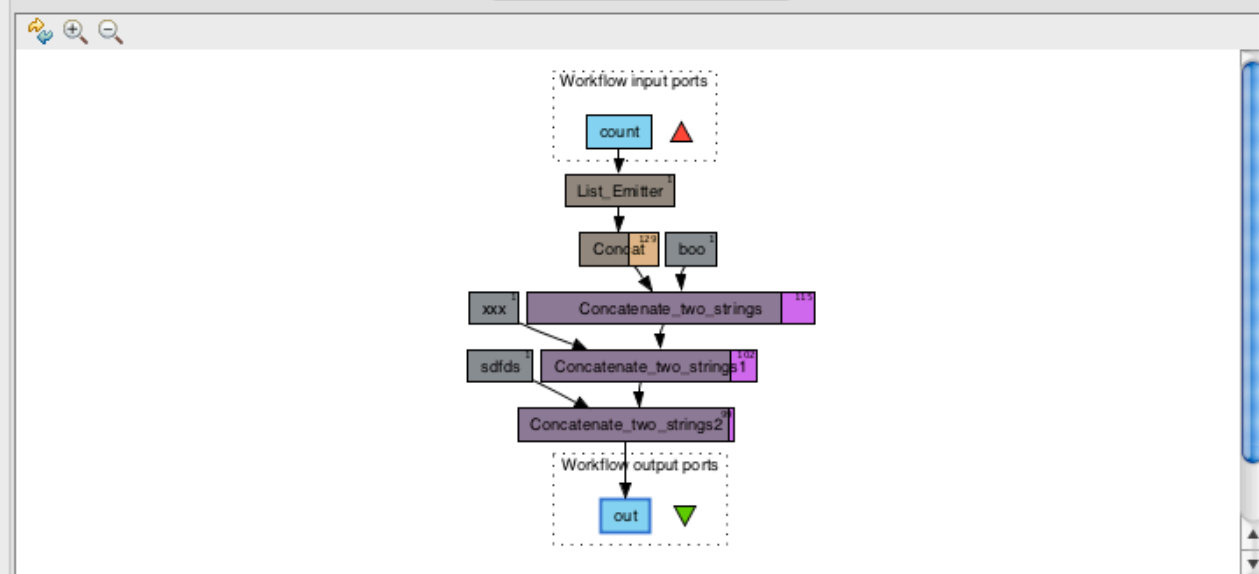
Remove

Click on a run to see its values
Click on a service in the diagram
to see intermediate values (if available)

lists_iterate 2010-11-29 14:15:15
 BiomartAndEMBOSSAnalysis 2010-11-29 14:13:35
 BiomartAndEMBOSSAnalysis 2010-11-29 14:13:27
 BiomartAndEMBOSSAnalysis 2010-11-29 14:13:09
 Spreadsheet_Import_Example 2010-11-04 17:26:41
 Workflow65 2010-10-27 16:46:05
 Workflow65 2010-10-27 16:43:29
 Workflow65 2010-10-27 16:40:45
 Workflow65 2010-10-27 16:40:09
 Workflow65 2010-10-27 16:39:16
 Workflow65 2010-10-27 16:37:56
 Workflow65 2010-10-27 16:36:32
 Workflow64 2010-10-27 16:35:34
 Workflow65 2010-10-27 16:33:23
 Workflow65 2010-10-27 16:33:05
 Workflow65 2010-10-27 16:32:40
 Workflow65 2010-10-27 16:32:26
 Workflow65 2010-10-27 16:31:20
 Workflow65 2010-10-27 16:30:20

Graph

Progress report


 Running

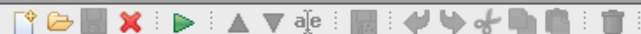
Workflow results

Save all values

Click in tree to view values

Value type

- ▼ List with 101 values
 - Value 1
 - Value 2
 - Value 3
 - Value 4
 - Value 5
 - Value 6
 - Value 7
 - Value 8
 - Value 9
 - Value 10



Workflow runs

Remove

Click on a run to see its values
Click on a service in the diagram
to see intermediate values (if available)

lists_iterate 2010-11-29 14:15:15

BiomartAndEMBOSSAnalysis 2010-11-29 14:13:35

BiomartAndEMBOSSAnalysis 2010-11-29 14:13:27

BiomartAndEMBOSSAnalysis 2010-11-29 14:13:09

Spreadsheet_Import_Example 2010-11-04 17:26:41

Workflow65 2010-10-27 16:46:05

Workflow65 2010-10-27 16:43:29

Workflow65 2010-10-27 16:40:45

Workflow65 2010-10-27 16:40:09

Workflow65 2010-10-27 16:39:16

Workflow65 2010-10-27 16:37:56

Workflow65 2010-10-27 16:36:32

Workflow64 2010-10-27 16:35:34

Workflow65 2010-10-27 16:33:23

Workflow65 2010-10-27 16:33:05

Workflow65 2010-10-27 16:32:40

Workflow65 2010-10-27 16:32:26

Workflow65 2010-10-27 16:31:20

Workflow65 2010-10-27 16:30:20

Graph Progress report

Name	Status	Queued itera...	Iterations done	Iterations w/...	Average time...	First iteration...	Last iteration...
lists_iterate	Running	-	-	-		14:15:49	
boo - ,b	Finished	0	1	0	23 ms	14:15:50	14:15:50
Concat	Running	160	39	0	139 ms	14:15:54	
Concat	Running	0	39	0	31 ms	14:15:54	
Concat	Running	2	36	0	38 ms	14:15:54	
Concat	Running	2	33	0	30 ms	14:15:54	
List_Emit	Finished	0	1	0	1.8 s	14:15:51	14:15:53
sdfds -	Finished	0	1	0	176 ms	14:15:50	14:15:50
xxx - X	Finished	0	1	0	17 ms	14:15:50	14:15:50

 Running

Workflow results

Click in tree to view values

List with 101 values

- Value 1
- Value 2
- Value 3
- Value 4
- Value 5
- Value 6
- Value 7
- Value 8
- Value 9
- Value 10

Value type

Sharing workflows

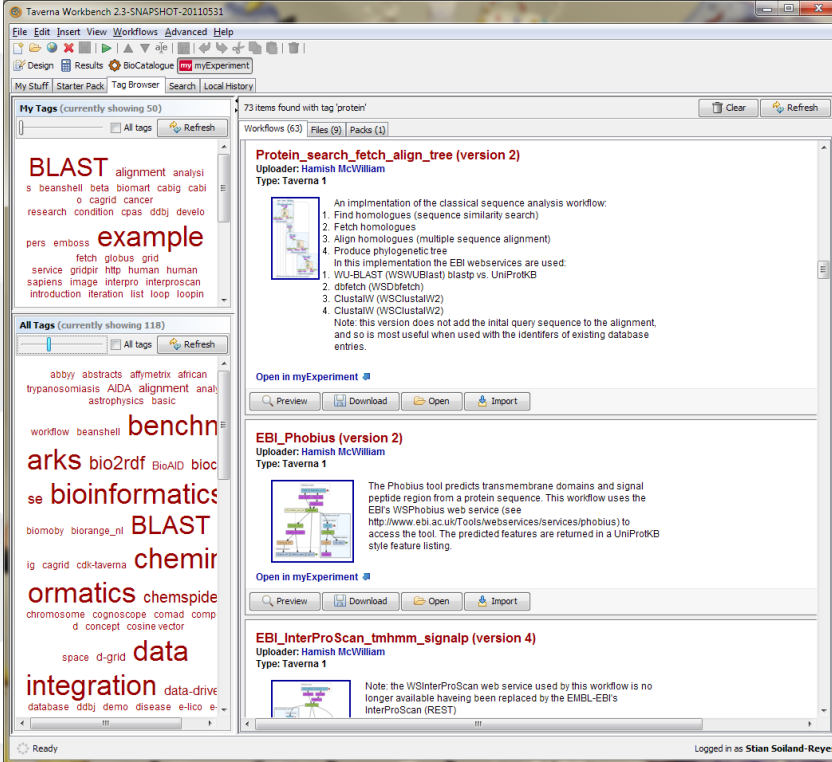
- myExperiment.org allows users to share, find, download and rate workflows
- “Facebook for the scientist”
- 4000+ members, 1400+ workflows
- Open source code, can set up own instance

my experiment



myExperiment integration

- Search and browse workflows
 - By tags
 - Free text search
 - Own/group workflows
 - Packs, e.g. “Examples”
- Upload/share workflows



The screenshot displays the Taverna Workbench interface. The main window shows a search results page for workflows. The left sidebar contains a 'My Tags' section with a list of tags including 'blast', 'alignment', 'analysis', 'beanshell', 'beta', 'biomart', 'cabig', 'cabi', 'cagid', 'cancer', 'research', 'condition', 'cpas', 'dobj', 'develo', 'pers', 'emboss', 'example', 'fetch', 'globus', 'grid', 'service', 'gridpir', 'http', 'human', 'human', 'sapiens', 'image', 'interpro', 'interproscan', 'introduction', 'iteration', 'list', 'loop', 'loopin'. Below this is an 'All Tags' section with a list of tags including 'abby', 'abstracts', 'affymetric', 'african', 'trypanosomiasis', 'AIDA', 'alignment', 'anal', 'astrophysics', 'basic', 'workflow', 'beanshell', 'bench', 'arks', 'bio2rdf', 'BioAID', 'bioc', 'se', 'bioinformatics', 'biomoby', 'biorange', 'BLAST', 'ig', 'cagid', 'cdk-taverna', 'chemir', 'ormatics', 'chemspide', 'chromosome', 'cognoscope', 'comad', 'comp', 'd', 'concept', 'cosine', 'vector', 'space', 'd-grid', 'data', 'integration', 'data-drive', 'database', 'dobj', 'demo', 'disease', 'e-lico', 'e-'. The main content area shows a list of workflows, with the first one being 'Protein_search_fetch_align_tree (version 2)' by Hamish McWilliam. The workflow description includes a list of steps: 1. Find homologues (sequence similarity search), 2. Fetch homologues, 3. Align homologues (multiple sequence alignment), 4. Produce phylogenetic tree. Below the description are buttons for 'Open in myExperiment', 'Preview', 'Download', 'Open', and 'Import'. The second workflow is 'EBI_Phobius (version 2)' by Hamish McWilliam, which predicts transmembrane domains and signal peptide regions. The third workflow is 'EBI_InterProScan_tmhmm_signalp (version 4)' by Hamish McWilliam, which uses the WSIInterProScan web service.

Taverna workflow features

- ⦿ Nested workflows
 - Reuse existing components
- ⦿ Implicit iterations
 - With customizable list handling
- ⦿ Pipelining
 - Process partial iteration results early
- ⦿ Parallelisation
 - Run as soon as data is available
- ⦿ Retries, failover, looping
 - For stability and conditional testing
- ⦿ Plugin-extensible execution control
 - Ideas: caching, error detection, dynamic service lookup



Extensible UI and engine

- ⦿ Plugins can provide new “perspectives”
 - e.g.: BioCatalogue, myExperiment
- ⦿ Provide service-specific customization
 - e.g.: BioMart interface replicates web site
- ⦿ Adding new functionality
 - New service types, eg: ...
 - Execution control like looping/branching
 - Design helpers, “Find matching service”



Workflow limitations

- ⦿ Initially designed for **dataflows**
 - Not suitable for business processes like “*HR procedure for hiring new staff*”
 - *Long-running workflows require Taverna Server*
 - .. But suitable for **coordinating** command line and grid executions, the data might just be job references
 - Execution control extensible, eg:
 - Looping, Branching
 - Dynamic service lookup
 - Data manipulation, Error detection

Data and provenance handling

- ⦿ **Data references** passed between services in workflow
 - http, file, sftp, gridftp, etc (extensible)
- ⦿ Data downloaded/uploaded or references **translated** when needed
- ⦿ **Provenance** captured for workflow runs
 - **Trace** execution steps, view **intermediate values** while running
 - Export as Open Provenance Model (OPM) / RDF
 - Proof and **origin** of produced outputs
 - Extensible **annotations**
- ⦿ Wf4Ever: reproducible **research objects**
 - Workflow/data as a scientific publication → preservation
 - Need to capture more service data and metadata



Data limitations

- ⦿ Running Workbench limited by:
 - Local disk space for storing data
 - Network speeds for up/download
 - Firewall access
 - → Execute wf using **Taverna Server** or **command line** remotely with ssh/job submission
- ⦿ No standardized WS reference mechanism
 - Agree on mechanism within WS 'family' with shared disk (eg. deconstruct local path from HTTP URI)

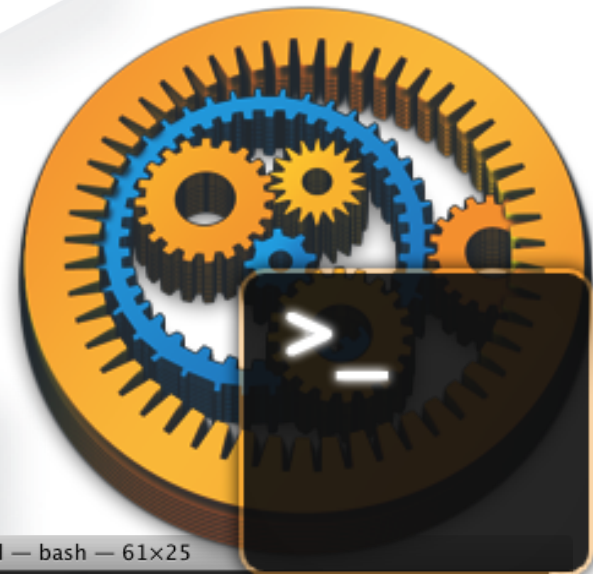


Parameter sweeps

- ⦿ **Implicit iterations** with **pipelining** provides an intuitive way to set up parameter sweeps
- ⦿ Advanced **looping** and extensible **execution control** allows iterative & recursive reductions/approximations



Taverna command line



- Executes from a Windows/Linux/OSX shells
- Takes a predefined workflow with files as inputs and outputs
- Quick way to “productionize” a workflow

```
Terminal — bash — 61x25
heater:~ stain$ executeworkflow ebi_interproscan.t2flow
The number of inputs provided does not match the number of input ports.
2 inputs were expected which are:
'Email_address' 'Sequence_or_ID'
0 inputs were provided
heater:~ stain$ executeworkflow ebi_interproscan.t2flow -inputvalue Email_address nobody@soiland-reyes.com -inputvalue Sequence_or_ID uniprot:wap_rat
Outputs will be saved to the directory: /Users/stain/EBI_InterProScan_output
heater:~ stain$ ls EBI_InterProScan_output/
InterProScan_GFF
InterProScan_XML_result
InterProScan_text_result
Job_ID
status
heater:~ stain$ cat EBI_InterProScan_output/InterProScan_text_result
```

Taverna Server

- REST/SOAP interface to execute workflows
- Client libraries for Ruby and Java
- Two demonstration web interfaces
 - Ruby
 - Java Portlets
- Upcoming:
 - Security delegation
 - AWS image



Taverna portlet

- Example portlet interface
- Executes workflows using Taverna Server

Workflow Submission

Select a workflow to run:

Workflow: EBI_InterProScan_T2

Description: Perform an InterProScan analysis of a protein sequence using the EBI's WSInterProScan service (see <http://www.ebi.ac.uk/Tools/webservices/services/interproscan>). The input sequence to use and the user e-mail address are inputs, the other parameters for the analysis (see Job_params) are allowed to default. InterProScan searches a protein sequence against the protein family and domain signature databases integrated into InterPro (see <http://www.ebi.ac.uk/interpro/>). A set of matches to the signatures are returned, which are annotated with the corresponding InterPro and GO term assignments for these signature matches.

Workflow inputs:

Name	Type	Description	Value
Email_address	single value	Your email address (for EBI statistics on job submissions)	Paste the value here: <input type="text"/> Or load the value from a file: <input type="button" value="Choose File"/> No file chosen
Sequence_or_ID	single value	Sequence or ID Example value: uniprot:wap_rat	Paste the value here: <input type="text"/> Or load the value from a file: <input type="button" value="Choose File"/> No file chosen



To view results of a workflow submission job, click on its id, once its status becomes "Finished".

To refresh job statuses, click on the "Refresh" button.

Workflow submission jobs:

Refresh

Job Id	Workflow name	Start date	Status	Delete job
02eac9ec-e590-4628-8830-6367ea5c5794	biomoby_tutorial_workflow_877727	Mon, 29 Nov 2010 17:30:45 GMT	Operating	
d0280ad0-1119-4d08-aa76-45fc829308a7	ebi_interproscan_for_taverna_2_317472	Mon, 29 Nov 2010 17:27:56 GMT	Finished	

Job Id: d0280ad0-1119-4d08-aa76-45fc829308a7
 Workflow: ebi_interproscan_for_taverna_2_317472

Download the results as a [single Baclava X](#)
 You can view the file with Taverna's DataView

Results:

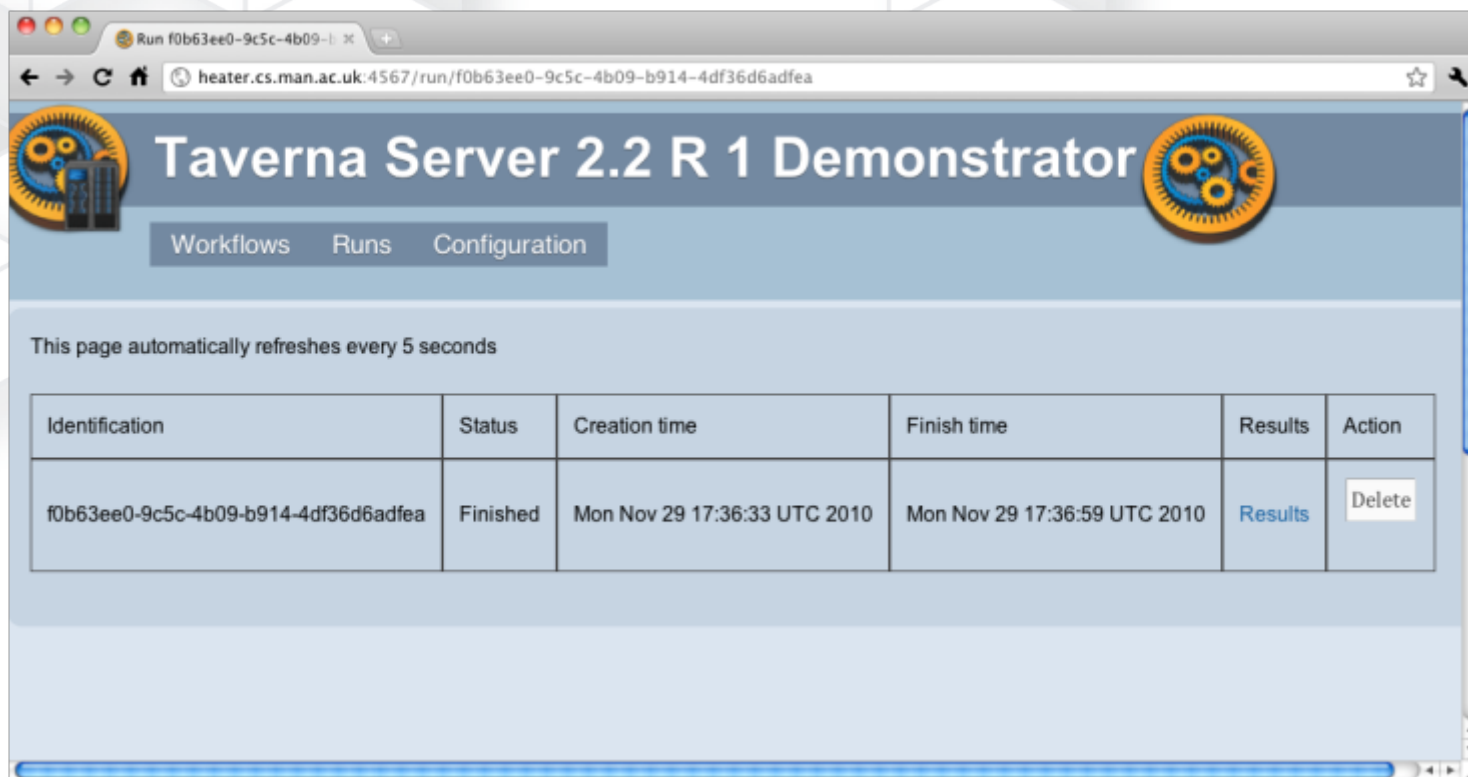
Output port	Data
Job_ID - single value <i>text/plain</i>	Value
InterProScan_text_result - single value <i>text/plain</i>	Value
InterProScan_XML_result - single value <i>text/xml</i>	Value
status - single value <i>text/plain</i>	Value
InterProScan_GFF - single	Value

Data preview
<pre> sp P01174 WAP_RATPrintScan 4DISULPHCORE 76 85 1.0E-6 . . InterProScan ; IPR015874 4-disulphide core sp P01174 WAP_RATPrintScan 4DISULPHCORE 101 108 1.0E-6 . . InterProScan ; IPR015874 4-disulphide core sp P01174 WAP_RATPrintScan 4DISULPHCORE 108 117 1.0E-6 . . InterProScan ; IPR015874 4-disulphide core sp P01174 WAP_RATPrintScan 4DISULPHCORE 118 126 1.0E-6 . . InterProScan ; IPR015874 4-disulphide core sp P01174 WAP_RATProfileScan WAP 27 73 0.0 . . InterProScan ; IPR008 Whey acidic protein, 4-disulphide core sp P01174 WAP_RATProfileScan WAP 76 127 0.0 . . InterProScan ; IPR008 Whey acidic protein, 4-disulphide core sp P01174 WAP_RATHMMPfam WAP 79 126 5.299999999999999E-7 . . InterP IPR008197 Whey acidic protein, 4-disulphide core sp P01174 WAP_RATHMMPanther PTHR19441 4 133 3.4999946686394883E-34 . InterProScan sp P01174 WAP_RATsuperfamily WAP 77 129 1.699999222840496E-8 . . InterProScan ; IPR008197 Whey acidic protein, 4-disulphide core </pre>



Ruby web interface

- Example customized web interface
- Uses Ruby gem t2-server



Identification	Status	Creation time	Finish time	Results	Action
f0b63ee0-9c5c-4b09-b914-4df36d6adfea	Finished	Mon Nov 29 17:36:33 UTC 2010	Mon Nov 29 17:36:59 UTC 2010	Results	<input type="button" value="Delete"/>



Grids and clusters

⦿ Taverna have been integrated with several leading grid and middleware infrastructures, such as:

- PBS
- caGrid/Globus
- EGEE/gLite
- NorduGrid's ARC
- JSDL/GridSAM

⦿ Plans for SAGA integration




Taverna on the cloud

Use-case:

- SNP analysis and annotation of genome sequenced from breeds of cows in Africa – why are some of them resistant to X?
- Amazon EC2 with Taverna Server and local services
- Ruby on Rails web interface
- Runs through 31 chromosomes in 2 hours using 10 instances - \$10



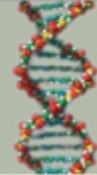


Next Gen Sequence Analysis

Using Taverna

[Edit registration](#)
[Logout](#)

Inputs Experiments



[Start experiment](#)

[Delete experiment](#)

[Create a new experiment](#)

[Back to all experiments](#)

Experiment: Harry's Boran Data (Liverpool)

[Start experiment](#)

Description

SNP sequencing data from Boran (March 2010). Chromosomes 1, 2, 7, 12, and 15.

Details

Start time: 2010-09-24 17:29:47 UTC	End time: 2010-09-24 17:30:56 UTC	Job status: Completed
Use polyphen: true Liftover: bosTauMd3ToBosTau4.over.cha Species: Cow		

Job Execution Info

Completed

Nodes 3 currently running
[Completed](#) [Completed](#) [Completed](#)

Polyphen datasets loaded 3 currently running
[Completed](#) [Completed](#) [Completed](#)

Reference genome instances loaded across nodes 3 currently running
[Completed](#) [Completed](#) [Completed](#)

Input files (5)
[Completed](#) [Completed](#) [Completed](#) [Completed](#) [Completed](#)

Inputs

Input data: [Better data](#)

Input data statistics:

Chromosome	SNP count	Files
------------	-----------	-------

Taverna 3 roadmap

- ⦿ OSGi plugin system
- ⦿ Workflow language: Scuf12
 - Compound format; embedding metadata, dependencies, independent API for creating/inspecting workflows
- ⦿ Components
 - Finding/sharing command line tool descriptions
 - Richer way of finding compatible services



Open source, open development

- ⦿ Taverna suite of tools are all **open source**, free to use and **customize**
- ⦿ Large user **community**, active mailing lists
- ⦿ Lead developers: **myGrid** in Manchester UK
- ⦿ **Contributors** from across the world
- ⦿ **PAL** programme
- ⦿ myGrid provides **training**, tutorials and **documentation**



Who uses Taverna?

- ⦿ Bioinformatics: EMBL-EBI, ONDEX
- ⦿ Astronomy: HELIO, AstroGrid, SAMPO
- ⦿ Engineering: NASA Jet Propulsion Lab (JPL)
- ⦿ Chemistry: CDK, CIC
- ⦿ Biodiversity: BioVel
- ⦿ Preservation: Wf4Ever, SCAPE
- ⦿ BioMedicine/Cancer research: caGrid
- ⦿ Data/text mining: eLico, AID



Taverna in numbers

⦿ Taverna:

- 361 organisations
- 48 countries
- 70,000+ downloads
- ~4000 source

⦿ myExperiment:

- 4000+ registered users
- 56 countries
- 1400+ workflows

⦿ BioCatalogue:

- 2000+ services
- 150+ service providers
- 500+ members
- 27 countries



Acknowledgements

MANCHESTER
1824



EPSRC

JISC



Lilly

Microsoft®

nbic
netherlands
bioinformatics
centre



UNIVERSITY OF
Southampton
School of Electronics
and Computer Science



E · S · R · C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL



<http://www.mygrid.org.uk/>

<http://www.taverna.org.uk/>





More information

- ◎ <http://www.mygrid.org.uk/>
- ◎ <http://www.taverna.org.uk/>
- ◎ <http://www.myexperiment.org/>
- ◎ <http://www.biocatalogue.org/>

