Panel on "Workflow as the Methodology of Science"
Tuesday June 20 2006 WORKS Workshop
HPDC Paris France 12pm - 1.30pm
http://www.isi.edu/works06
Moderator Geoffrey Fox

A recent NSF workshop http://vtcpc.isi.edu/wiki/index.php/Main_Page proposed that workflow could be viewed as underlying support for the scientific methodology emerging in many fields and involving distributed interdisciplinary data deluged scientific methodology as an end (instrument, conjecture) to end (publication, archived results) process. This vision for workflow mixes the coupled execution of related services characteristic of most scientific workflow with the more asynchronous longer term processes familiar in some business workflow. Can one usefully link these different styles of workflow and further enhance scientific productivity?

One challenge is reproducibility of this full process which is core to the scientific method and requires rich provenance, interoperable persistent repositories with linkage of open data and publication as well as distributed simulations, data analysis and new algorithms. The distributed reproducible science methodology can be supported by publishing all steps in a sort of electronic logbook that is the "script" of the full scientific workflow. It would need to capture the scientific process (data analysis) as a rich cloud of resources including emails, presentations, wikis as well as databases, compiler options, build time/runtime configurations etc. One could need to separate wheat from chaff in this electronic record (logbook) keeping only that required to make process reproducible and allowing selective execution (checking) of components of the log.

Is this a reasonable picture for a future workflow requirement and what are the new research challenges it engenders?

The presentations at NSF meeting can be found at http://vtcpc.isi.edu/wiki/index.php/Documents and give us a starting point!

Contributors:

E. Deelman, USC/ISI	Summary of NSF Workshop
S. Jha, University College of London	Application perspective
D. De Roure, University of Southampton	Provenance
I. Foster, ANL & Univ. of Chicago	Lessons from current Science Grids
John Ibbotson, IBM Hursley	Web service and business workflow