Chain ReAKTing:
Collaborative Advanced Knowledge Technologies in the Comb-e-Chem Grid

Michelle Bachler, Simon Buckingham Shum, Jessica Chen-Burger, Jeff Dalton, David De Roure, Marc Eisenstadt, Jeremy Frey, Jiri Komzak, Danius Michaelides, Kevin Page, Stephen Potter, Nigel Shadbolt and Austin Tate

Abstract

The CoAKTinG (Collaborative Advanced Knowledge Technologies in the Grid) project has developed a set of integrated tools to enhance collaboration between e-Scientists. As one of three case studies, these tools are being applied within the Comb-e-Chem e-Science pilot project. Two levels of integration are being explored: straightforward deployment of generic CoAKTinG tools, and a "deep" integration between these tools and the Comb-e-Chem grid. The deeper integration supports the 'publication at source' research objective of Comb-e-Chem, in which a digital record is maintained through the information processing chain that starts in the laboratory, supporting retrospective use in the e-Science process. In this paper we provide an overview of the tools and we focus in particular on the adaptation of one of the tools for the Comb-e-Chem application.

CoAKTinG

The objective of CoAKTinG ('Collaborative Advanced Knowledge Technologies in the Grid') has been to advance the state of the art in collaborative mediated spaces for distributed e-Science through the novel application of advanced knowledge technologies. It comprises four tools: meeting capture and replay, instant messaging and presence notification (BuddySpace), graphical meeting and group memory capture (Compendium) and intelligent 'to-do' lists (Process Panels). These are integrated through exchanging and storing events according to a set of CoAKTinG ontologies.

In its last phase, CoAKTinG has been conducting a number of case studies. One of these has been with Comb-e-Chem, to put the tools in the hands of e-Scientists. The discussions between the two teams have led to two notions of integration: 'shallow', where the tools are deployed as they are, and 'deep' where the tools are more intimately integrated with the Comb-e-Chem systems.

Comb-e-Chem

The CombeChem project aims to enhance structure property correlation and prediction by increasing the amount of knowledge about materials via synthesis and analysis of large compound libraries. Automation of the measurement and analysis is required in order to do this efficiently and reliably while ensuring that wide dissemination of the information occurs together with all the necessary associated background (raw) data that is needed to specify the provenance of the material. The project aims for a complete end-to-end connection between the laboratory bench and the intellectual chemical knowledge that is published as a result of the investigation; this necessitates that all steps in the process are enhanced by a suitable digital environment. Comb-e-Chem has achieved many parts of this ambitious programme, e.g. the smart laboratory (smarttea.org), grid-enabled instrumentation, data tracking for analysis, methodology for publication@source, process and role based security and high throughput computation.

The CoAKTinG tools provide support for the e-Science process in Comb-e-Chem and they also enable the digitisation of 'missing links' in the processing chain which form part of the typical collaborative scientific processes that we are attempting to enhance using the grid infrastructure: support of the experimental process, tracking and awareness of people and machine states, capturing of the discussions about data as well as the traditional metadata, and enriched meta-data regarding these components to support interlinking.

Deep Integration

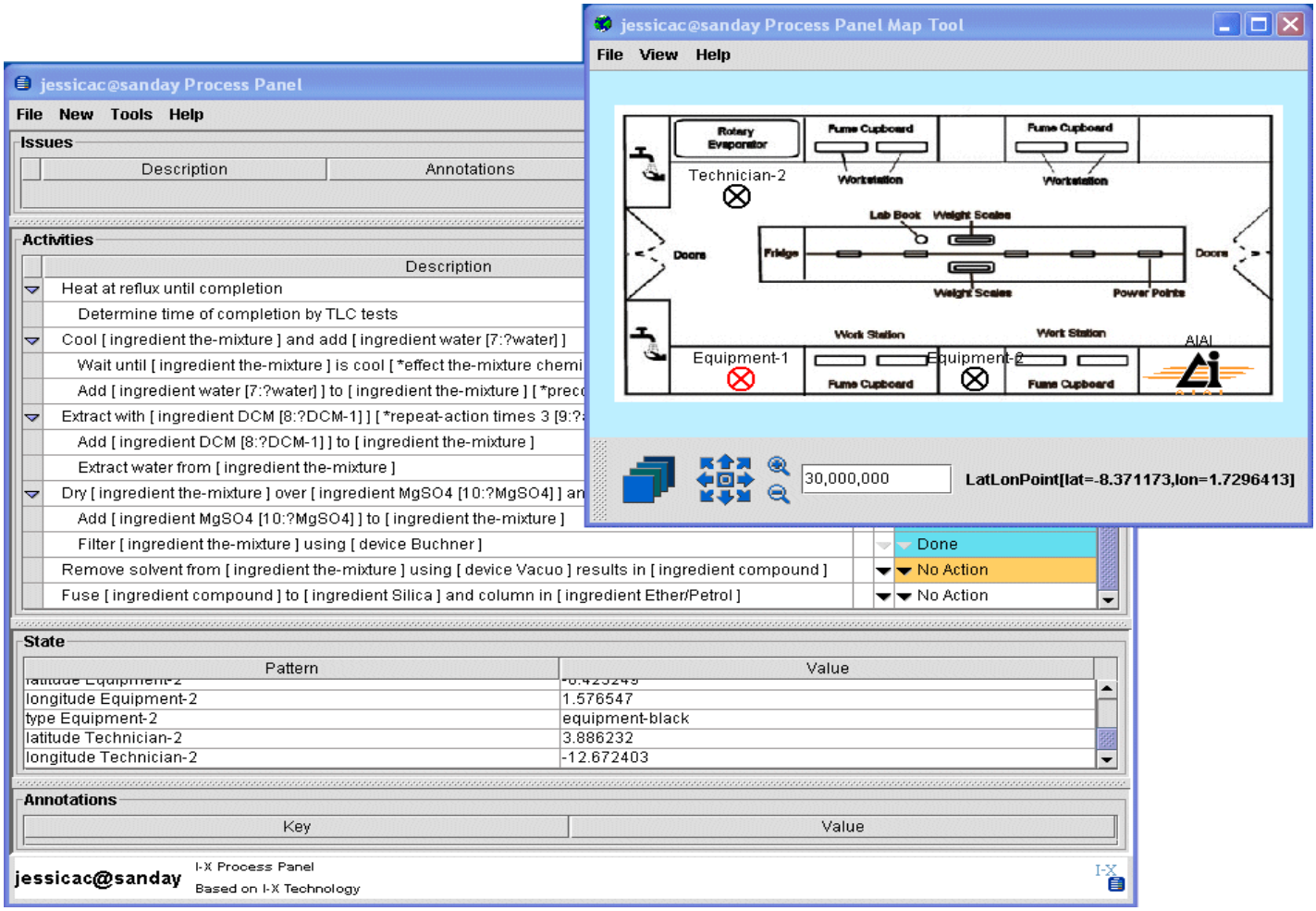
The BuddySpace systems can be adapted to show and track the interactions between the staff and equipment using the National Crystallographic Service (NCS), providing information to their users about the state of the service. Compendium provides the harness to ensure more adequate capture of the discussions in analysis, while Process Panels provide the means to initiate and track key tasks and issues. Additionally the ideas from CoAKTinG provide different techniques to achieve the necessary multi-user interaction in real time over the

network and give Comb-e-Chem the opportunity to implement the 'video interaction' collaboration part of Comb-e-Chem using event based ontologies to annotate real time streaming media and content.

These various components are valuable complements to Comb-e-Chem individually but jointly are even more powerful. For example, Process Panels can exploit the presence information derived from Buddy Space with respect to instrument status and operator availability to offer more informed task delegation options. This completes the chain of digital support and capture, maximising the potential for re-use of the digital information in support of the scientific process.

Process panels

Here we focus on one particular aspect of the deep integration – the application of the Process Panel tool to the laboratory, building on the process capture work of Comb-e-Chem's Smart Tea team. A picture speaks a thousand words:



This figure shows a screen capture of an IX Process Panel and its Map Tool resulting from our initial experiment. The Map Tool depicts a real Chemistry lab where both fixed and mobile entities are represented. The positions of mobile entities such as movable equipment and technicians are updated automatically through the (World) State sub-panel. By sharing information with BuddySpace, (dynamic) properties of devices are also described in the same panel. At this particular point in time, it shows Technician-2 is in front of the Rotary Evaporator and about to carry out the sub-process "Remove solvent from the-mixture using Vacuo results in Compound", having completed the previous steps in this process. In our investigation, the process decomposition facility of the IX Activity sub-panel supports views of different levels of abstraction that fits nicely with different chemists' (and labs') practice. Activities, issues, annotations and constraints may be recorded directly or via Compendium where in-depth discussion has taken place. Static and dynamic process editing provide great flexibility as processes are modifiable at run-time in response to unexpected changes. The ability to store, retrieve and refine process models is important in the Chemistry domain where existing processes are constantly reviewed and modified to discover or synthesise new chemical compounds. This facility alone makes I-X a valuable back-end component for integration with the existing Comb-e-Chem Grid.