

Getting Started with Kepler Workflow Run Manager 2.5

November 2015

Getting Started with the Workflow Run Manager

The Workflow Run Manager was created to assist Kepler users in managing 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 Runs	5
3.2 Searching Runs	5
3.3 Importing Runs	6
3.4 Browsing and Downloading Remote 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. 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 (“provenance”) 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). You may also use the Workflow Run Manager to connect to, browse, and download runs stored on the Kepler repository (<http://library.kepler-project.org/kepler/>).

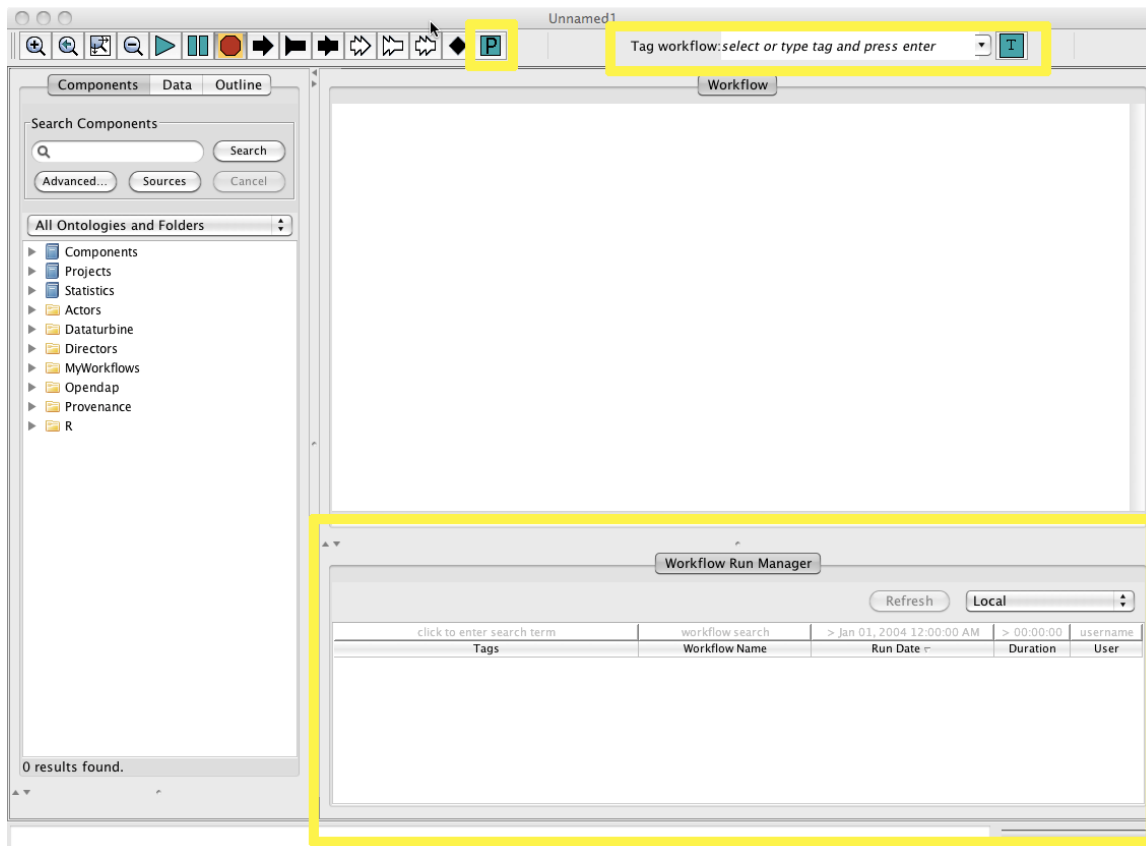
2. Downloading and Installing the Workflow Run Manager

From the Kepler application menu select Tools => Module Manager. From the Module Manager dialog, select the Available Modules tab. Select ‘workflow-run-manager-2.5’ from the list of Available Suites, then click the right arrow button to move the workflow-run-manager-2.5 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-2.5 suite (and please also refer to the Reporting documentation). Click the 'Apply and Restart' button to download the suite you have selected and restart Kepler.

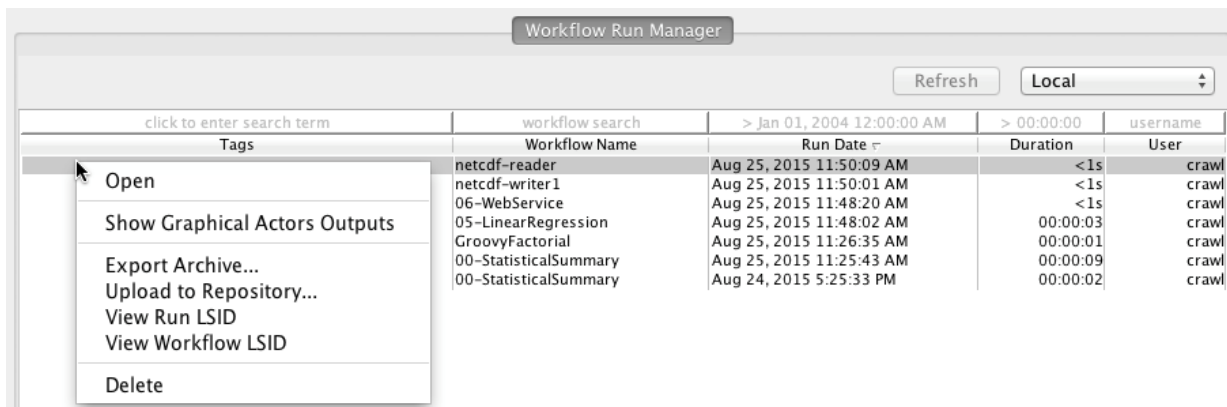
3. Using the Workflow Run Manager

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 (using a supported Director), a row will appear in the Workflow Run Manager table. Selecting 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 -- you may always right-click on this run to reopen the workflow as it was.

Selecting **Show Graphical Actors Outputs** will display the outputs from any graphical actors used in this workflow execution. The output displayed is identical to the output displayed when the workflow was executed. For example, if the workflow had the *Display* actor, then selecting this option will show the *Display* window containing the text that was displayed for the workflow execution. Similarly, if the workflow displayed a plot using the *Sequence Plotter* actor, then selecting this option will display the same plot.

The outputs from multiple workflow executions can be displayed simultaneously by first selecting the multiple rows in the table, and then selecting **Show Graphical Actors Outputs**. This allows you to, e.g., compare outputs from multiple executions of the same workflow.

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 later be 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 this action 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 data associated with the run from provenance, including the workflow snapshot, the run metadata, 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 Runs

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, and 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:

mcl	workflow search	> August 1, 2011 12:00:00 AM	> 00:00:01	username
Tags	Workflow Name	Run Date	Duration	User
McLaughlin, VegetationCover	mclaughlin_par	Aug 5, 2011 5:02:43 PM	00:00:07	derik
McLaughlin	mclaughlin_soilmoisture	Aug 5, 2011 5:00:51 PM	00:00:02	derik
McLaughlin, VegetationCover	mclaughlin_par	Aug 5, 2011 5:00:24 PM	00:00:03	derik
McLaughlin	heatIndexAndDewpoint	Aug 5, 2011 4:59:29 PM	00:00:16	derik
McLaughlin	growingDegreeDays	Aug 5, 2011 4:52:05 PM	00:01:19	derik

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 range. A 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”.

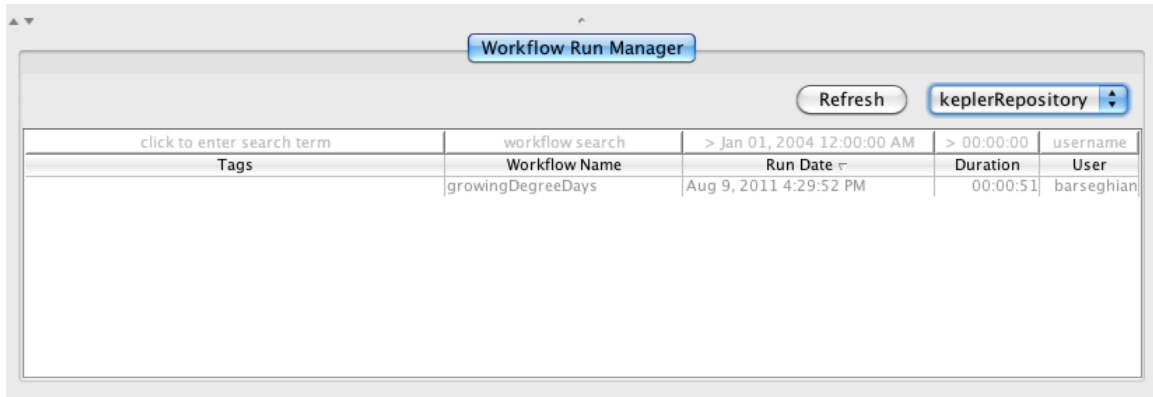
The **Duration** column may be searched using a range, similar to Run Date, utilizing “>”, “<”, or two durations separated by a “-“. For example, to search for those executions that took longer than 5 seconds to complete, you need only enter “>5”.

3.3 Importing Runs

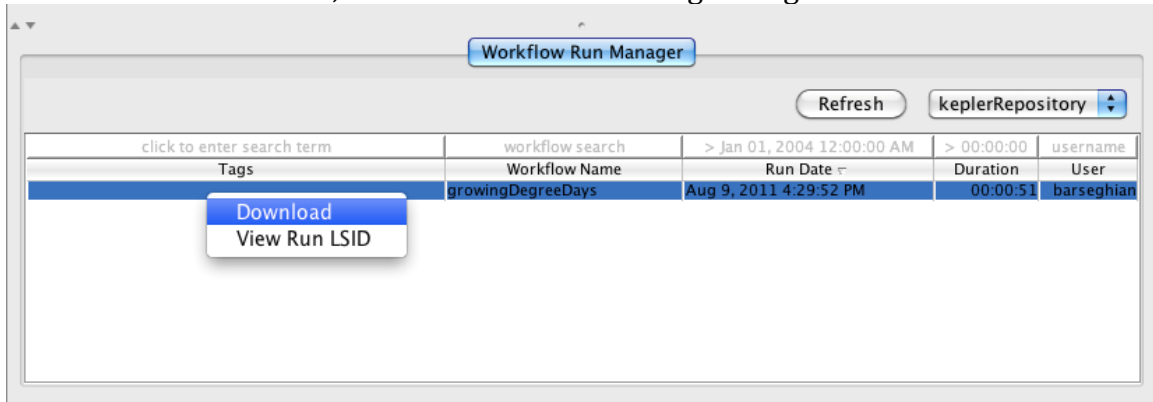
To open a KAR exported using the **Export Archive...** option, simply use the Kepler File => Open application menu, or right-click on the KAR if it is in a folder in your Components Library (e.g. MyWorkflows) and select Open. Note that if you have deleted the run row from which this KAR was created, 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.

3.4 Browsing and Downloading Remote Runs

The Workflow Run Manager may also be used to browse, download, and open remote KARs containing run-metadata (“run-KARs”). When you change from Local to keplerRepository using the drop-down list, the remote runs are shown in gray.



You can then select runs, and Download them using the right-click context menu.



Once downloaded, the run-row will be highlighted in light-blue. Right-clicking on the downloaded run-row will show the regular context menu options (Open, etc). Using Delete will delete your downloaded copy of this remote run-KAR, but will **not** delete the repository copy. You can verify this, as well as periodically check for new run-KARs on the remote repository, using the Refresh button.