

What Makes Scientific Workflows Scientific?

Bertram Ludäscher

University of California, Davis
ludaesch@ucdavis.edu

A *scientific workflow* is the description of a process for accomplishing a scientific objective, usually expressed in terms of tasks and their dependencies [5]. While workflows have a long history in the database community as well as in business process modeling (where they are also known as *business workflows*), and despite some early works on scientific workflows [3,10], the area has only recently begun to fully flourish (e.g., see [1,2,9,7,4,11]). Similar to scientific data management which has different characteristics from traditional business data management [8], scientific workflows exhibit new challenges and opportunities that distinguish them from business workflows. We present an overview of these challenges and opportunities, covering a number of issues such as different models of computation, scalable data and process management, and data provenance and lineage handling in scientific workflows.

References

1. Fox, G.C., Gannon, D. (eds.): *Concurrency and Computation: Practice and Experience*. Special Issue: Workflow in Grid Systems, vol. 18(10). Wiley & Sons, Chichester (2006)
2. Gil, Y., Deelman, E., Ellisman, M., Fahringer, T., Fox, G., Gannon, D., Goble, C., Livny, M., Moreau, L., Myers, J.: *Examining the Challenges of Scientific Workflows*. *Computer* 40(12), 24–32 (2007)
3. Ioannidis, Y.E., Livny, M., Gupta, S., Ponnokanti, N.: *ZOO: A Desktop Experiment Management Environment*. In: *VLDB*, pp. 274–285 (1996)
4. Ludäscher, B., Altintas, I., Berkley, C., Higgins, D., Jaeger, E., Jones, M., Lee, E.A., Tao, J., Zhao, Y.: *Scientific Workflow Management and the Kepler System*. *Concurrency and Computation: Practice & Experience* 18(10), 1039–1065 (2006)
5. Ludäscher, B., Bowers, S., McPhillips, T.: *Scientific Workflows*. In: Özsu, M.T., Liu, L. (eds.) *Encyclopedia of Database Systems*. Springer, Heidelberg (to appear, 2009)
6. Ludäscher, B., Goble, C. (eds.): *ACM SIGMOD Record: Special Issue on Scientific Workflows*, vol. 34(3) (September 2005)
7. Oinn, T., Addis, M., Ferris, J., Marvin, D., Senger, M., Greenwood, M., Carver, T., Glover, K., Pocock, M., Wipat, A., Li, P.: *Taverna: a tool for the composition and enactment of bioinformatics workflows*. *Bioinformatics* 20(17) (2004)
8. Shoshani, A., Olken, F., Wong, H.K.T.: *Characteristics of Scientific Databases*. In: *VLDB*, pp. 147–160. Morgan Kaufmann, San Francisco (1984)
9. Taylor, I., Deelman, E., Gannon, D., Shields, M. (eds.): *Workflows for e-Science: Scientific Workflows for Grids*. Springer, Heidelberg (2007)
10. Wainer, J., Weske, M., Vossen, G., Medeiros, C.B.: *Scientific Workflow Systems*. In: *Proceedings of the NSF Workshop on Workflow and Process Automation in Information Systems: State of the Art and Future Directions* (1996)
11. Yu, J., Buyya, R.: *A Taxonomy of Scientific Workflow Systems for Grid Computing*. In: Ludäscher and Goble [6]