



THE NATIONAL SCIENCE DIGITAL LIBRARY

Biodiversity Analysis Pipeline

Paul Allen

Cornell Lab of Ornithology



Biodiversity Analysis Pipeline

- Web-based
 - workflow execution, like **Hydrant**
 - authoring environment, like **Yahoo! Pipes**
 - controlled “vocabulary” of actors
 - workflow sharing & community, like **myExperiment**
 - [results accessible as web resource/service](#)
- Kepler integration
 - running headless to execute workflows
 - scale by adding CPUs running Kepler

Pipes: editing 'Photos Near Napa Wineries' - Mozilla Firefox

File Edit View History Bookmarks Tools Help del.icio.us

http://pipes.yahoo.com/pipes/pipe.edit?_id=bMPFtO342xGrr53VyzUFzw

kepler project vergil

Photos Near Napa Wineries*

Layout Expand All Collapse All Back to My Pipes New Save a copy

Sources

- Fetch CSV
- Feed Auto-Discover
- Fetch Feed
- Fetch Data
- Fetch Page
- Fetch Site Feed
- Flickr
- Google Base
- Item Builder
- Yahoo! Local
- Yahoo! Search

User inputs

- Operators
- Url
- String
- Date
- Location
- Number
- Favorites
- My pipes
- Deprecated

Yahool Local

Find within of

Search near (e.g. Napa)

Name:
Prompt: Search near (e.g. Napa:
Position:
Default:
Debug:

For Each: Annotate

For each item in input feed, set attribute to output from

Flickr

Find

Rename

Mappings
- Copy As

What to Find (e.g. Wine)

Name:
Prompt: What to Find (e.g. Wine
Position:
Default:
Debug:

Rename

Mappings
- Rename

Sort

Sort by
- in order

Pipe Output

Debugger: Location Input (0 items)

Done

start

Pipes: editing 'Photos...

https://cuweblogin.ci...

9:23 AM



Search bar with dropdown menu set to 'All' and a 'Search' button.

Home » Workflows



Buttons for 'Upload New Workflow' and 'View All Workflows'.

Top 50 tags for Workflows [See All Tags]

List of tags including affymetrix, AIDA, beanshell, BioAID, bioassist_nl, bioinformatics, biomoby, biorange_nl, BLAST, cel, data-driven, ddbj, demo, design pattern, disambiguation, disease, example, floss, gene identifier, gene symbol, genotype, iteration, kegg, Kegg Pathways, microarray, mygrid, nbiconworkflows, open source software, oss, pathway, pathway-driven, pathways, pdb, phenotype, protein, protein sequence, pubmed, scaffold, sequence, shim, similarity, social sciences, taverna, test, text mining, text_mining, text_mining_network, uniprot, utility, VL-e

Latest Last Updated Most Viewed Most Downloaded

Uploader:



Saeedeh

Escherichia coli : From cDNA Microarray Raw Data to Pathways and Published Abstracts (v1)

View Download (v1)

Created: 08/05/08 @ 15:25:29 | Updated: 12/05/08 @ 09:01:37

Credits: Saeedeh Paul Fisher

Attributions: HUMAN Microarray CEL file to candidate pathways

License: Creative Commons Attribution-Share Alike 3.0 License



This workflow takes in a CDNA raw file and a normalisation method then returns a series of images/graphs which represent the same output

New/Upload

Workflow dropdown menu and GO button



Paulallen

- My Profile
My Inbox
My Memberships
My History
My News

My Stuff

0 friends | 0 groups

Done

Biodiversity Analysis Pipeline

- Audience
 - NSDL users
 - students, teachers (grades 6-12)_
 - resource/land managers (decision support tool)_
 - citizen scientists
- Things I worry about
 - big datasets
 - caching results of intermediate actors
 - versioning of workflows, actors, data
 - workflow authoring without Vergil

Biodiversity Analysis Pipeline

- Contributions
 - reduce learning curve
 - composite actors
 - sharing workflows
 - result dissemination
 - “always on” web accessible results
 - workflows as web service
- Collaboration