

Getting Started with the Kepler Tagging module

Tagging 2.1.0

May 1, 2010

Getting Started with the Kepler Tagging Module

The *Getting Started with the Kepler Tagging Module* guide is a tutorial style manual for scientists who would like to use the *Tagging suite* for Kepler.

The Tagging Module was created to assist scientists who use Kepler in managing a large number of workflows and workflow runs.

Table of Contents

| | |
|--|---|
| 1. Introduction | 1 |
| 1.1 What is Tagging for Kepler and How is it used ? | 1 |
| 2. Downloading and Installing the Tagging module | 2 |
| 3. Basic Components of Semantic Tagging for Kepler | 2 |
| 3.1 Workflow Runs and the Tagging Interface | 3 |
| 3.2 Navigation between Workflows and the Tagging Interface | 3 |
| 3.3 Tagging Features | 3 |
| 4. Additional modules that can be used with the Workflow Run Manager ... | 4 |
| 4.1 Workflow Run Manager features | 4 |
| 5. Saving Tags in Workflows and Workflow Runs | 4 |
| 6. Publishing Tags in Workflows and Workflow Runs | 4 |
| 6.1 Saving Tags for Workflows, Workflow Runs, and Reports | 4 |
| 6.2 Publishing Tags with Workflows, Workflow Runs, and Reports | 4 |

1. Introduction

The “Getting Started with the Kepler Tagging Module” guide introduces the main components and functionality of the Tagging Module for Kepler. The Tagging module is used in conjunction with Kepler, a software application for the analysis and modeling of scientific data. Other modules that are often used with Tagging are the Workflow Run Manager module, Reporting, and Provenance. The Reporting Suite incorporates all of these modules into a single installation.

This guide provides a brief introduction to the tagging interface, as well as providing some guidance on how to use the features of the tagging system to effectively manage and organize large numbers of workflows and workflow runs.

1.1 What is Tagging for Kepler and How is it used?

The Tagging module adds additional features to the Kepler application. It provides a way to use semantic tagging for workflows and workflow runs. Tags have a label, which is just a text string, and optionally, a parent tag. For example, the tag ‘Blue’ might have as a parent tag ‘Color.’ Tags can even have more than one parent tag, but

this is uncommon. Tags can be used to facilitate many methods of personal organization; it's up to you. Here are a few ideas to get you started:

- Categorize workflows based on what input dataset they are configured to use
- Categorize workflow runs based on whether results were normal or anomalous
- Use tags to keep track of workflow runs referenced in a publication

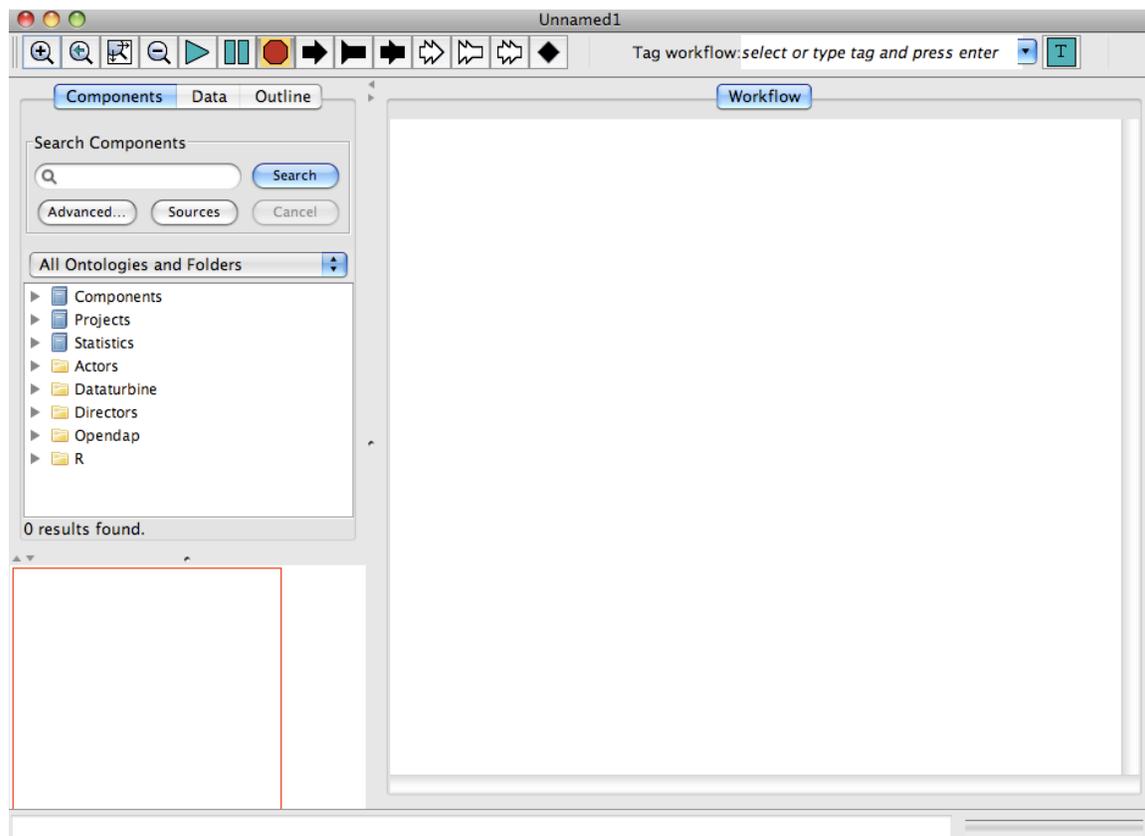
The sky's the limit. Just think of tags as a way to put extra information you want to keep track of.

2. Downloading and Installing the Tagging Module

From the main Kepler Menu select Tools > Module Manager. From the Module Manager dialog, select the Available Modules tab. Select 'tagging-2.1.0' from the list of Available Suites, then click the right arrow button to move the tagging-2.1.0 suite to the list of Selected Modules. Click the 'Apply and Restart' button to retrieve the suite you have selected, and re-start Kepler.

3. Basic Components of Semantic Tagging for Kepler.

Once you have installed the Tagging module, you will notice that the Kepler interface now has some new items. Notice the new text entry field in the toolbar. This is used to add new tags.



Open up a new canvas. To add a tag, simply click in the text field, type something, and press Return or click the tagging button (the green 'T' button to the right of the text field).

This will expand the tag bar, a separate area below the normal toolbar that shows the current tags on an object. You will see your brand new tag in this area. You just added a tag to the blank workflow. You can confirm what type of object you are tagging by looking at the label to the left of the tag text field. You can delete your tag by clicking the 'X' on the right side of the tag.

There are many pre-existing tags available, defined in ontologies bundled with Kepler. These can be selected by clicking the drop-down button next to the text field. The tags are color-coded by ontology, and the colors can be customized from the Tag Manager (Tools->Tag Manager...). Remember that to actually tag the object, you have to either hit Return after selecting a tag or click the Tag button.

3.1 Workflow Runs and the Tagging Interface.

Tagging workflow runs is similar to tagging workflows themselves. To tag an existing workflow run, just click on the entry in the Workflow Run Manager table. This lets the tagging system know which workflow run or workflow to tag. If the label of the tag input field indicated that you were tagging workflows previously, when you click it should indicate that you are now tagging a workflow run. Now just add a tag as usual. You will notice that it is displayed in the tag bar, as well as in the "Tags" column of the Workflow Run Manager table.

If you tried to tag a workflow run that came from a workflow that had tags on it, you might have noticed that your workflow run already had some tags to start with. Workflow runs inherit all tags of the workflow they came from. There is a slight difference between tags that a workflow run inherits from its workflow and those that are tags on the run itself; namely, tags inherited from a workflow cannot be deleted from the run. These tags do not display the 'X' that would normally allow their deletion.

Power User Tip: You can select multiple rows in the Workflow Run Manager table and tag a number of rows at once!

3.2 Navigation between Workflows with the Tagging Interface.

When the Tagging module is installed, Kepler will always start in the Workflow Editor view. Since each open workflow has its own set of all the UI elements involved in tagging, it is relatively simple to intuit what a given tagging request will do.

3.3 Tagging Features

When the Tagging module is installed, Kepler will always start in the Workflow Editor view. Since each open workflow has its own set of all the UI elements involved in tagging, it is relatively simple to intuit what a given tagging request will do. If you are in workflow tagging mode, then the workflow currently open in the window will be tagged. If you're in run tagging mode, then the currently selected run will be tagged. One thing to note is that if you have multiple windows open to the same workflow (looking inside composites), then tagging one will add a tag to all of those windows.

4. Additional modules that can be used with the Tagging module.

4.1 Workflow Run Manager features

The Workflow Run Manager is a way to quickly see information about and search through a large number of workflow runs. Tags on each of the workflow runs are displayed in a column of the Workflow Run Manager. A user can search on this column, just like the rest in the WRM.

5. Saving Tags for Workflows and Workflow Runs

Tags for workflows and workflow runs are both stored in a database for efficient querying by the Workflow Run Manager. Tags on workflows are also embedded into the MOML markup defining the workflow.

6 Publishing Tags for Workflows and Workflow Runs

If Alice gives Bob a workflow, either as a MOML file or as a KAR, which has tags on it, there are a few caveats. If the tags all come from ontologies that are packaged with Kepler, then there is no problem; they will appear on Bob's machine precisely the same as they would on Alice's. However, if Alice used custom tags on her workflow, then they will appear on Bob's computer as 'orphans.' Orphaned tags are tags that know what their label is, but not what ontology they came from, nor this tag's relationships with other tags in the ontology. Unless the ontology had complex relationships between concepts, this is rarely a problem, though.