

# Extension of Galaxy to Utilize Web Services and A Semantic Suggestion Engine

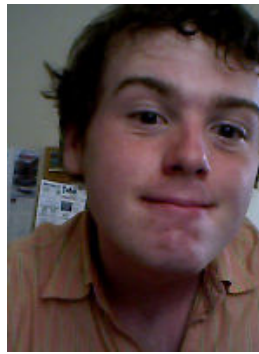
Alok Dhamanaskar, Akshay Choche, Michael E. Cotterell, Jie Zheng,  
Christian Stoeckert Jr., Jessica C. Kissinger and John A. Miller  
University of Georgia and University of Pennsylvania, USA



Alok  
Dhamanaskar



Akshay  
Choche



Michael  
Cotterell



Chris  
Stoeckert



John  
Miller

# The Motivation:

Galaxy is currently restricted to tools/data resources that are wrapped specifically for Galaxy and manually added either via and tool wrapper and addition to tools\_conf.xml or addition via the tool shed.

Wouldn't it be nice if Galaxy could utilize any tool that is publicly exposed via a Web service with just a click?


# What are Web services?

- A web service is any piece of software that makes itself available over the internet and uses a standardized XML messaging system. XML is used to encode all communications to a web service. For example, a client invokes a web service by sending an XML message, then waits for a corresponding XML response. Because all communication is in XML, web services are not tied to any one operating system or programming language--Java can talk with Perl; Windows applications can talk with Unix applications.
- Web services come in two flavors:
  - SOAP services described by WSDL XML documents
  - RESTful services described by WADL XML documents

## The BioCatalogue: providing a curated catalogue of Life Science Web Services


### Latest Activity

#### Last 7 days

 [Seyed Mahmoud Hashemi](#)  
 joined the BioCatalogue




#### Older

-  [EMBL-EBI Support](#) **added** a tag annotation to Service: [JDispatcherService](#)
-  [EMBL-EBI Support](#) **added** a publication annotation to the Soap Service of Service: [JDispatcherService](#)
-  [EMBL-EBI Support](#) **added** a publication annotation to the Soap Service of Service: [JDispatcherService](#)
-  [EMBL-EBI Support](#) **added** a contact annotation to the Service Deployment of Service: [JDispatcherService](#)
-  [EMBL-EBI Support](#) **added** a usage condition annotation to

The BioCatalogue currently has **2486 services**, **188 service providers** and **753 members** 

*"Web Services are hard to find"*


### DISCOVER

-  Find the right Web Service
-  Powerful search and filtering
-  Information from providers and community

[More info](#)

*"My Web Services are not visible"*




### REGISTER

-  Easily register Web Services
-  Instantly available to everyone
-  Providers can advertise, describe and monitor their Services

[More info](#)

*"Web Services are poorly described"*




### ANNOTATE

-  Anyone can describe and annotate
-  Ongoing expert curation
-  Social curation by the community

[More info](#)

*"Web Services are volatile"*

### MONITOR

-  Services change and get outdated
-  BioCatalogue monitors Services
-  Monitors availability and reliability

[More info](#)

### Site Announcements

**Soaplab EMBOSSE services end of life: 31st January 2013**

By [Robert Haines](#) (7 months ago)

**KEGG WSDL/SOAP services end of life: 31st December 2012**

By [Robert Haines](#) (7 months ago)

**BioCatalogue Scheduled Downtime 14th August 2012**

By [Robert Haines](#) (10 months ago)

**Many DDBJ web services unavailable until further notice**

By [Robert Haines](#) (about 1 year ago)

**Seven Web Services Specifications are Proposed Recommendations (Call for Review)**

By [Robert Haines](#) (about 1 year ago)

[More](#)

<http://www.biocatalogue.org>

# Enabling Galaxy to Invoke Web services

## Prerequisites:

- SSH/Terminal Access to Galaxy Installation
- Java 7 SE or EE (Oracle or OpenJDK)
- Python 2.6+
- JPytype Python Library



# Enabling Galaxy to Invoke Web services

Installation is Easy!

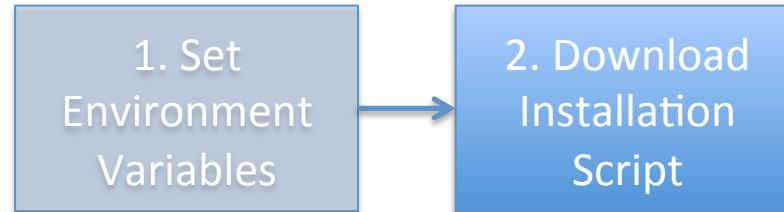
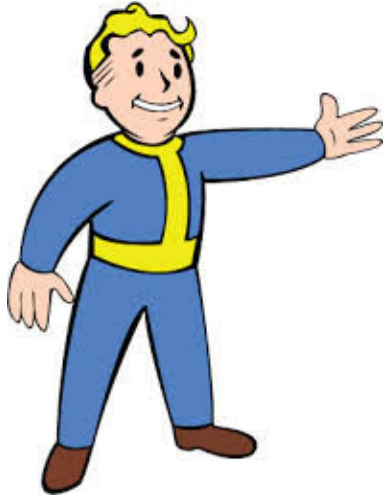


1. Set  
Environment  
Variables

```
galaxy@galaxy:~/galaxy-dist$ export JAVA_HOME=/usr/lib/jvm/java-7-openjdk-amd64
galaxy@galaxy:~/galaxy-dist$ export GALAXY_HOME=$(pwd)
galaxy@galaxy:~/galaxy-dist$
```

# Enabling Galaxy to Invoke Web services

Installation is Easy!



```
galaxy@galaxy:~/galaxy-dist$ cd
galaxy@galaxy:~$ wget http://mango.ctegd.uga.edu/jkissingLab/SWS/Galaxy
--2013-06-21 12:06:16-- http://mango.ctegd.uga.edu/jkissingLab/SWS/Gal
Resolving mango.ctegd.uga.edu (mango.ctegd.uga.edu)... 128.192.99.13
Connecting to mango.ctegd.uga.edu (mango.ctegd.uga.edu)|128.192.99.13|:
HTTP request sent, awaiting response... 200 OK
Length: 41656528 (40M) [application/zip]
Saving to: `WSExtensionTool.zip'

100%[=====

2013-06-21 12:06:17 (111 MB/s) - `WSExtensionTool.zip' saved [41656528/

galaxy@galaxy:~$ unzip WSExtensionTool.zip
```



# Enabling Galaxy to Invoke Web services

Installation is Easy!



```
galaxy@galaxy:~$ cd WS\ Extension\ Tool/  
galaxy@galaxy:~/WS Extension Tool$ python install.py  
Preparing to install the new WS Extension tool  
Copied the folder 'WebServiceToolWorkflow_REST_SOAP' to /home/galaxy/galaxy  
The "Add Web Service Tool" added successfully  
The "Web Service Tools" added successfully  
The section "Web Service Workflow Tools" is added successfully  
Workflow disabling Code added to __init__.py  
Refresh Code added to __init__.py  
Refresh code is added to app.py  
****Installation of tool is complete. Now reload galaxy to use the installed tool  
galaxy@galaxy:~/WS Extension Tool$
```

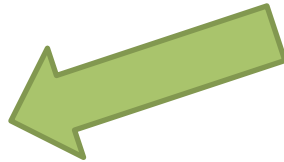


# Enabling Galaxy to Invoke Web services

Installation is Easy!



Now the Web Service tool has been installed and added to Galaxy's tool\_conf.xml file.



**However, you need to  
Restart Galaxy  
for tool\_conf.xml changes to take effect!**

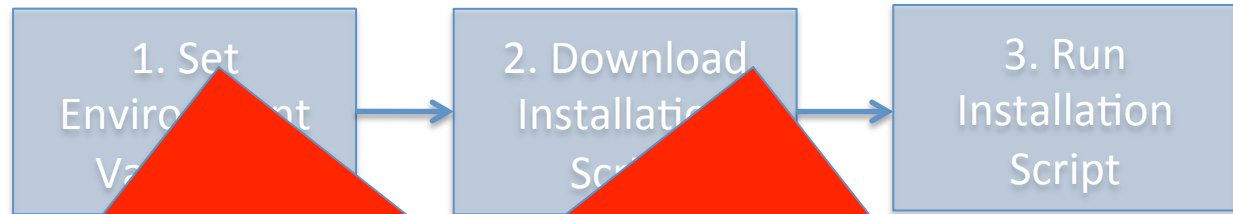
## WEB SERVICE TOOLS

### Add Web service tool(s)

- Step a: Enter the URL (WSDL/WADL location) of tool (web service) description document
- Step b: Select desired Web Service Tools / Operations to be added.
- Step c: Add tool(s) to Galaxy

# Enabling Galaxy to Invoke Web services

Installation is Easy!



tool  
has  
added  
file.

tool\_conf.  
xml

Web  
Service  
Tool

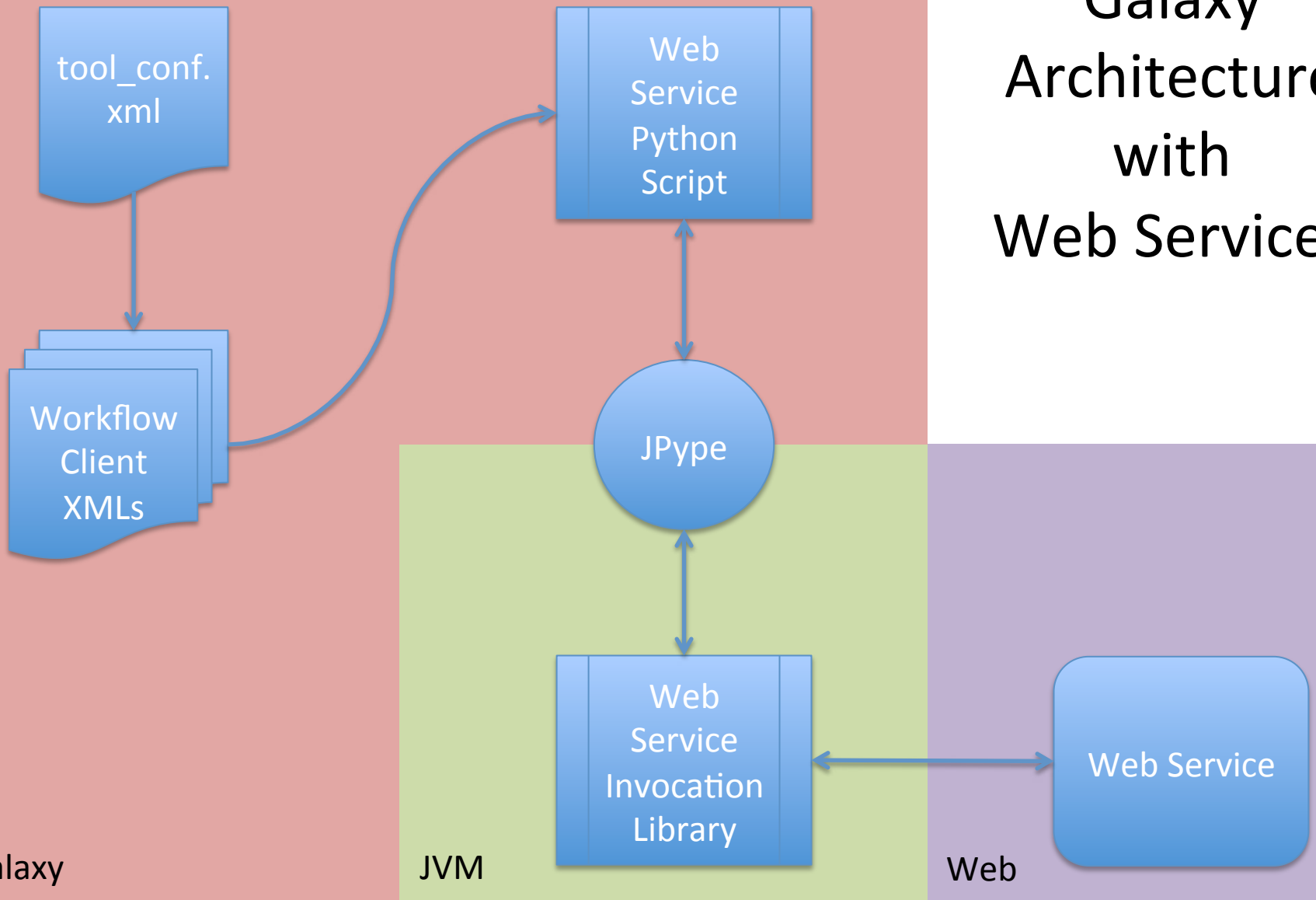
## WEB SERVICE TOOLS

### Add Web service tool(s)

- Step a: Enter the URL (WSDL/WADL location) of tool (web service) description document
- Step b: Select desired Web Service Tools / Operations to be added.
- Step c: Add tool(s) to Galaxy

However, you need to  
**Restart Galaxy**  
for tool\_conf.xml changes to take effect!

# Galaxy Architecture with Web Services



# Using Web Services in Galaxy

## WEB SERVICE TOOLS

### Add Web service tool(s)

- Step a: Enter the URL (WSDL/WADL location) of tool (web service) description document
- Step b: Select desired Web Service Tools / Operations to be added.
- Step c: Add tool(s) to Galaxy

Step a: Enter the URL (WSDL/WADL location) (version 1.1.0)

Enter the url of the tool's description document:

<http://mango.ctegd.uga.edu/jkissingLab/SWS/Wsanno>

Execute



The following job has been successfully added to the queue:

### **30: Step a: Enter the URL (WSDL/WADL location)**

You can check the status of queued jobs and view the resulting data by refreshing the **History** pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

# Using Web Services in Galaxy

## WEB SERVICE TOOLS

### Add Web service tool(s)

- Step a: Enter the URL (WSDL/WADL location) of tool (web service) description document
- Step b: Select desired Web Service Tools / Operations to be added.
- Step c: Add tool(s) to Galaxy

## Step b: Select desired Web Service Tools / Operations (version 1.1.0)

### 1. The previous step was (No action required) :

30: Step a: Enter the.. location) ⇅

You can also select any previously performed Step a from the history to avoid repeated execution of Step a.

### 2. The tool specified by the Web service is:



<http://mango.ctegd.uga.edu/jkissingLab/SWS/Wsannotation/resources/clustal>  
Displays the tool (Web service) chosen in Step a. No action required.

### 3. Select functions of the Web service tool to add to galaxy:

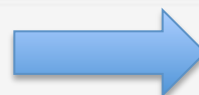
Select All

Unselect All

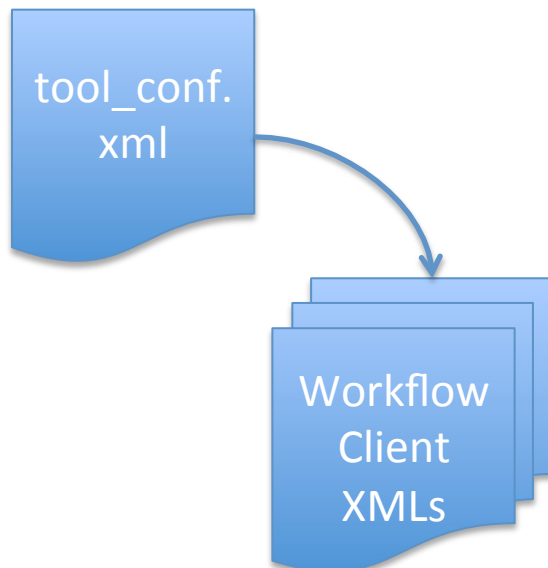
- ☐ run
- ☐ getStatus
- ☐ getResultTypes
- ☐ getResult
- ☐ getParameters
- ☐ getParameterDetails

Some tools have more than one function. It is important to select the appropriate functions.

Execute



**Generates Workflow Client XML Files**





# Using Web Services in Galaxy

## WEB SERVICE TOOLS

### Add Web service tool(s)

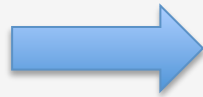
- Step a: Enter the URL (WSDL/WADL location) of tool (web service) description document
- Step b: Select desired Web Service Tools / Operations to be added.
- Step c: Add tool(s) to Galaxy

Step c: Add tool(s) to Galaxy (version 1.1.0)

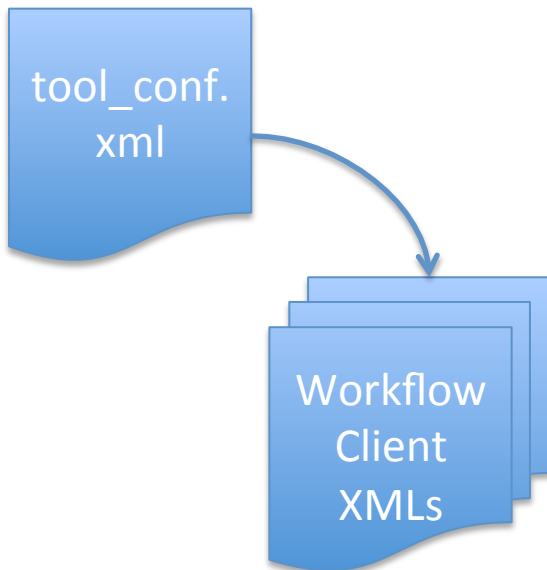
Execute to :

☒ Add Tool

Execute



**Does Nothing!**



Need to  
**Restart Galaxy**  
for tool\_conf.xml changes to take effect





# Using Web Services in Galaxy

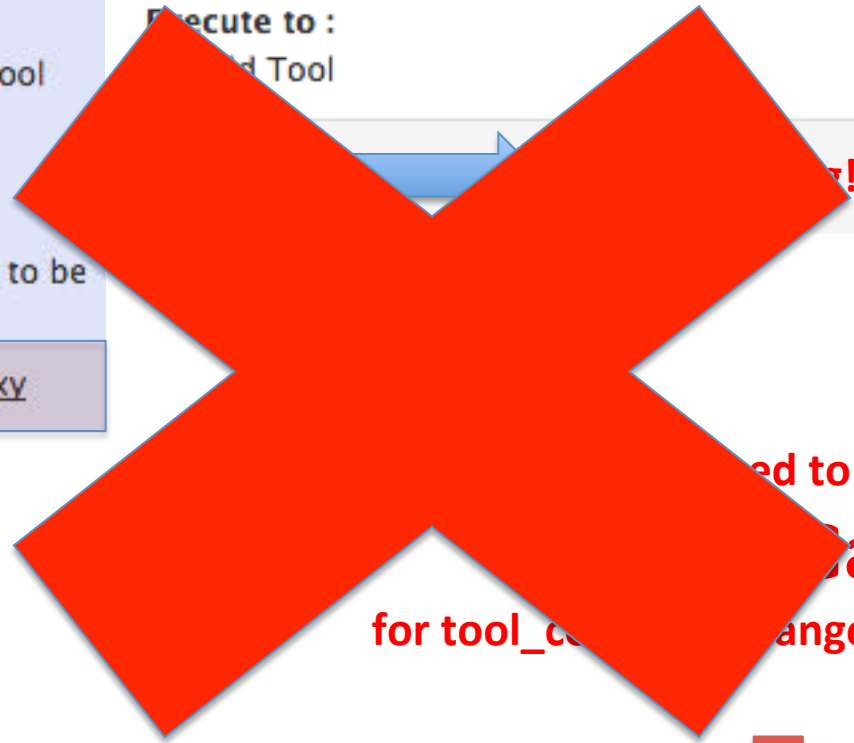
## WEB SERVICE TOOLS

### Add Web service tool(s)

- Step a: Enter the URL (WSDL/WADL location) of tool (web service) description document
- Step b: Select desired Web Service Tools / Operations to be added.
- Step c: Add tool(s) to Galaxy

Step c: Add tool(s) to Galaxy (version 1.1.0)

Execute to :  
Add Tool



ed to  
**Galaxy**  
for tool\_c...anges to take effect

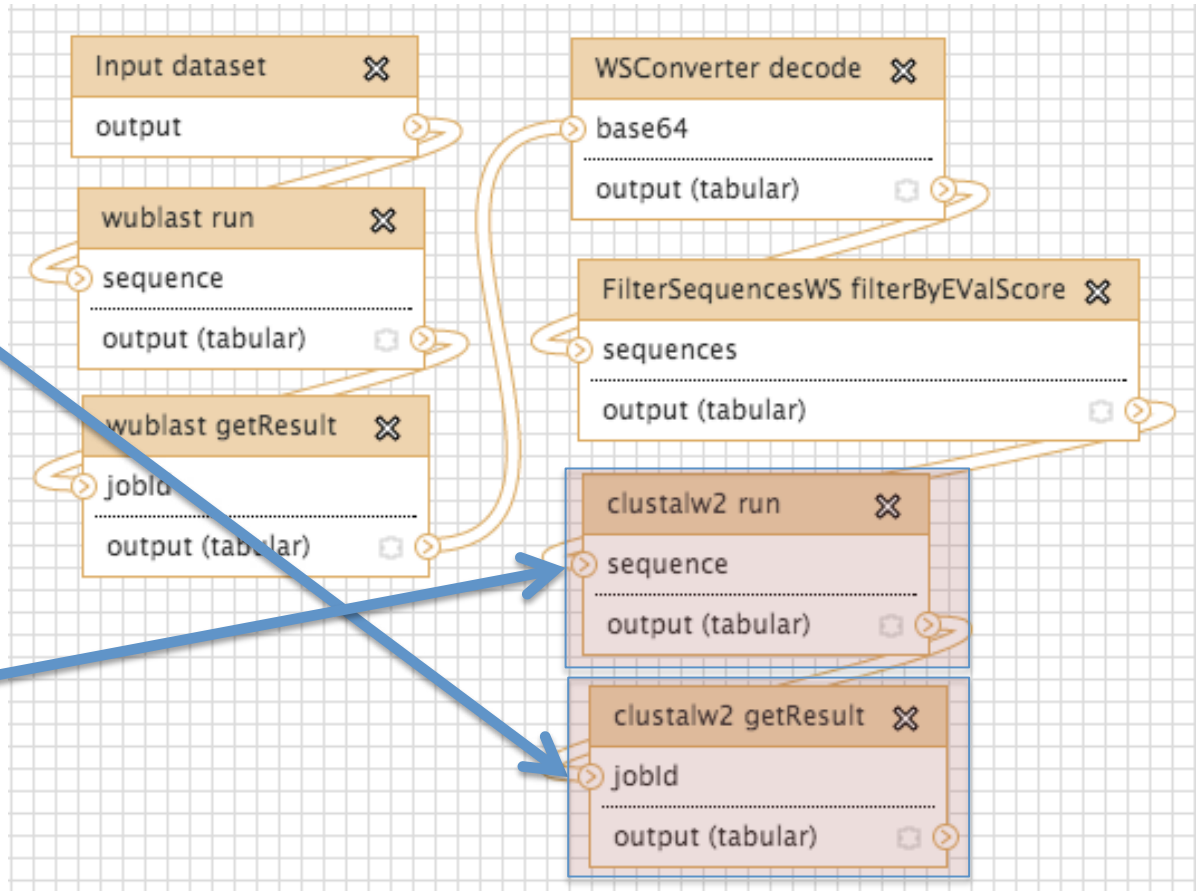
tool\_conf.  
xml

Workflow  
Client  
XMLs

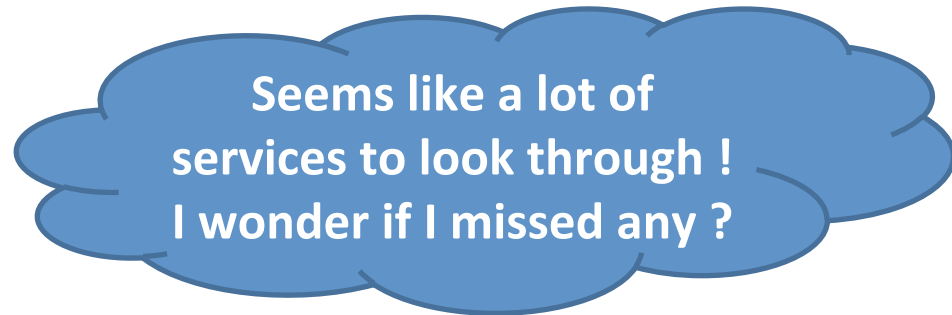
# Web Services can be used in Workflows

## Select Web Service Workflow Tool

- WSConverter encode  
URL\$<http://mango.ctegd.uga.edu/>
- WSConverter decode  
URL\$<http://mango.ctegd.uga.edu/>
- clustalw2 getResult  
URL\$<http://mango.ctegd.uga.edu/>
- clustalw2 getResultTypes  
URL\$<http://mango.ctegd.uga.edu/>
- clustalw2 getStatus  
URL\$<http://mango.ctegd.uga.edu/>
- clustalw2 run  
URL\$<http://mango.ctegd.uga.edu/>



# Find Services / Tools for My Workflow



**Service Discovery**

**Sequence Alignment Search**

Total Services = 483  
with maybe 1000s of  
operations

#354

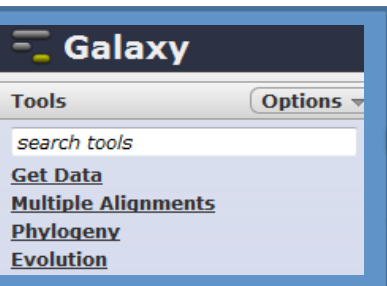
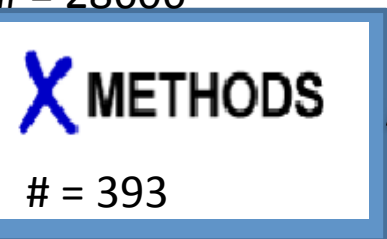
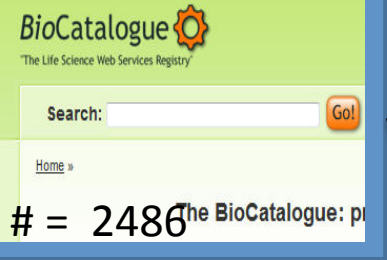
#93

#0

#36

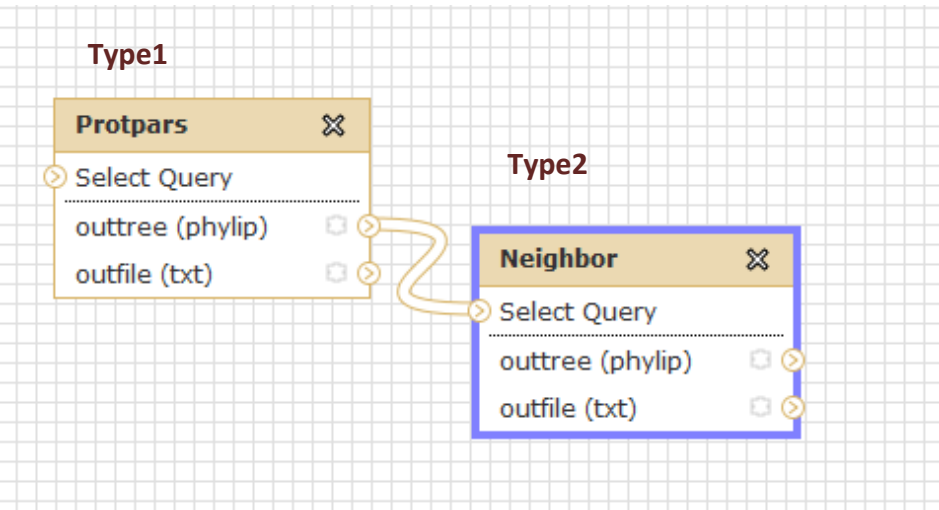
Popular Web service  
Registries

Tools available in Galaxy +  
Tools in the Galaxy Tool-shed



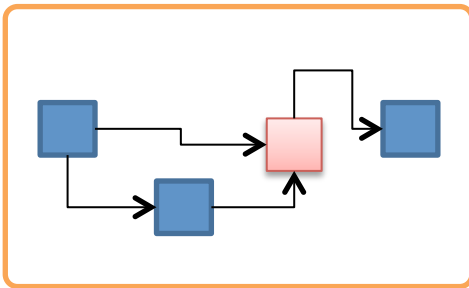
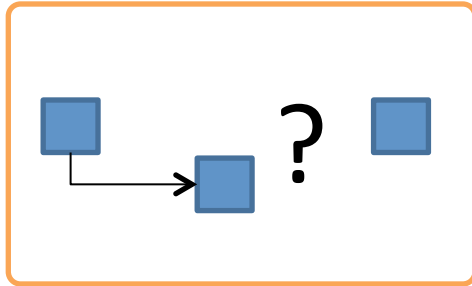
# Find Input-Output Compatible Services

I selected 2 of those operations, entered them in my workflow designer, but when I tested, it didn't work.



**Input output Matching Problem**

# Connect the Newly Added Service into the Workflow



Wish I could figure out which service operation(s) might fill the gap in my workflow.

Why can't the software make a **suggestion?**



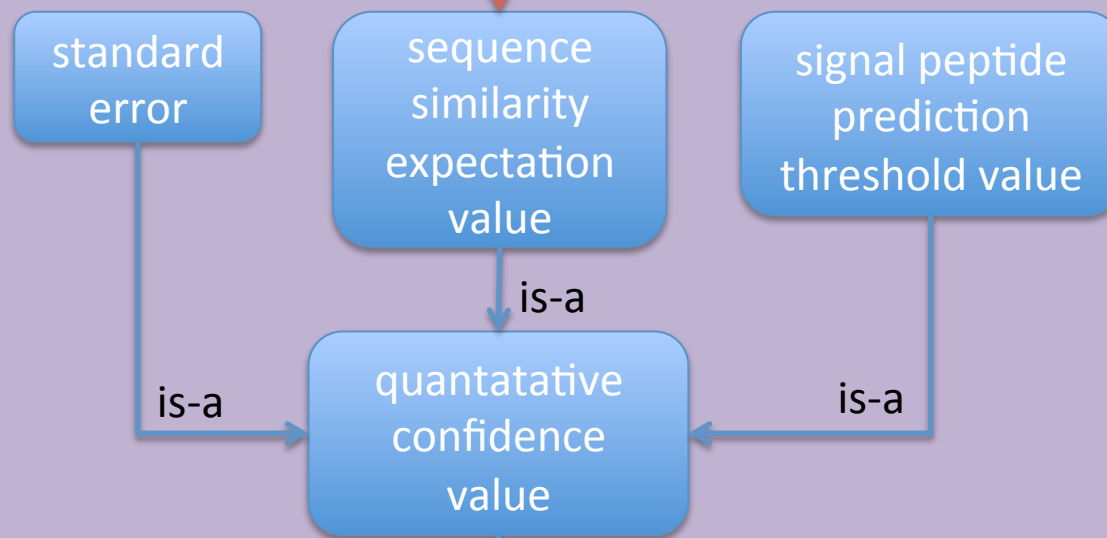
Workflow design / service composition

# Semantically Annotated Web Services

## wublast.sawsdl – Semantically Annotated WSDL for WU-BLAST

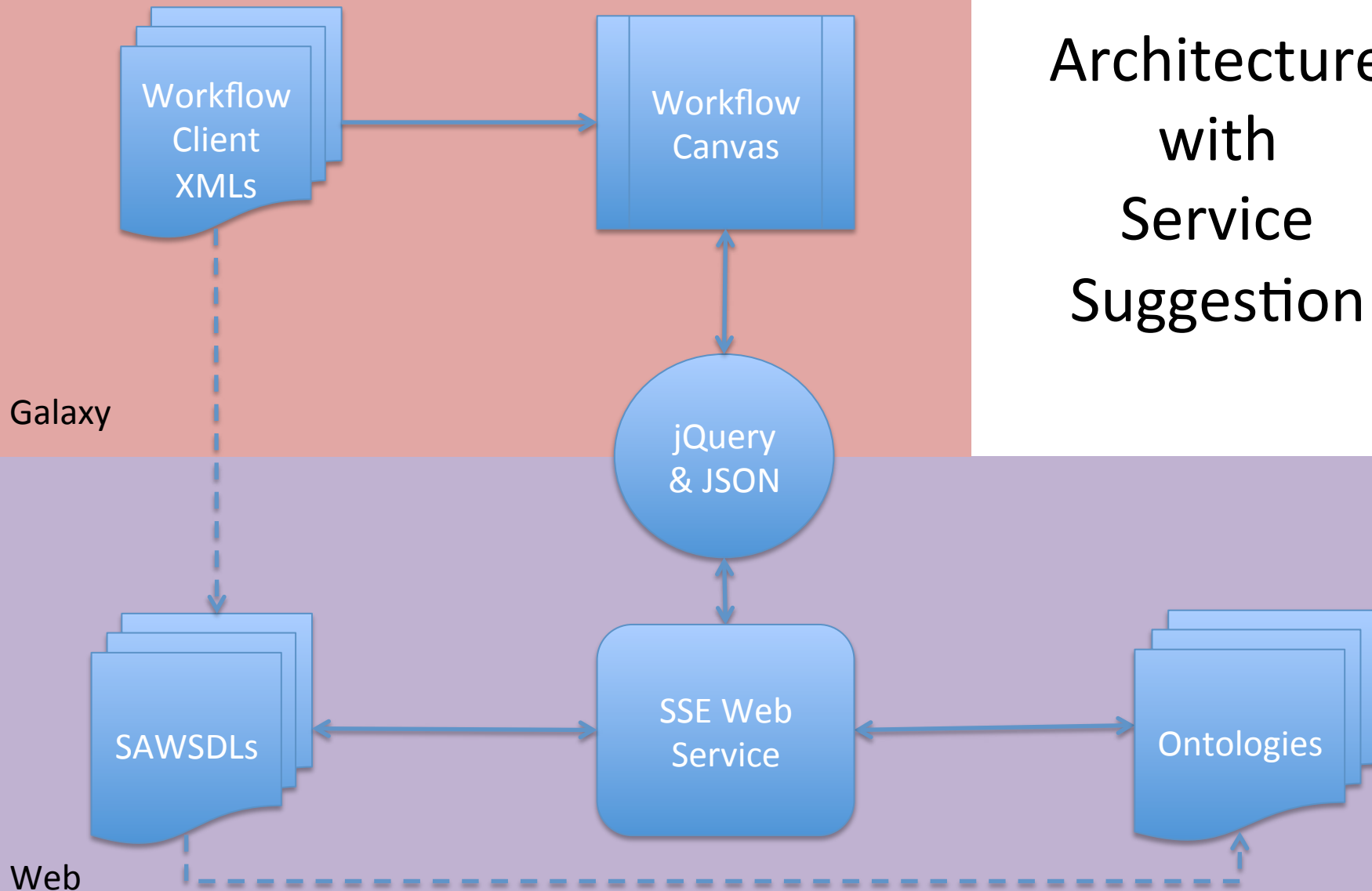
```
▼<xsd:element minOccurs="0" maxOccurs="1" name="exp" nillable="true" type="xsd:string"
sawsdl:modelReference="http://purl.obolibrary.org/obo/OBIws_0000082">
  ▼<xsd:annotation>
    ▼<xsd:documentation xml:lang="en">
      Expectation value threshold [Limits the number of scores and alignments reported
      times the match is expected to occur by chance.]
    </xsd:documentation>
  </xsd:annotation>
</xsd:element>
```

[http://purl.obolibrary.org/obo/OBIws\\_0000082](http://purl.obolibrary.org/obo/OBIws_0000082)

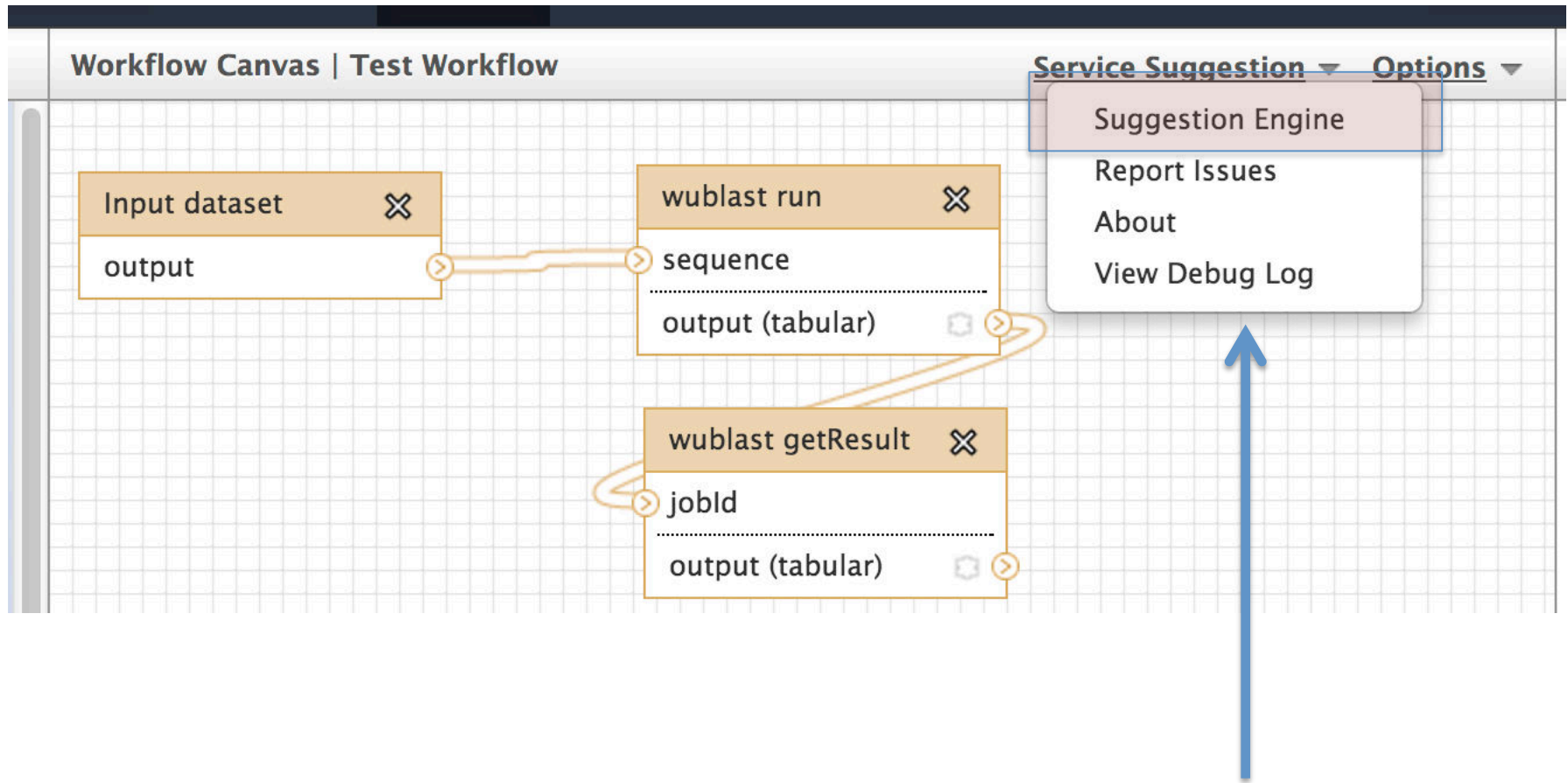




# Architecture with Service Suggestion



# Service Suggestion



Modified the UI using minimal mako modifications  
and some jQuery magic!

# Service Suggestion

1. Choose the type of suggestion you want



**Forward Suggestion**



**Backward Suggestion**



**Bidirectional Suggestion**



## Details

### Service Suggestion Engine

Version: 1.6

**[ - ] Setup & Run Query**

#### Choose Suggestion Type:

Choose the kind of suggestions you want by indicating where you think they should go in the workflow.

Forward Suggestion ▾



**i** Click [here](#) for more information about the different types of service suggestion.

#### Tools to Feed From:

Select the tools in current workflow that you think should feed into suggested service operations.

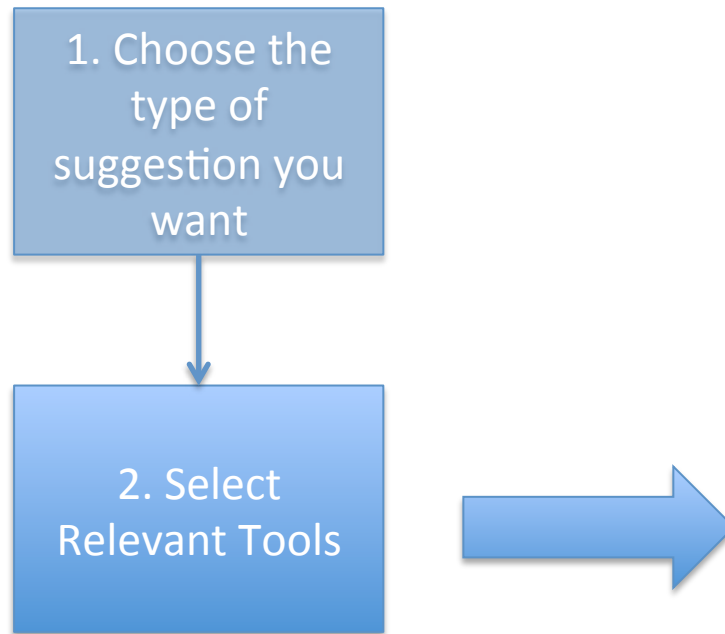
Step 2 – wublast getResult ✕

Step 2 – wublast getResult ▾

#### Goal / Purpose:

In simple terms, describe what you are trying to do/find (e.g., multiple

# Service Suggestion



The tools currently on the canvas will appear in one or more dropdown lists (depending on suggestion type).

Select specific ones to utilize for suggestion or choose “all”.

## Choose Suggestion Type:

Choose the kind of suggestions you want by indicating where you think they should go in the workflow.

Forward Suggestion ▾



**i** Click [here](#) for more information about the different types of service suggestion.

## Tools to Feed From:

Select the tools in current workflow that you think should feed into suggested service operations.

Step 2 – wublast getResult ✕

Step 2 – wublast getResult ▾

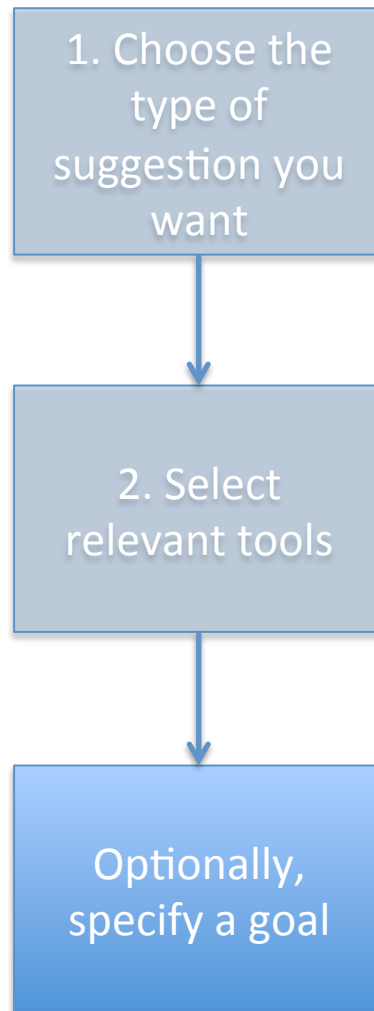
## Goal / Purpose:

In simple terms, describe what you are trying to do/find (e.g., multiple sequence alignment).

**i** Click [here](#) for more information about specifying goals.

Run Query / Get Suggestions

# Service Suggestion



Choose the kind of suggestions you want by indicating where you think they should go in the workflow.

Forward Suggestion



**i** Click [here](#) for more information about the different types of service suggestion.

## Tools to Feed From:

Select the tools in current workflow that you think should feed into suggested service operations.

Step 2 – wublast getResult ✕

Step 2 – wublast getResult

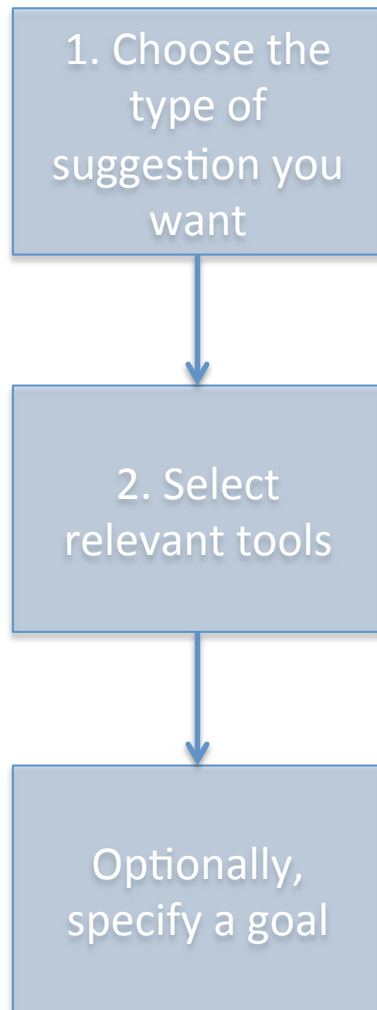
## Goal / Purpose:

In simple terms, describe what you are trying to do/find (e.g., multiple sequence alignment).

**i** Click [here](#) for more information about specifying goals.

Run Query / Get Suggestions

# Service Suggestion




**If you're not sure what something means then click on "help" to display a modal help dialog.**

Choose the kind of suggestions you want by indicating where you think they should go in the workflow.

Forward Suggestion ▾



 Click [here](#) for more information about the different types of service suggestion.

## Tools to Feed From:


Select the tools in current workflow that you think should feed into suggested service operations.

Step 2 – wublast getResult ✕

Step 2 – wublast getResult ▾

## Goal / Purpose:

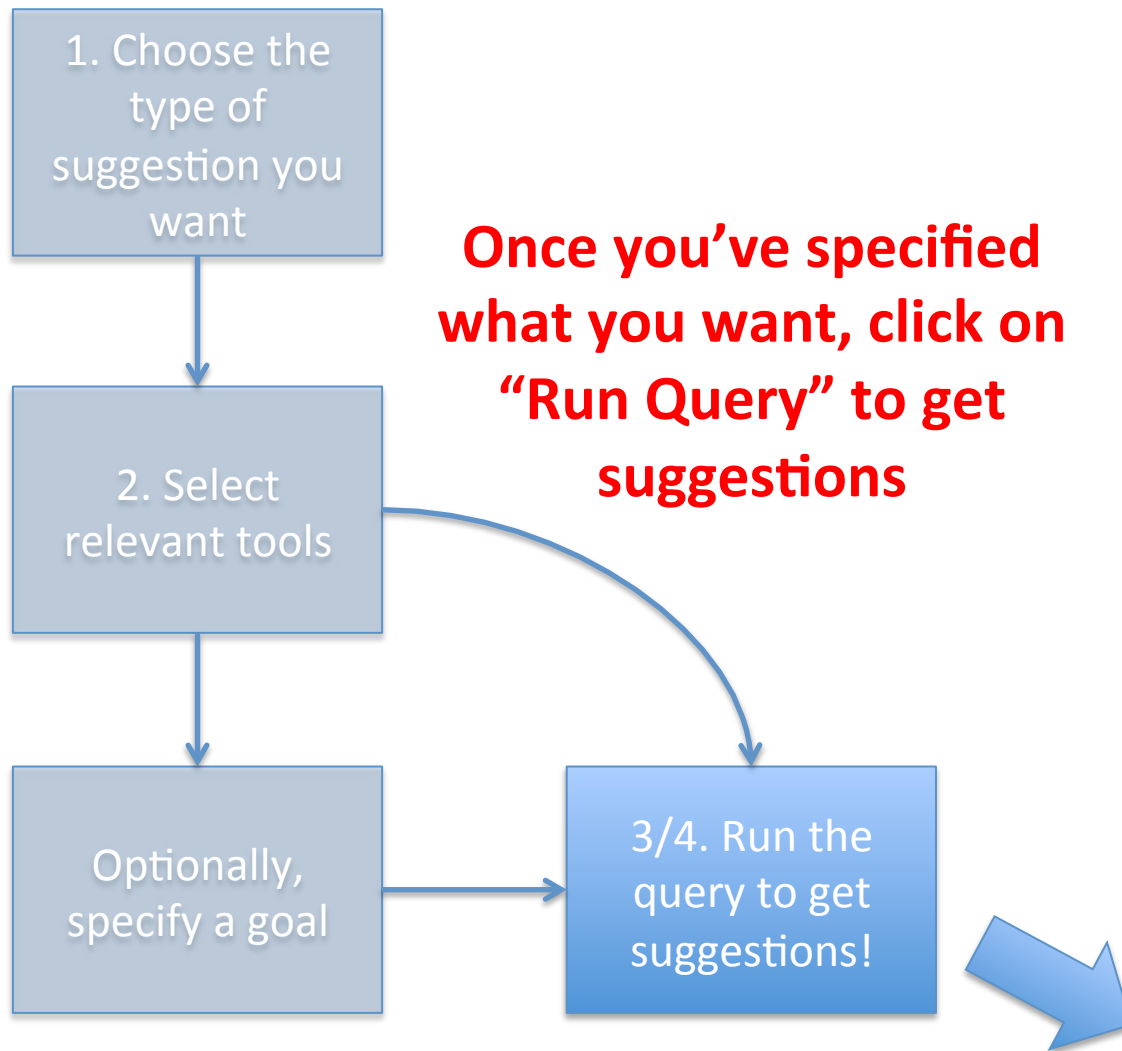
In simple terms, describe what you are trying to do/find (e.g., multiple sequence alignment).

 Click [here](#) for more information about specifying goals.

Run Query / Get Suggestions



# Service Suggestion



Choose the kind of suggestions you want by indicating where you think they should go in the workflow.

Forward Suggestion



**i** Click [here](#) for more information about the different types of service suggestion.

## Tools to Feed From:

Select the tools in current workflow that you think should feed into suggested service operations.

Step 2 - wublast getResult ✕

Step 2 - wublast getResult

## Goal / Purpose:

In simple terms, describe what you are trying to do/find (e.g., multiple sequence alignment).

**i** Click [here](#) for more information about specifying goals.

Run Query / Get Suggestions

Service Suggestion ▾ Options ▾

Input dataset ✕  
output

wublast run ✕  
sequence  
output (tabular)

wublast getResult ✕  
jobId  
output (tabular)

Details

Service Suggestion Engine

Version: 1.6

[ - ] Setup & Run Query

[ - ] Query Results

Ranked Suggestion Results

These are SEE suggested tools. Click on a tool to add it to your workflow. Tools are ranked by a relevance score (1->0).

Showing 5 of 9 results. | [Show All](#)


▪ [WSConverter decode](#)  
0.3924244219779627

▪ [clustalw2 run](#)  
0.22178261375220856

▪ [WSConverter encode](#)  
0.2202941177288691

▪ [wublast run](#)  
0.20485147913275623

▪ [clustalw2 getResult](#)  
0.18738075580970934

 Click [here](#) for more information about how results are ranked.

Annotated services are displayed in the results, ranked by how well they satisfy the suggestion query.

Service Suggestion ▾Options ▾

Input dataset ✕  
output

WSConverter decode ✕  
base64  
output (tabular)

wublast run ✕  
sequence  
output (tabular)

wublast getResult ✕  
jobId  
output (tabular)

Service Suggestion Engine

Version: 1.6

[ - ] Setup & Run Query

[ - ] Query Results

Ranked Suggestion Results

These are SEE suggested tools. Click on a tool to add it to your workflow. Tools are ranked by a relevance score (1->0).

Showing 5 of 9 results. | [Show All](#)

WSConverter decode

0.3924244219779627

clustalw2 run

0.22178261375220856

WSConverter encode

0.2202941177288691

wublast run

0.20485147913275623

clustalw2 getResult

0.18738075580970934

i

Click [here](#) for more information about how results are ranked.

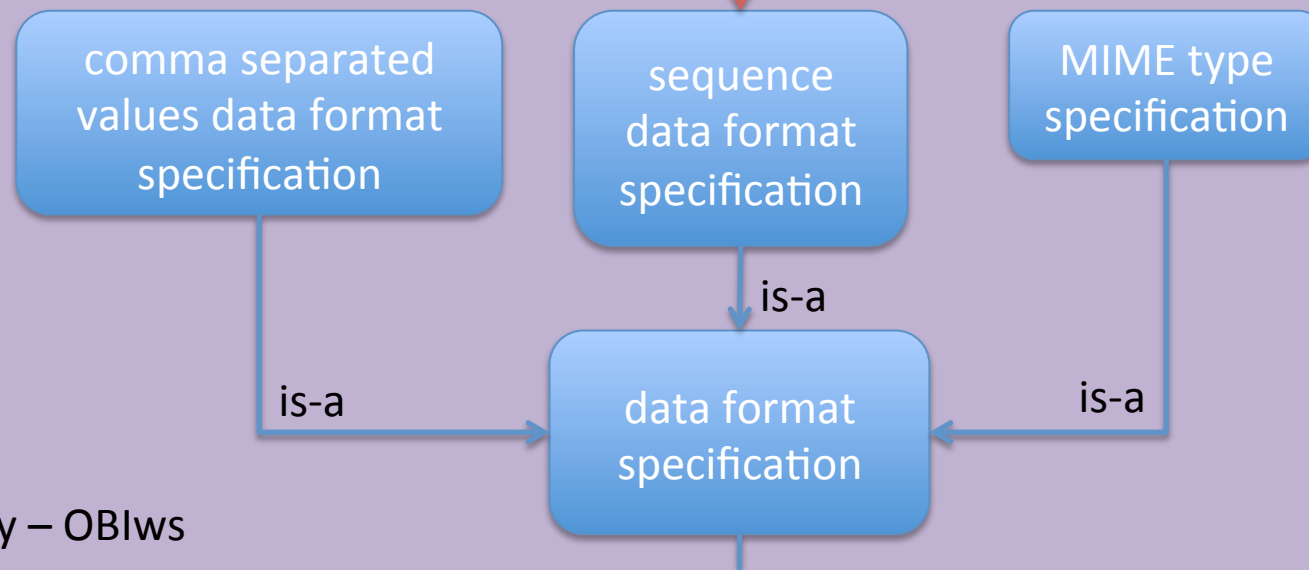
Clicking on a tool adds it to the canvas!

# Why not annotate Galaxy Tools to enable the SSE?

## tabular\_to\_fasta.xml – Example of a Semantically Annotated Galaxy Tool

```
<description>converts tabular file to FASTA format</description>
<command interpreter="python">tabular_to_fasta.py $input $title_col $seq_col $output </command>
<inputs>
  <param name="input" type="data" format="tabular" label="Tab-delimited file"/>
  <param name="title_col" type="data_column" data_ref="input" multiple="True" numerical="False" label="Title column(s)" help="Multi-
    select list - hold the appropriate key while clicking to select multiple columns"/>
  <param name="seq_col" type="data_column" data_ref="input" numerical="False" label="Sequence column" />
</inputs>
<outputs>
  <data name="output" format="fasta" modelReference="http://purl.obolibrary.org/obo/OBIws_0000106" />
</outputs>
```

**All we need to do is add modelReferences to a tool's XML file!**



# Annotating Services with RadiantWeb

Simplified View  
of  
XML File

Tree View  
of  
Ontology

**Radiant Web - Semantic Annotation Tool**

**WSDL/WADL Viewer**

WSDL location:

C:\fakepath\wublast.wsdl

WU-BLAST stands for Washington University Basic Local Alignment Search Tool. The emphasis of this tool is to find regions of sequence similarity or homology quickly, with minimum loss of sensitivity. This will yield functional and evolutionary clues about the structure and function of your novel sequence. Dr Warren Gish at Washington University released this first "gapped" version of BLAST allowing for gapped alignments and statistics.

Identified Operations in the WSDL (portType : JDispatcherService)

- getParameters
- getStatus
- check web service execution status objective
- inputs (getStatusRequest)
- Outputs (getStatusResponse)
  - status(\*)
  - web service execution status
- getResultTypes
- getResult
- run
- getParameterDetails

**Ontology Viewer**

Owl Location:

Search:

Definition

Ontology 1 **Ontology 2**

- Thing
  - Obsolete Class
  - entity
    - occurrent
      - continuant
        - spatial region
        - dependent continuant
        - independent continuant

**Legend**

✓ Approve Suggested Term ✗ Reject Suggested Term ✕ Remove annotation Suggested Term Pre-existing annotation Approved Term Operation  
Messages (Input/Output) ComplexType (Inputs/Outputs) SimpleType (Inputs/Outputs)

Version 1.0 The project is developed at [The University of Georgia](#). For feedback and suggestions [contact us](#)



# Annotating Services with RadiantWeb

## Annotation Recommendations

The screenshot displays the Radiant Web Semantic Annotation Tool interface. The browser address bar shows the URL `192.168.1.78:8084/radiant/sawSDL.jsp`. The page title is "Radiant Web - Semantic Annotation Tool".

The interface is divided into two main panels:

- WSDL/WADL Viewer:**
  - WSDL location: `C:\fakepath\wublast.wsdl`
  - Buttons: `Browse`, `Load WS`
  - Current WSDL: `C:\fakepath\wublast.wsdl`
  - Buttons: `Save WSDL`, `Recommend Terms`, `WSDLXML view`
  - Description: WU-BLAST stands for Washington University Basic Local Alignment Search Tool. The emphasis of this tool is to find regions of sequence similarity or homology quickly, with minimum loss of sensitivity. This will yield functional and evolutionary clues about the structure and function of your novel sequence. Dr Warren Gish at Washington University released this first "gapped" version of BLAST allowing for gapped alignments and statistics.
  - Identified Operations in the WSDL (portType : JDispatcherService):
    - `getParameters` (Operation) with `Recommend Terms` button.
    - `getStatus` (Operation) with `Recommend Terms` button.
    - `check web service execution status objective` (Suggested Term) with `Recommend Terms` button.
    - `Inputs (getStatusRequest)` (Message) with `Recommend Terms` button.
    - `Outputs (getStatusResponse)` (Message) with `Recommend Terms` button.
    - `status(*)` (Operation) with `Recommend Terms` and `Add SchemaMapper` buttons.
    - `web service execution status` (Suggested Term) with `Recommend Terms` button.
    - `getResultTypes` (Operation) with `Recommend Terms` button.
    - `getResult` (Operation) with `Recommend Terms` button.
    - `run` (Operation) with `Recommend Terms` button.
    - `getParameterDetails` (Operation) with `Recommend Terms` button.
- Ontology Viewer:**
  - Owl Location: `C:\fakepath\webService.owl`
  - Buttons: `Browse`, `Load OWL`
  - Search:  `Set Options`
  - Definition:
  - Ontology 1, Ontology 2, Ontology 3
  - Thing
    - Obsolete Class
    - entity
      - occurent
      - continuant
        - spatial region
        - dependent continuant
        - independent continuant

**Legend:**

- ✓ Approve Suggested Term
- ✗ Reject Suggested Term
- ✕ Remove annotation
- Suggested Term
- Pre-existing annotation
- Approved Term
- Operation
- Messages (Input/Output)
- ComplexType (Inputs/Outputs)
- SimpleType (Inputs/Outputs)

Version 1.0 The project is developed at [The University of Georgia](#). For feedback and suggestions [contact us](#)



# Annotating Services with RadiantWeb

Documentation

Inputs/Outputs

**Radiant Web - Semantic Annotation Tool**

**WSDL/WADL Viewer**

WSDL location:

C:\fakepath\wublast.wsdl

WU-BLAST stands for Washington University Basic Local Alignment Search Tool. The emphasis of this tool is to find regions of sequence similarity or homology quickly, with minimum loss of sensitivity. This will yield functional and evolutionary clues about the structure and function of your novel sequence. Dr Warren Gish at Washington University released this first "gapped" version of BLAST allowing for gapped alignments and statistics.

Identified Operations in the WSDL (portType : JDispatcherService)

- ☐
- ☐
- ☒ ☒
- ☐
- ☐
- ☒ ☒
- ☐
- ☐
- ☐
- ☐

**Ontology Viewer**

Owl Location:

Search:

Definition

**Ontology 1** **Ontology 2** **Ontology 3**

- ☐ Thing
  - ☐ Obsolete Class
  - ☐ entity
    - ☐ occurrent
    - ☐ continuant
      - ☐ spatial region
      - ☐ dependent continuant
      - ☐ independent continuant

**Legend**

☒ Approve Suggested Term ☒ Reject Suggested Term ☒ Remove annotation ☒ Suggested Term ☒ Pre-existing annotation ☒ Approved Term ☒ Operation

☒ Messages (Input/Output) ☒ ComplexType (Inputs/Outputs) ☒ SimpleType (Inputs/Outputs)

Version 1.0 The project is developed at [The University of Georgia](http://The University of Georgia). For feedback and suggestions [contact us](http://contact us)

# Adding Web services to Galaxy

## Advantages

- Increased tool/database access
- Don't need to wrap the tools/data access
- Distribute the load - Utilize someone else's computers
- Further empower the end user

## Disadvantages

- Finding the Web services
- Moving data
- Hidden analysis step (Get job ID)
- All possible result formats (many not compatible with existing Galaxy tools)
- Security
- Banned from sites (depends on who is viewed as the user)

# Adding a Semantic Suggestion Engine

## Advantages

- Helps users find Galaxy tools with the desired functionality
- Helps with workflow construction
- Annotated tools are better documented tools. You will now what that parameter does and understand inputs and outputs

## Disadvantages

- The community needs to annotate its tools
- Not all semantic terms will exist
- You give up some display area in the history column when the SSE is being used. It can be collapsed when not in use.

Thank you!

<http://mango.ctegd.uga.edu/jkissingLab/SWS/>