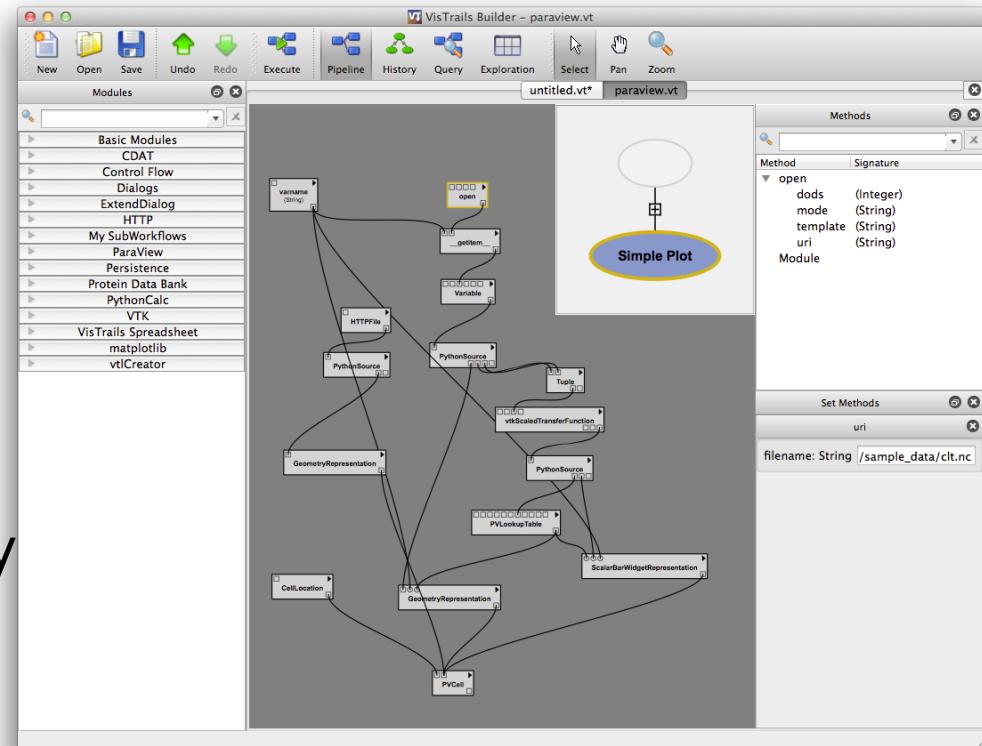


# VisTrails



# VisTrails Overview

- VisTrails is a scientific workflow management system focused on:
  - Exploration
  - Visualization
  - Analysis
- Has extensive provenance infrastructure
  - Allows reproducibility
  - Maintains complete history
- Can easily integrate libraries and packages



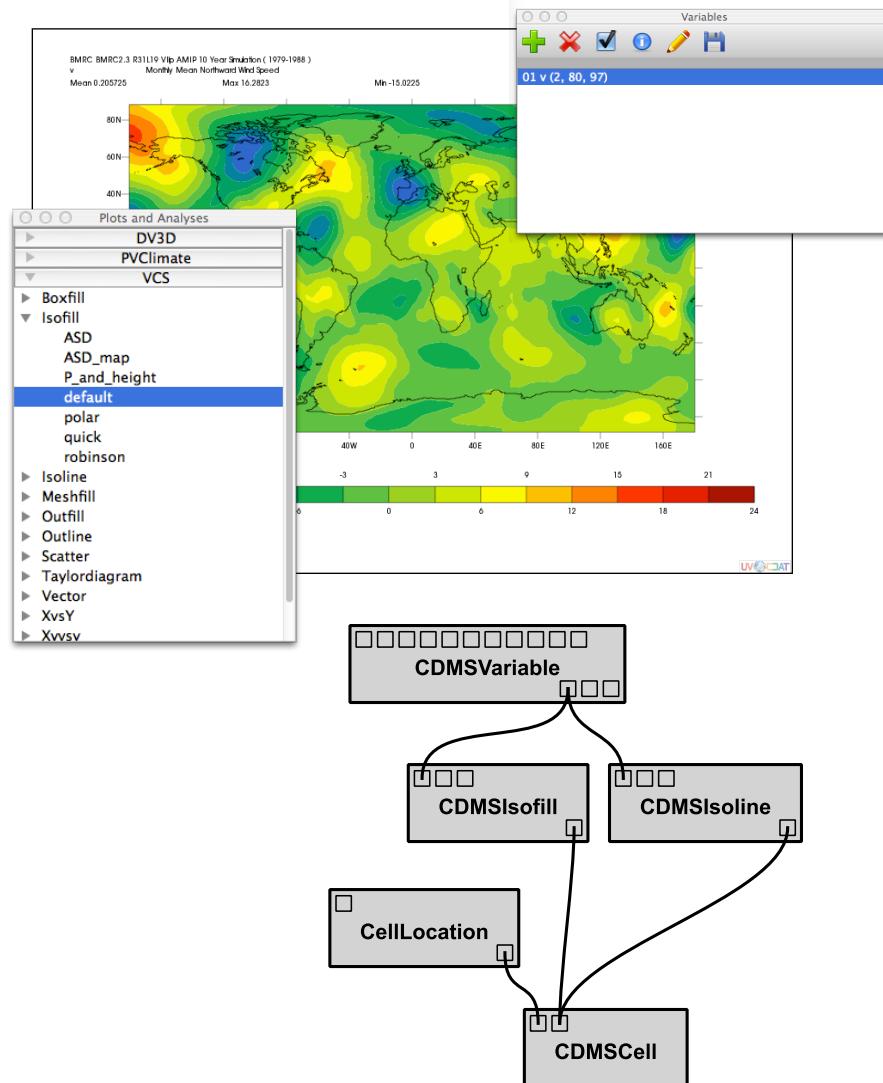
# VisTrails in UV-CDAT

---

- Each visualization in UV-CDAT is produced by executing an underlying VisTrails workflow
  - UV-CDAT builds workflows automatically so users don't have to
  - All of the provenance is automatically captured as well
- Any operations (e.g. regridding) and changes (e.g. colormap changes) involving data or visualizations are also automatically recorded
  - Both parameter changes and structural workflow modifications
  - Each action can be undone and replayed at will
- Users can access full VisTrails functionality from UV-CDAT
  - Allows advanced workflow customizations
  - Can view detailed provenance

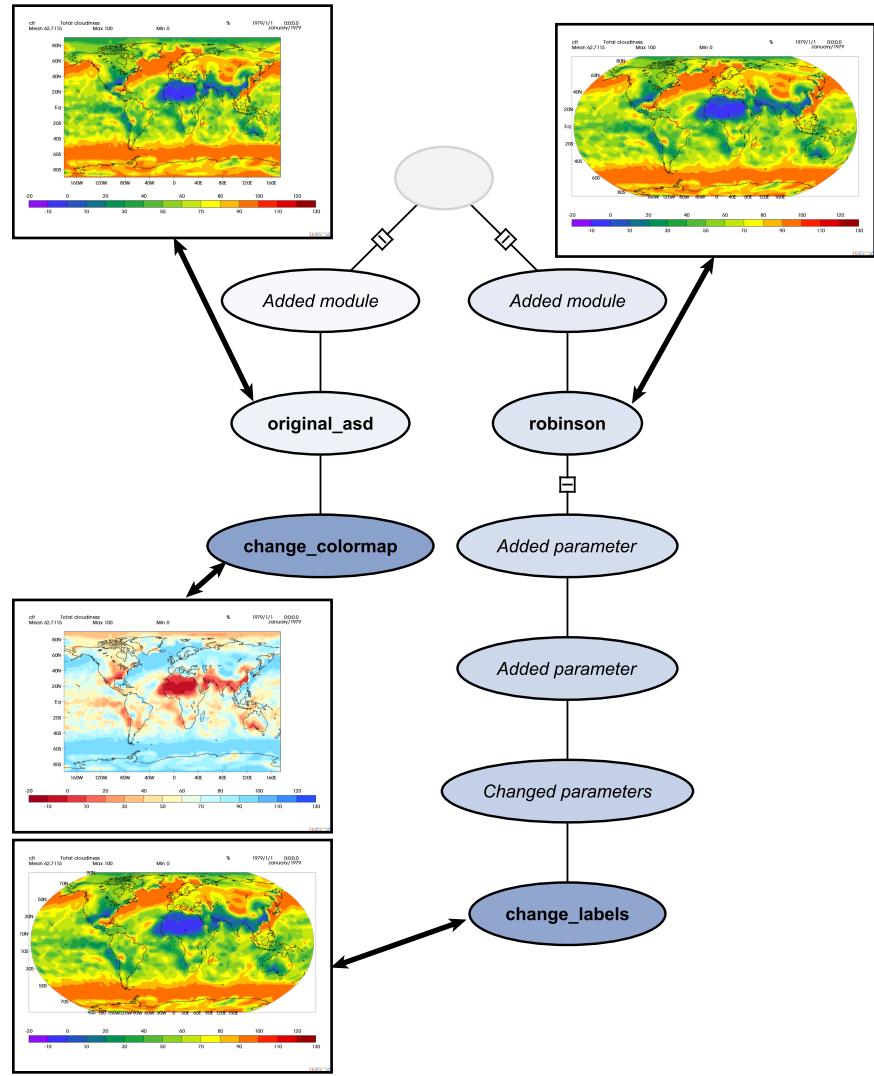
# Creating Workflows

- As soon as enough plots and variables have been dropped into a cell, the workflow is created and executed
- Adding plots and variables and changing parameters updates the underlying workflow
- Complex workflows can be created with a few drags and clicks



# VisTrails Provenance: Capturing Version History

- Each UV-CDAT cell is linked with a specific version of a workflow
- As users make changes, the VisTrails library automatically and transparently captures and records this history
- From the version tree (right), users can explore past analyses and step through each change that was made to a visualization



# VisTrails Provenance: Workflow Execution Logs

Log Details      Go to this pipeline

Successful      Error      Cached  
Not executed      Suspended

Pipeline	Start	End
► ROOT + 2*	2013-02-13 16:21:39....	2013-02-13 16:21:41....
► ROOT + 2*	2013-02-13 16:22:33....	2013-02-13 16:22:34....
► robinson + 4*	2013-02-13 16:24:54....	2013-02-13 16:24:54....
► robinson + 5*	2013-02-13 16:25:15....	2013-02-13 16:25:16....
► robinson + 6*	2013-02-13 16:25:30....	2013-02-13 16:25:31....
► robinson + 7*	2013-02-13 16:25:46....	2013-02-13 16:25:46....
► robinson + 8*	2013-02-13 16:25:57....	2013-02-13 16:25:58....
► robinson + 9*	2013-02-13 16:26:16....	2013-02-13 16:26:17....
► robinson + 10*	2013-02-13 16:27:04....	2013-02-13 16:27:05....
└ CDMSVariable	2013-02-13 16:27:04....	2013-02-13 16:27:04....
└ CDMSSlsofill	2013-02-13 16:27:04....	2013-02-13 16:27:04....
└ CellLocation	2013-02-13 16:27:04....	2013-02-13 16:27:04....
└ CDMSCell	2013-02-13 16:27:04....	2013-02-13 16:27:05....
► change_labels + 1*	2013-02-13 16:31:59....	2013-02-13 16:32:00....
► change_labels + 3*	2013-02-13 16:32:35....	2013-02-13 16:32:36....

robinson + 10\*  
Start: 2013-02-13 16:27:04.789245  
End: 2013-02-13 16:27:05.307975  
User: dakoop  
Completed: Yes

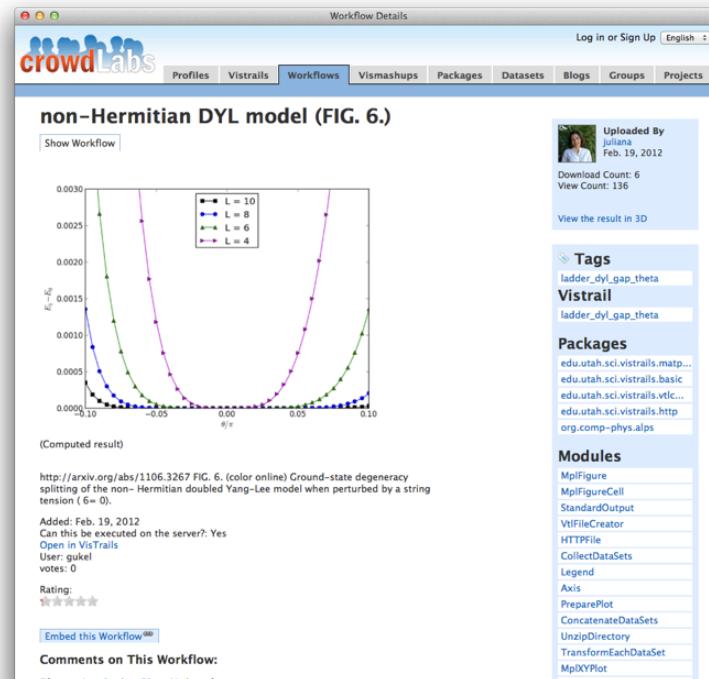
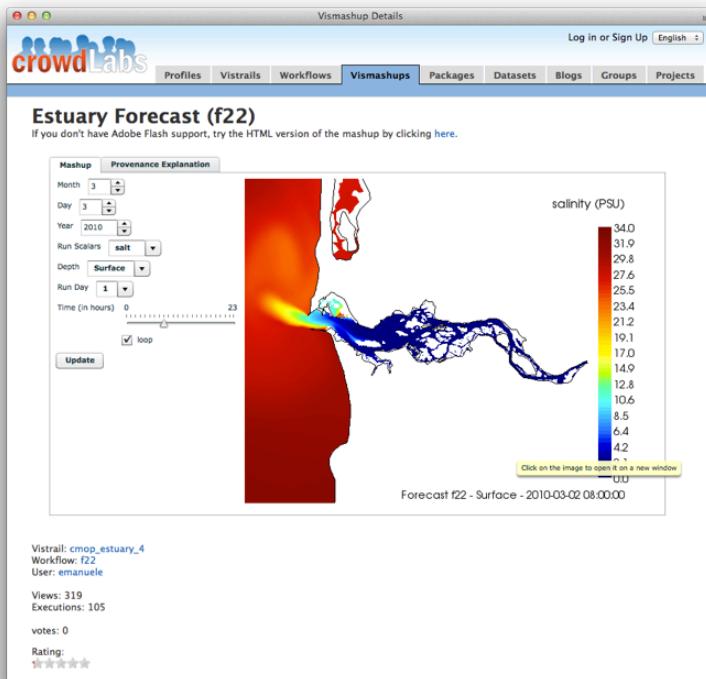
Annotations:  
\_reason\_: Pipeline Execution

```
graph TD; CDMSVariable[CDMSVariable] --> CellLocation[CellLocation]; CellLocation --> CDMSSlsofill[CDMSSlsofill]; CDMSSlsofill --> CDMSCell[CDMSCell];
```

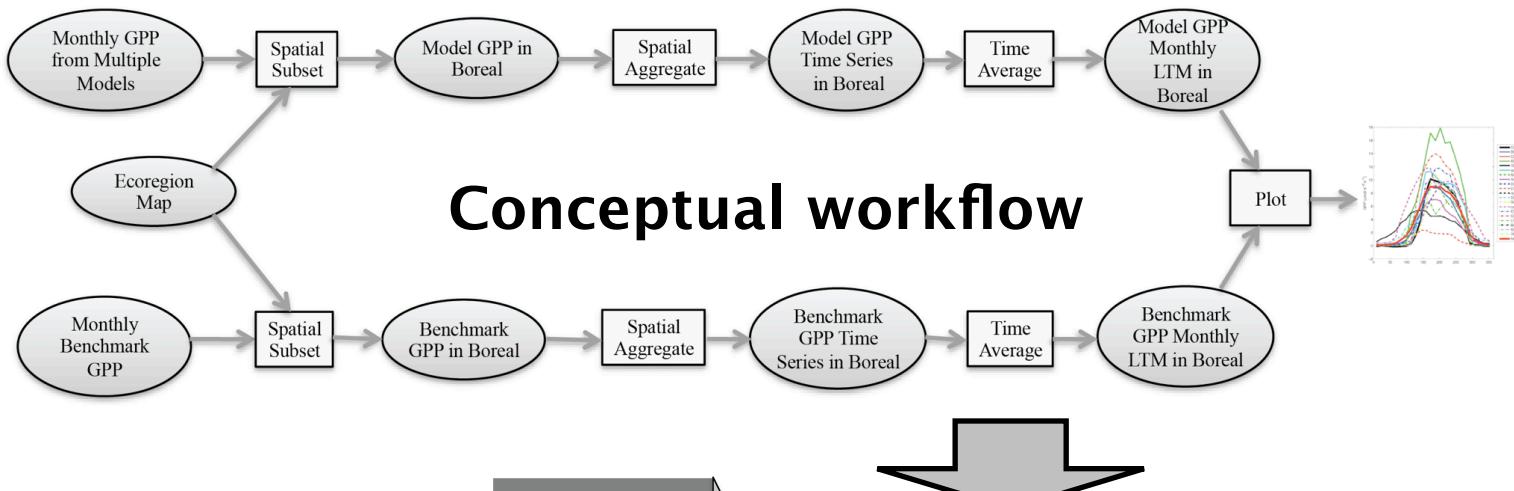
- VisTrails captures each step during the execution of an analysis
- Users can explore past executions and locate earlier results by searching this execution provenance
- VisTrails provides a graphical interface for browsing this provenance information (left)

# crowdLabs: A Social Visualization Repository

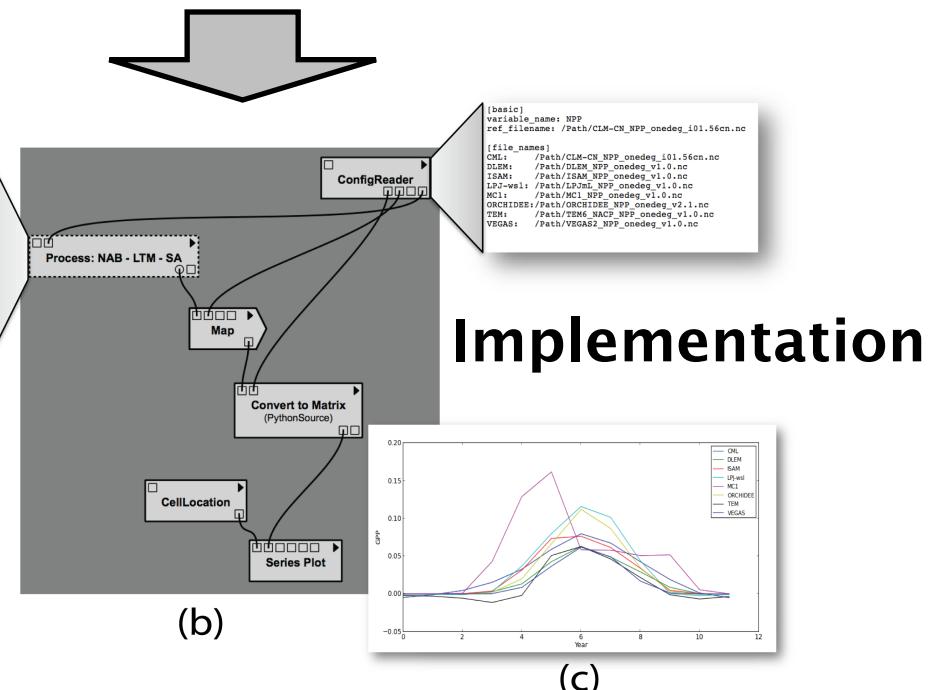
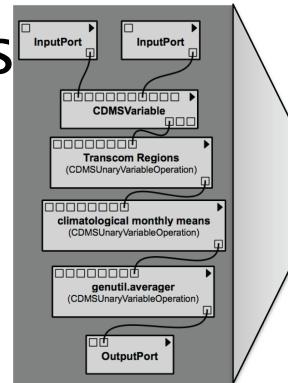
- Users can upload their work as well as download, investigate, and comment on others' work on [www.crowdlabs.org](http://www.crowdlabs.org)
- Can link from published papers to provenance and interactive visualizations (e.g. <http://arxiv.org/abs/1106.3267>)
- Planning to expand crowdLabs support for UV-CDAT workflows



# Multiple Model Inter-comparison using MsTMIP Data



Complex analyses leverage more advanced VisTrails features including subworkflows and looping

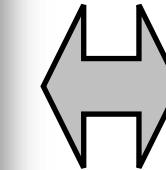
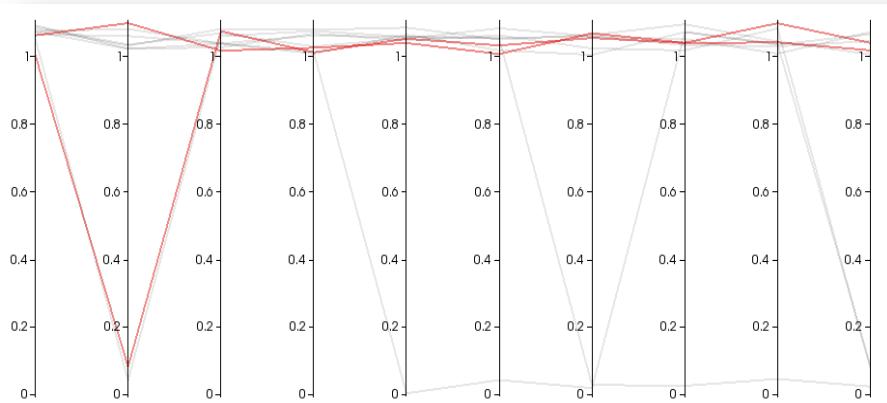


# Model Inter-comparison: Correlating Data Sources

- Model structure is represented by **Parallel Coordinates**.
- Model output is visualized using **Dimensionality Reduction**.
- Linked views are used to correlate both data sets.

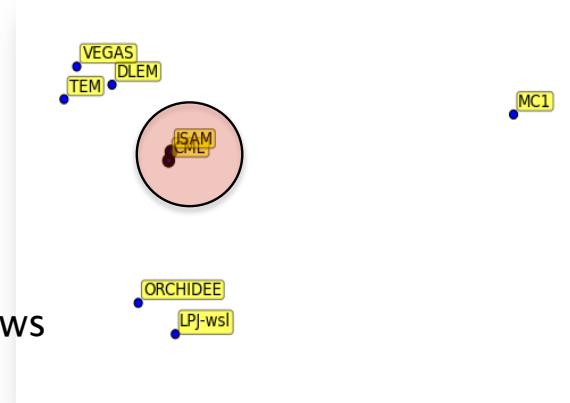
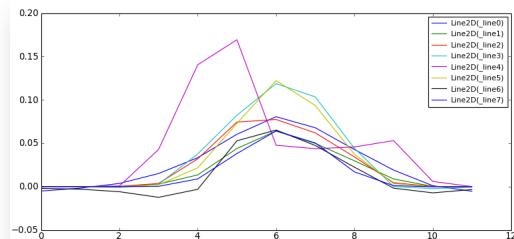
Model Structure

CLASS-CTEM-N+	CLM	DLEM	Ecosys	IRC/DayCent	ISAM	JULES	LPJ-wsl	MC1	ORCHIDEE
Reflectance/Transmittance/Absorptance computed by model (yes =1, no = 0)	1	1	1	0	1	1	0	0	1
RTS = 3-D (yes = 1; no = 0)	0	0	0	1	0	0	0	0	0
RTS = 2-stream (yes = 1; no = 0)	0	1	0	0	0	1	0	0	0
RTS = Beer's law (yes = 1; no = 0)	1	1	1	0	0	0	1	0	0
RTS = Albedo (yes = 1; no = 0)	1	1	1	0	0	0	0	0	1
model partitions net radiation into latent & sensible heat (yes = 1; no = 0)	1	1	0	1	0	1	0	0	0
Model simulates ground heat flux (yes = 1; no = 0)	1	1	0	1	0	1	0	0	1
Canopy stomatal conductance - shaded leaves (yes = 1; no = 0)	1	1	1	1	0	1	1	0	0
Canopy stomatal conductance - sun leaves (yes = 1; no = 0)	1	1	1	1	0	1	0	0	0
Canopy stomatal conductance - whole canopy (yes = 1; no = 0)	1	0	0	0	0	0	0	1	0
Stomatal conductance scheme = Jarvis-type (yes = 1; no = 0)	0	0	0	0	0	0	0	0	0
Stomatal conductance scheme = Ball Berry (yes = 1; no = 0)	1	1	1	0	0	1	1	0	1
Stomatal conductance connected to photosyn (yes = 1; no = 0)	1	1	1	1	0	1	1	0	1



Linked Views

Model Output



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

UVICDAT

# Adding Packages to UV-CDAT

- Use the VisTrails API
  - Programmatically create workflows with a few lines of code
  - Create custom UI widgets to modify and control plots
  - Flexible spreadsheet package supports advanced interaction with 3D visualizations

```
registry = get_module_registry()
descriptor = registry.get_descriptor_by_name
pkg = 'gov.llnl.uvcdat.cdms'

variableDesc = descriptor(pkg, 'CMDSVariable')
add_module_from_description(variableDesc)
...
```

