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

- 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

• 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

Conceptual workflow

Complex analyses leverage more advanced VisTrails features including subworkflows and looping

Implementation
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**

<table>
<thead>
<tr>
<th>Class</th>
<th>CLM</th>
<th>DLEM</th>
<th>Ecosys</th>
<th>MG/GeoClim</th>
<th>IPSAM</th>
<th>JULES</th>
<th>LPJwsl</th>
<th>MC1</th>
<th>ORCHIDEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLASS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CLM</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DLEM</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ecosys</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MG/GeoClim</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>IPSAM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>JULES</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LPJwsl</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MC1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ORCHIDEE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Model Output**

**Linked Views**
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

```python
registry = get_module_registry()
descriptor = registry.get_descriptor_by_name(pkg = 'gov.llnl.uvcdat.cdms'
variableDesc = descriptor(pkg, 'CMDSVariable')
add_module_from_description(variableDesc)
...```