

Getting Started with the Kepler Workflow Run Manager

Workflow Run Manager Suite includes:

Workflow Run Manager
Provenance
Tagging

Workflow Run Manager 2.1.0

September 10, 2010

Getting Started with the Workflow Run Manager

The *Getting Started with the Kepler Workflow Run Manager* guide is for scientists who would like to use the *Workflow Run Manager* suite in Kepler.

The Workflow Run Manager was created to assist scientists who would like to use Kepler to manage workflow execution history stored on their computer by the Provenance module.

Table of Contents

1. Introduction.....	2
1.1 What is the Workflow Run Manager ?	2
2. Downloading and Installing the Workflow Run Manager	2
3. Using the Workflow Run Manager.....	3
3.1 Tagging.....	5
3.2 Searching runs.....	5
3.3 Importing runs	6

1. Introduction

The Getting Started with the Workflow Run Manager guide introduces the main components and functionality of the Workflow Run Manager. The Workflow Run Manager is an add-on module suite for Kepler, a software application for the analysis and modeling of scientific data, and it provides functionality for managing workflow execution history stored locally on your computer by the Provenance module, mainly via an additional GUI component within Kepler.

1.1 What is the Workflow Run Manager ?

The Workflow Run Manager provides a GUI for managing workflow execution history (“workflow runs”) stored by the Provenance module. When executing a workflow with provenance recording enabled, the Provenance module stores data and artifacts generated by the execution in a database on your computer (currently located in: ~/KeplerData/modules/provenance/).

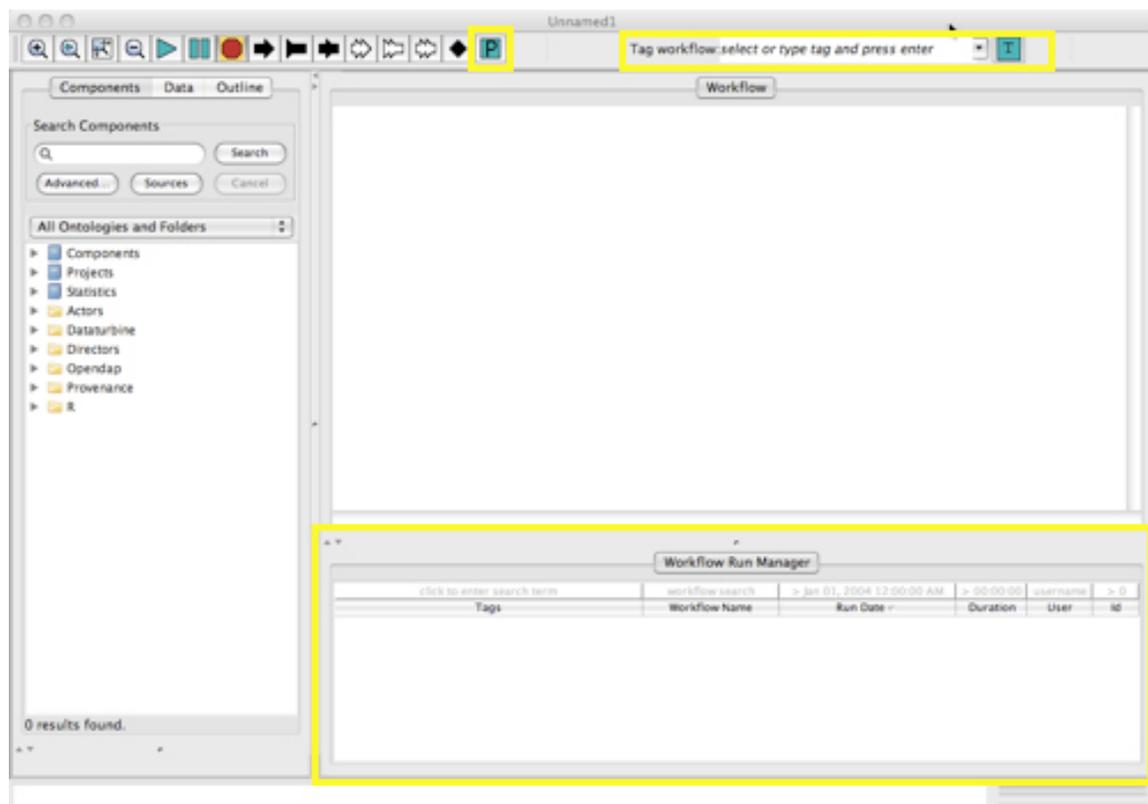
Workflow runs may searched, tagged, deleted, exported, imported, or uploaded to the Remote Repository you currently have selected. (To configure which remote repository is selected, navigate to Tools => Preferences => Components).

2. Downloading and Installing the Workflow Run Manager

From the Kepler application menu select Tools => Module Manager. From the Module Manager dialog, select the Available Modules tab. Select 'workflow-run-manager-2.1.0' from the list of Available Suites, then click the right arrow button to move the workflow-run-manager-2.1.0 suite to the list of Selected Modules. If you would like to use the Workflow Run Manager in conjunction with Reporting, instead choose the reporting-2.1.0 suite (and please also refer to the Reporting documentation). Click the 'Apply and Restart' button to retrieve the suite you have selected and restart Kepler.

3. Using the Workflow Run Manager

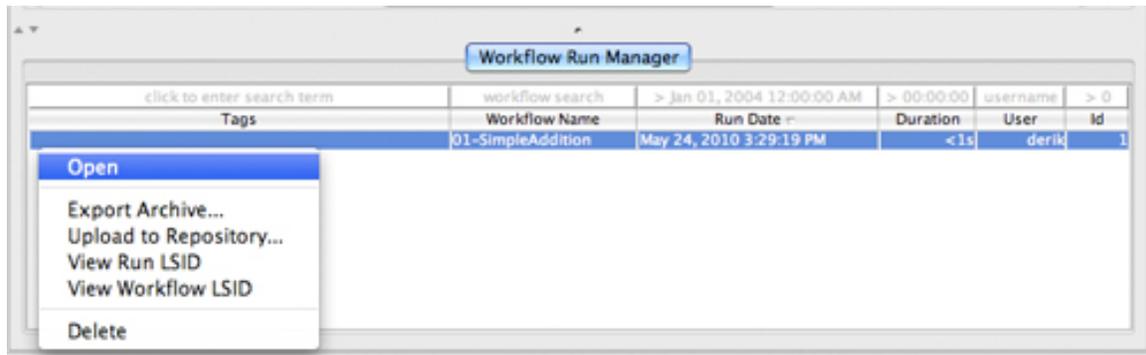
Once you have installed the Workflow Run Manager, you will notice the Kepler GUI has new items. If you're using the Workflow Run Manager suite, on restart a new Kepler window will have three additions: a toolbar button to toggle Provenance recording, a Tagging toolbar, and a Workflow Run Manager table. These are highlighted in the screenshot below.



If you are using the Reporting suite, the Workflow Run Manager will instead appear on the Reporting View. Please see the Reporting suite documentation for more details.

By default, provenance recording is on, and may be turned off by clicking the green "P" provenance button in the toolbar. The button will turn red, denoting recording is now off. When on, for every workflow execution, a row will appear in the Workflow

Run Manager table. Selecting rows and then right-clicking on rows will bring up a context menu from which you may Open, Export, or Delete a run:



Selecting **Open** will open the workflow associated with this execution. This is a snapshot of the workflow *as it was for this execution*. You may make changes and e.g. save or run the workflow, but the snapshot of the workflow associated with this particular execution in provenance will never change (i.e. you may always reopen it to find it as it was).

Selecting **Export Archive...** will allow you to save the selected run to a KAR file. A KAR file is a JAR (Java ARchive) file – a file that aggregates many files into one. If you are in the Workflow Run Manager suite, using Export Archive will create a KAR file that contains a copy of the workflow snapshot associated with the run, and a “WorkflowRun” file that contains metadata about the execution. If you are using the Reporting suite, the KAR will also contain, if they exist, the report design and generated report associated with the selected execution. Such a KAR file may be later opened using Kepler’s File => Open Archive (KAR) application menu item.

Selecting **Upload to Repository...** is the same as Export Archive action, only the KAR will also be uploaded to a remote repository you may have selected. You will be prompted to enter your credentials, or to upload anonymously, and whether or not you would like your KAR to be publically accessible. To configure to which remote repository Upload to Repository will upload, use the Tools => Preferences application menu, and navigate to the Components tab. Click the checkbox in the Save column for a repository of Type “remote”.

Selecting **Delete** allows you to delete the selected run(s). This deletes as much data associated with the run from provenance as possible, including the workflow snapshot, and any report design and report instance. **WARNING:** This action cannot be undone.

The **View Run LSID** and **View Workflow LSID** menu items allow a user to see the unique identifiers for a run or a run’s workflow. These menus are included for advanced users, and their usage is beyond the scope of this document.

Note that some menu options are not allowed when multiple runs are selected, or when a row is selected that is still executing (shows “running...” in the Duration column).

3.1 Tagging

Workflow runs may be tagged using the Tagging interface. To tag runs, select the desired run rows in the Workflow Run Manager, and then use the Tagging toolbar to select an existing or enter a new tag, and then hit enter or click the tagging button (the green “T” toolbar button). You can tell that you are tagging workflow runs (instead of the workflow) because the words “Tag run:” will appear before the Tagging toolbar. These tags will appear in the Workflow Run Manager Tags column, may then be searched against, and will be saved into the workflow run metadata file if you export or upload the run. You may delete these tags by selecting the run-rows, and clicking the X button on the tags in the toolbar. Please see the Tagging module suite documentation for more details about tagging.

3.2 Searching runs

Above each column in the Workflow Run Manager table is a small field into which search terms may be entered. Clicking in these cells and then entering search terms will filter which run rows are shown. An example is shown below:

The screenshot shows a window titled "Workflow Run Manager" containing a table. Above each column header is a search filter field. The table has columns for Tags, Workflow Name, Run Date, Duration, User, and Id. The data rows show various workflow runs with their respective details.

Tags	Workflow Name	Run Date	Duration	User	Id
mclaughlin	heatIndexAndDewpoint	May 24, 2010 5:07:45 PM	00:00:06	derik	9
mclaughlin	growingDegreeDays	May 24, 2010 5:07:20 PM	00:00:03	derik	8
mclaughlin, VegetationCover	mclaughlin_par	May 24, 2010 5:05:44 PM	00:00:54	derik	7
mclaughlin, VegetationCover	mclaughlin_par	May 24, 2010 5:05:12 PM	00:00:06	derik	6
mclaughlin	mclaughlin_soilmoisture	May 24, 2010 5:04:47 PM	00:00:03	derik	5

The **Tags** search field accepts a comma-separated list of case-insensitive search terms.

The **Workflow Name** and **User** search fields accept one case-insensitive search term.

The **Run Date** search field accepts a single date (e.g. 5/24/10), or a time range. A time range may be specified in the following ways. Specifying times is optional:

- > 5/24/10
- < 5/24/10 05:37:00 PM
- 5/24/10 – 5/26/10
- 5/24/10 05:37:00 PM – 5/26/10

Dates may be entered in a variety of formats, e.g. May 24, 2010 or 05/24/2010, but times must be specified with hour and minutes (and optionally seconds) followed by a space, followed by either “AM” or “PM”.

The **Duration** column may be searched using a range, similar to Run Date, utilizing “>”, “<”, or two durations separated by a “-“. For example, to search for those executions that took longer than 5 seconds to complete, you need only enter “>5”.

The **Id** column may be searched using a single Id, or a range, similar to Run Date and Duration.

3.3 Importing runs

To open a KAR exported using the **Export Archive...** option, simply use the Kepler File => Open Archive (KAR) application menu, or right-click on the KAR if it is in a folder in your Components Library (e.g. MyWorkflows) and select Open. Note that if you have deleted the run row from which this KAR was created, or this is a KAR from a different system (e.g. emailed to you from a friend), opening the KAR will also add a new run row in your Workflow Run Manager. By clicking on this run row you may view the associated report instance. If this run already exists in your Workflow Run Manager, opening the KAR will not import the run. Imported runs are highlighted a light blue.