genSpace: Community-Driven Knowledge Sharing for the Discovery and Visualization of Workflows in geWorkbench
Swapneel Sheth, Gail Kaiser – {swapneel, kaiser} @ cs.columbia.edu
Programming Systems Lab, Department of Computer Science, Columbia University

**Motivation and Goals**
- Scientists collaborating together in the same lab on the same project share Data, Tools, Knowledge
- This model does not scale to communities
- Most current CSCW systems enable data sharing and/or tool sharing, but knowledge sharing (how/when/where/why to use tools and data) is limited to labor intensive approaches – pubs, email, wikis, chat, etc.
- We seek to enable automatic knowledge sharing without requiring “extra work” by scientists
- Our approach leverages popular social networking concepts, e.g., collaborative filtering (“people like you”) to disseminate knowledge on how to use geWorkbench

**genSpace Prototype**
- Implemented as a plugin for geWorkbench
- Logs, aggregates, and data mines users’ activities to recommend tools and tool sequences (workflows)
- Questions genSpace can answer
  - What do I do first?
  - Which tools work well together?
  - Where does this tool fit in a typical workflow?
  - Who do I know who also uses this tool?
  - How do I get help (from an expert who is online right now)?)
- Planned for next geWorkbench release

**Short Term Future Work**
- More precise monitoring – specific analysis parameters and options, visualization activities
- SMS with Annotated Workflow Attachments

**Long Term Future Work**
- Privacy and Security – Scientific research may involve competition & confidentiality, leverage collaborative networks to restrict visibility to only select colleagues
- Address “concept drift” as user participation, tool and workflow usage change over time
- Scaling to hundreds of users, millions of logs, offline access, fault tolerance – caching, incremental analysis

<table>
<thead>
<tr>
<th>Single Workflow Visualization</th>
<th>Workflow Relationship Visualization</th>
<th>Rating of Tools</th>
<th>Workflow and Tool Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Single Workflow Visualization" /></td>
<td><img src="image2" alt="Workflow Relationship Visualization" /></td>
<td><img src="image3" alt="Rating of Tools" /></td>
<td><img src="image4" alt="Workflow and Tool Statistics" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Realtime Workflow Visualization</th>
<th>genSpace Web Portal</th>
<th>Collaborative Network Visualization</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5" alt="Realtime Workflow Visualization" /></td>
<td><img src="image6" alt="genSpace Web Portal" /></td>
<td><img src="image7" alt="Collaborative Network Visualization" /></td>
</tr>
</tbody>
</table>