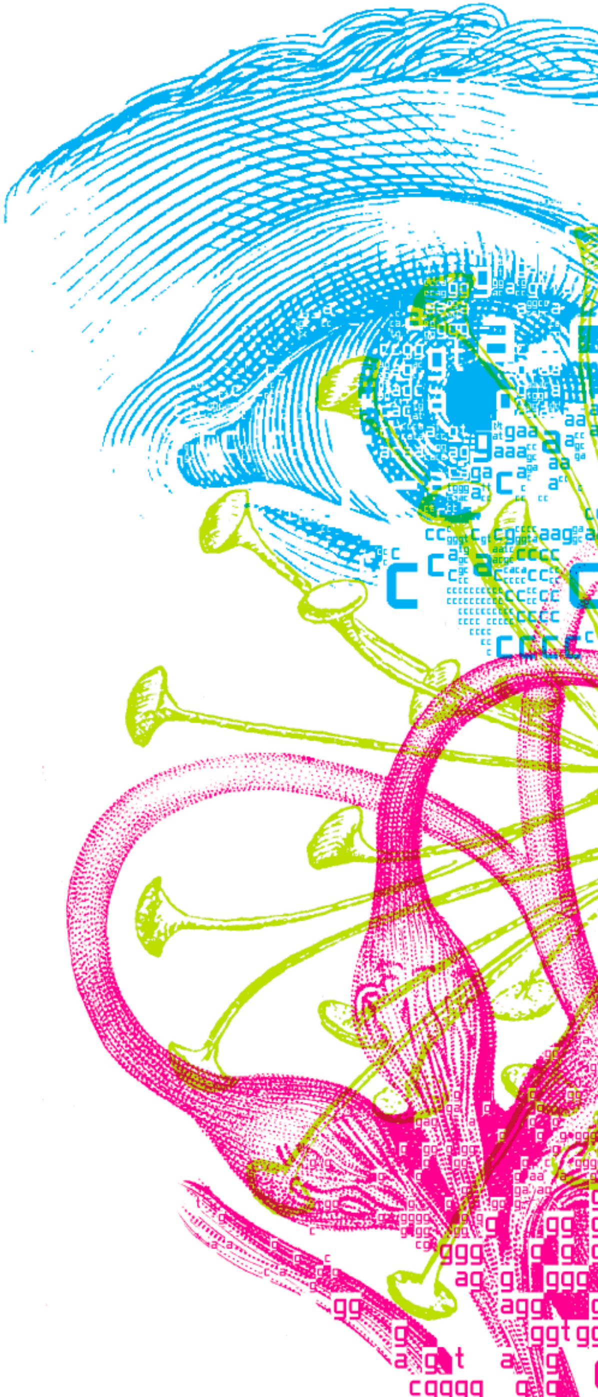
nbic
Netherlands
Bioinformatics
Centre



Enacting Taverna Workflows through Galaxy

Konstantinos Karasavvas

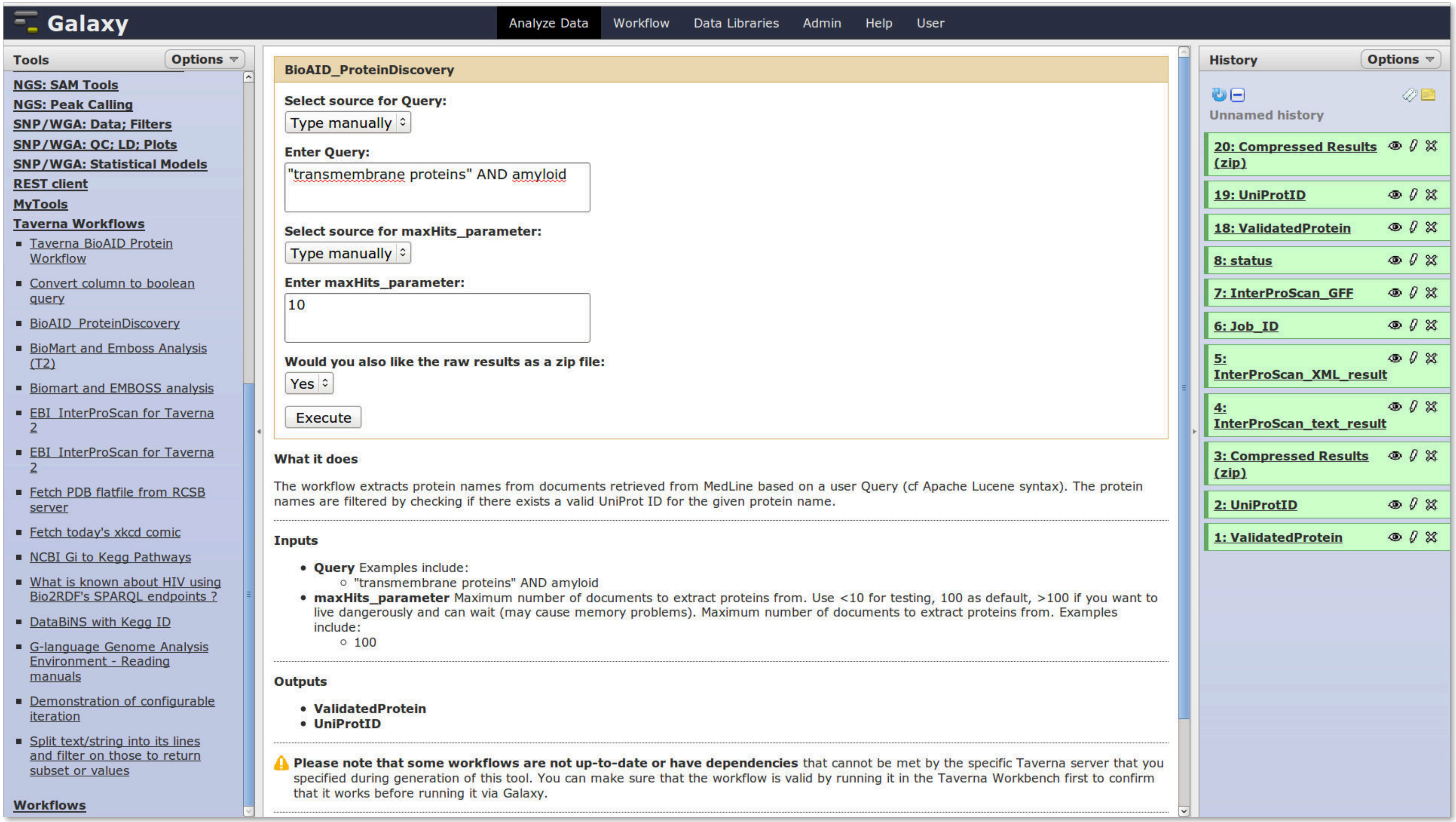
LUMC
Leiden University Medical Center

Introduction

- Software pipeline systems
  - Need to aggregate tools together
    - specify execution ordering
    - handle tools I/Os
    - pipelines, workflows, …

- Galaxy: web portal and framework for bioinformatics
- Taverna: workflow management system
- myExperiment: scientific social networking web site

Some Background: Galaxy



BioAID_ProteinDiscovery

Select source for Query:
Type manually

Enter Query:
"transmembrane proteins" AND amyloid

Select source for maxHits_parameter:
Type manually

Enter maxHits_parameter:
10

Would you also like the raw results as a zip file:
Yes

What it does

The workflow extracts protein names from documents retrieved from MedLine based on a user Query (cf Apache Lucene syntax). The protein names are filtered by checking if there exists a valid UniProt ID for the given protein name.

Inputs

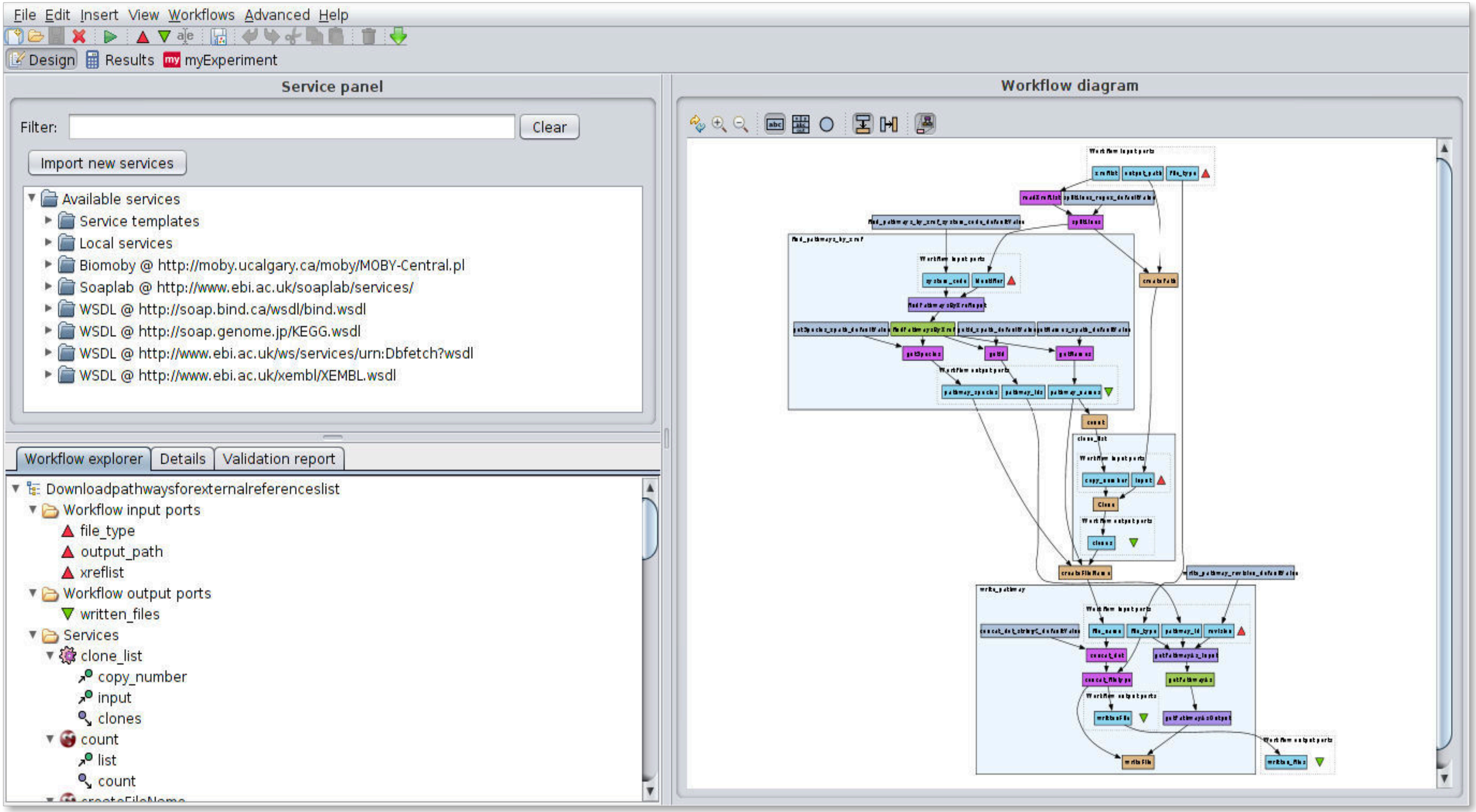
- **Query** Examples include:
  - "transmembrane proteins" AND amyloid
- **maxHits_parameter** Maximum number of documents to extract proteins from. Use <10 for testing, 100 as default, >100 if you want to live dangerously and can wait (may cause memory problems). Maximum number of documents to extract proteins from. Examples include:
  - 100

Outputs

- **ValidatedProtein**
- **UniProtID**

⚠ **Please note that some workflows are not up-to-date or have dependencies** that cannot be met by the specific Taverna server that you specified during generation of this tool. You can make sure that the workflow is valid by running it in the Taverna Workbench first to confirm that it works before running it via Galaxy.

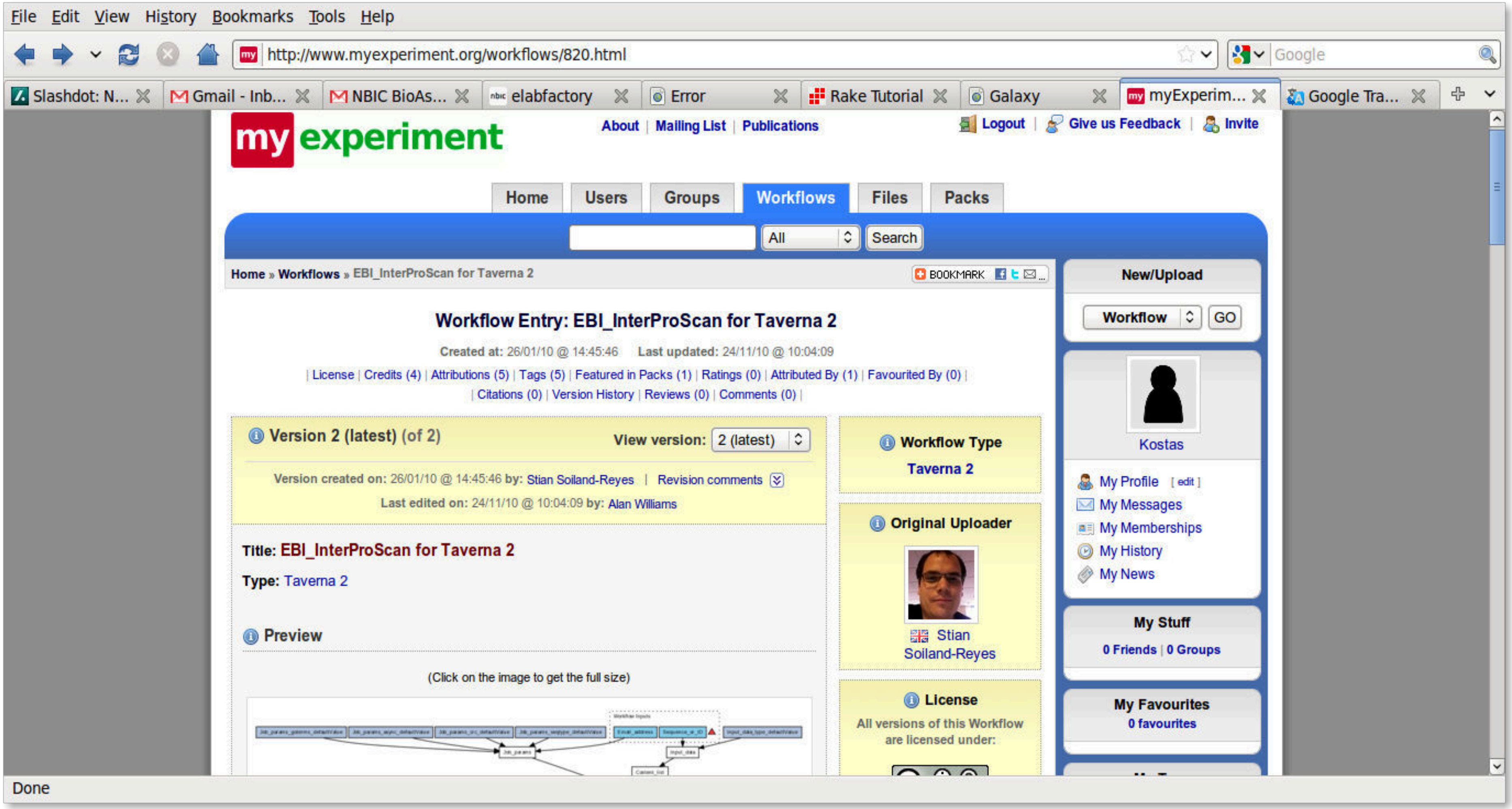
Some Background: Taverna Workbench



- Available services
  - Service templates
  - Local services
  - Biomoby @ http://moby.ucalgary.ca/moby/MOBY-Central.pl
  - Soaplab @ http://www.ebi.ac.uk/soaplab/services/
  - WSDL @ http://soap.bind.ca/wsdl/bind.wsdl
  - WSDL @ http://soap.genome.jp/KEGG.wsdl
  - WSDL @ http://www.ebi.ac.uk/ws/services/urn:Dbfetch?wsdl
  - WSDL @ http://www.ebi.ac.uk/xembl/XEMBL.wsdl

- Downloadpathwaysforexternalreferenceslist
  - Workflow input ports
    - file_type
    - output_path
    - xreflist
  - Workflow output ports
    - written_files
  - Services
    - clone_list
      - copy_number
      - input
      - clones
    - count
      - list
      - count

Some Background: myExperiment



Home » Workflows » EBI_InterProScan for Taverna 2

Workflow Entry: EBI_InterProScan for Taverna 2

Created at: 26/01/10 @ 14:45:46 Last updated: 24/11/10 @ 10:04:09

| License | Credits (4) | Attributions (5) | Tags (5) | Featured in Packs (1) | Ratings (0) | Attributed By (1) | Favourited By (0) |
| Citations (0) | Version History | Reviews (0) | Comments (0) |

Version 2 (latest) (of 2) View version: 2 (latest)

Version created on: 26/01/10 @ 14:45:46 by: Stian Soiland-Reyes | Revision comments

Last edited on: 24/11/10 @ 10:04:09 by: Alan Williams

Title: EBI_InterProScan for Taverna 2

Type: Taverna 2

Preview

(Click on the image to get the full size)

Workflow Type

Taverna 2

Original Uploader

Stian Soiland-Reyes

License

All versions of this Workflow are licensed under:

Kostas

My Profile [edit]

My Messages

My Memberships

My History

My News

My Stuff

0 Friends | 0 Groups

My Favourites

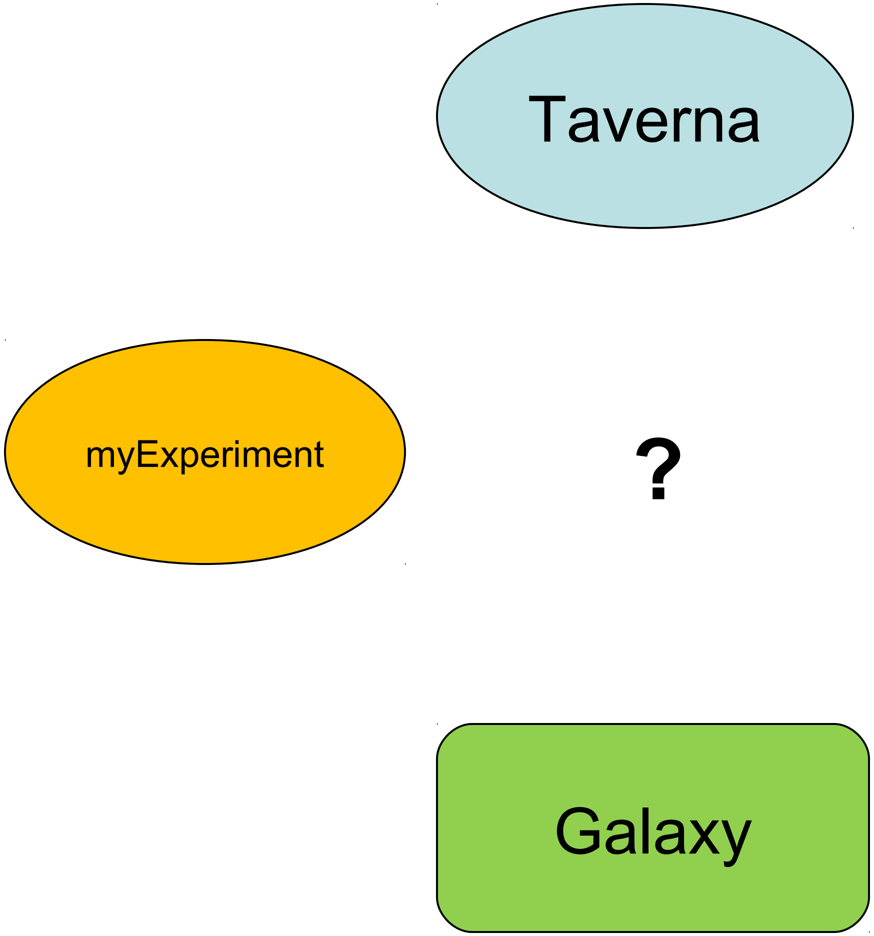
0 favourites

Taverna Workflows in Galaxy: motivation

- Galaxy
  - easy to use for biologists
  - e.g. NGS analysis, …
- Taverna
  - powerful expressive workflows
  - e.g. text mining, …
- active communities that add new functionality
  - some overlap but added value if combined
- approaches
  - incorporating Galaxy tools in Taverna
  - incorporating Taverna workflows in Galaxy

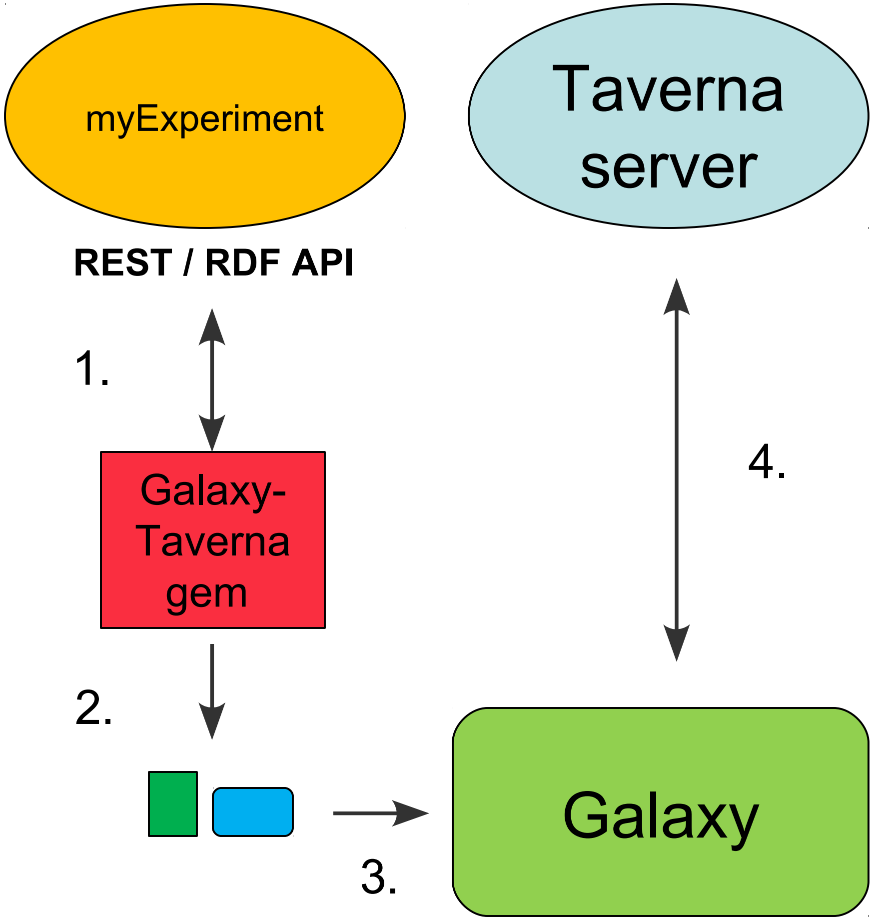
Taverna Workflows in Galaxy: components

- Taverna
  - a server (access)
- myExperiment
  - browse workflows
  - optional
- Galaxy
  - a server (admin)
  - tool to connect to Taverna
    - GUI plus config file
    - program (binary or script)

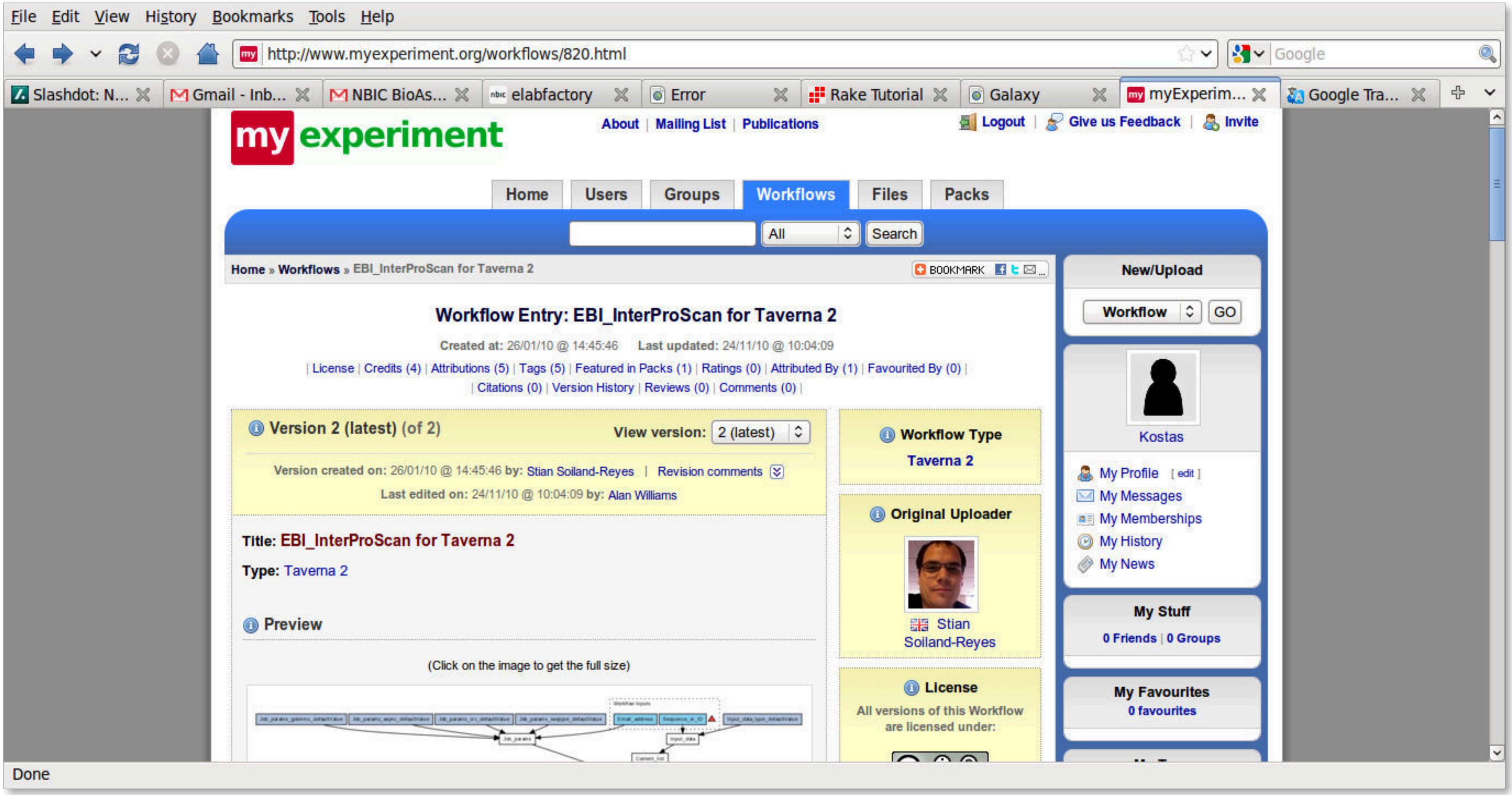


Taverna Workflows in Galaxy

- Galaxy-Taverna component
  - ruby gem
    - behind the scenes
  - generates a Galaxy tool
  - requires a workflow description
- Workflow description
  - myExperiment
  - workflow file (t2flow)
- Galaxy
  - tool needs to be manually installed



myExperiment



Home » Workflows » EBI_InterProScan for Taverna 2

Workflow Entry: EBI_InterProScan for Taverna 2

Created at: 26/01/10 @ 14:45:46 Last updated: 24/11/10 @ 10:04:09

| License | Credits (4) | Attributions (5) | Tags (5) | Featured in Packs (1) | Ratings (0) | Attributed By (1) | Favourited By (0) |
| Citations (0) | Version History | Reviews (0) | Comments (0) |

Version 2 (latest) (of 2) View version: 2 (latest)

Version created on: 26/01/10 @ 14:45:46 by: Stian Soiland-Reyes | Revision comments

Last edited on: 24/11/10 @ 10:04:09 by: Alan Williams

Title: EBI_InterProScan for Taverna 2

Type: Taverna 2

Preview

(Click on the image to get the full size)

Workflow Type

Taverna 2

Original Uploader

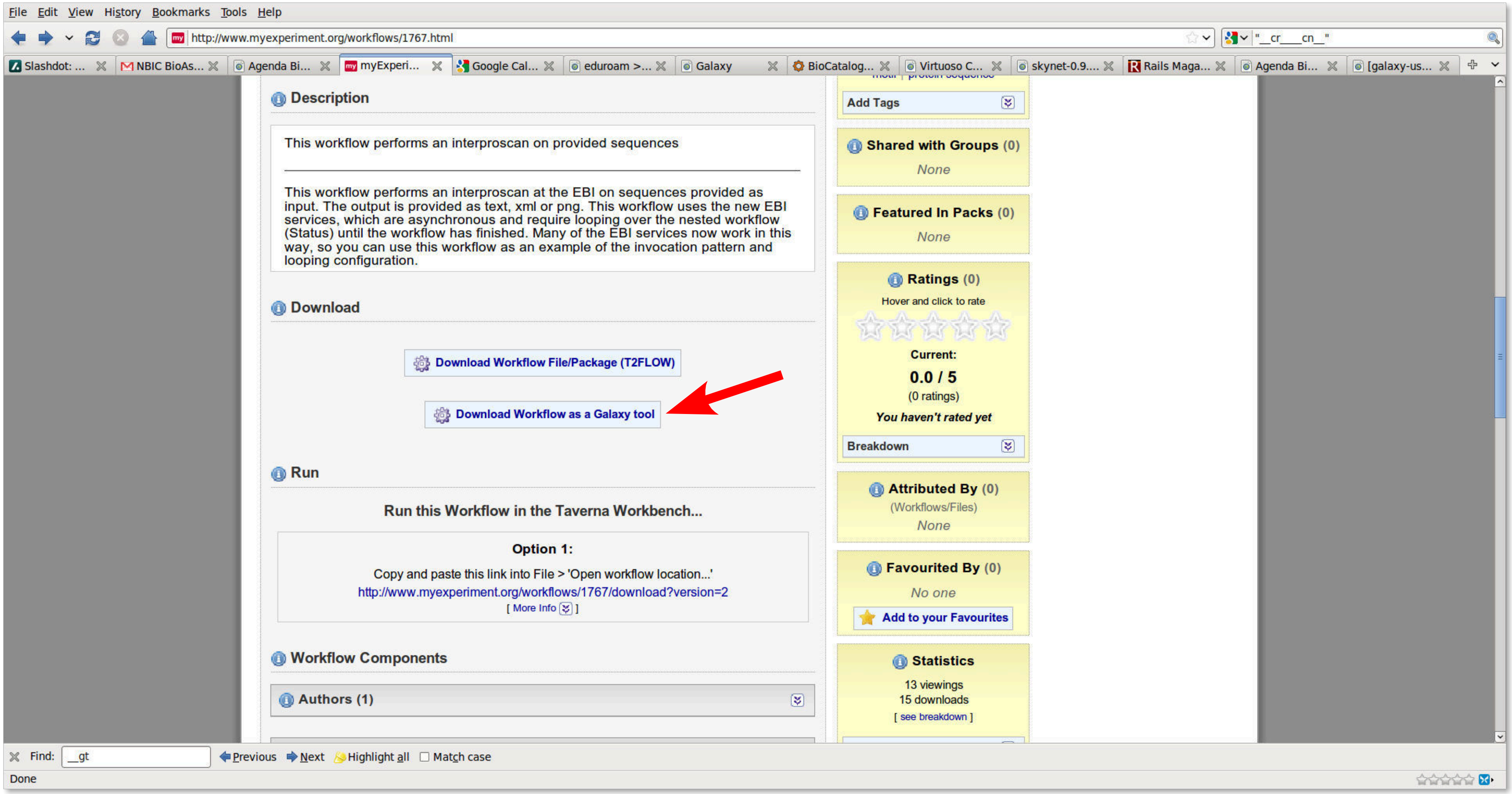
Stian Soiland-Reyes

License

All versions of this Workflow are licensed under:

Netherlands
Bioinformatics
Centre

Download Workflow as a Galaxy Tool



Description

This workflow performs an interproscan on provided sequences

This workflow performs an interproscan at the EBI on sequences provided as input. The output is provided as text, xml or png. This workflow uses the new EBI services, which are asynchronous and require looping over the nested workflow (Status) until the workflow has finished. Many of the EBI services now work in this way, so you can use this workflow as an example of the invocation pattern and looping configuration.

Download

Download Workflow File/Package (T2FLOW)

Download Workflow as a Galaxy tool

Run

Run this Workflow in the Taverna Workbench...

Option 1:

Copy and paste this link into File > 'Open workflow location...'

http://www.myexperiment.org/workflows/1767/download?version=2

[More Info]

Workflow Components

Authors (1)

Add Tags

Shared with Groups (0)

None

Featured In Packs (0)

None

Ratings (0)

Hover and click to rate

Current:

0.0 / 5

(0 ratings)

You haven't rated yet

Breakdown

Attributed By (0)

(Workflows/Files)

None

Favourited By (0)

No one

Add to your Favourites

Statistics

13 viewings

15 downloads

[see breakdown]

Taverna Workflows in Galaxy: requirements

- Galaxy server
  - admin access
  - Ruby environment and two gems
    - `$ sudo apt-get install ruby rubygems ruby1.8-dev libxml2-dev`
    - `$ sudo gem install t2-server --version 0.5.3`
    - `$ sudo gem install rubyzip`
- Taverna server
  - for testing, a taverna server is provided
- You can try it out (www.myexperiment.org)
  - trivial if you already have a Galaxy installation

Demo usage

- ISMB Technology track
  - Enacting Taverna Workflows through Galaxy
    - Lounge 1 on Monday, July 18: 12:15 p.m. - 12:40 p.m.

Future Work

- Taverna-Galaxy
  - integrate myExperiment as an external display application
  - requires dynamic loading of new tools
- Taverna workflows via a web interface
  - a web application to configure and run a workflow

- More information
  - https://trac.nbic.nl/elabfactory/wiki/eGalaxy
  - http://galaxy.psu.edu/
  - http://www.taverna.org.uk/
  - http://www.myexperiment.org/
- Questions?
  - kostas.karasavvas@nbic.nl