Preface

Shiyong Lu*

Department of Computer Science,
Wayne State University,
Detroit, MI 48202, USA
Fax: 313-577-6868
E-mail: shiyong@wayne.edu
*Corresponding author

Ewa Deelman

USC Information Sciences Institute,
4676 Admiralty Way, Suite 1001,
Marina del Rey, CA 90292, USA
E-mail: deelman@isi.edu

Zhiming Zhao

University of Amsterdam,
Science Park 107, 1098XG,
Amsterdam, The Netherlands
E-mail: z.zhao@uva.nl

Biographical notes: Shiyong Lu is an Associate Professor in the Department of Computer Science at Wayne State University and the Director of the Scientific Workflow Research Laboratory (SWR Lab). He received his PhD in Computer Science from the State University of New York at Stony Brook in 2002. His current research interests focus on scientific workflows and their applications. He is the Founding Chair of the IEEE International Workshop on Scientific Workflows (SWF) and a Founding Editorial Board Member of the *International Journal on Semantic Web and Information Systems*. He is a Senior Member of the IEEE.

Ewa Deelman is a Research Associate Professor at the USC Computer Science Department and a Project Leader at the USC Information Sciences Institute. Her research interests include the design and exploration of collaborative, distributed scientific environments, with particular emphasis on workflow management as well as the management of large amounts of data and metadata. At ISI, she is leading the Pegasus project, which designs and implements workflow mapping techniques for large-scale workflows running in distributed environments. She received her PhD from Rensselaer Polytechnic Institute in Computer Science in 1997 in the area of parallel discrete event simulation.

Zhiming Zhao obtained his PhD at the University of Amsterdam (UvA) in 2004 and currently works as a Researcher at the Computer Science Department at the University of Amsterdam. He studied Computer Science in Nanjing Normal University and East China Normal University in China, and obtained a Bachelor and MSc in 1993 and 1996 respectively. His research interest includes scientific workflow management, distributed interactive simulation systems, e-science and grid computing.

Scientific workflows have recently emerged as a new paradigm for scientists to formalise and structure complex scientific processes to enable and accelerate many significant scientific discoveries. A scientific workflow is a formal specification of a scientific process, which represents, streamlines and automates the analytical and computational steps that a scientist needs to go through from dataset selection and integration, computation and analysis, to final data product presentation and visualisation (Lin et al., 2009). The critical role of scientific workflows in science has been recognised by an NSF workshop on the challenges of scientific workflows (Deelman et al., 2006), which concluded that “workflows should become first-class entities in cyberinfrastructure architecture. For domain scientists, they are important because workflows document and manage the increasingly complex processes involved in exploration and discovery through computations. For computer scientists, workflows provide a formal and declarative representation of complex distributed computations that must be managed efficiently through their lifecycle from assembly, to execution, to sharing".

In this special issue on scientific workflows, we have accepted seven regular articles covering various aspects of scientific workflows, from language, system, enactment, typing, patterns, to interoperability. In this special issue, we briefly summarise these articles. The article entitled 'A declarative language and toolkit for scientific workflow implementation and execution' by Jamil et al., introduces a scientific workflow language and system featuring declarativity, a highly successful concept in databases. In the article entitled 'Context-aware scientific workflow systems using Kepler', Ngu et al. propose a context-aware architecture for scientific workflows, in which contexts are incorporated into the dataflow-oriented Kepler scientific workflow system, enabling the development of context-aware scientific workflows while avoiding the use of numerous low-level controlflow actors. In the article entitled 'Process space-based scientific workflow enactment', Sonntag et al. explore the use of process spaces, a middleware for the decentralised execution of scientific workflows. In the article entitled 'Polymorphic type framework for scientific workflows with relational data model', Curcin et al. propose a polymorphic-type framework for the inputs and outputs of workflow tasks suitable for use in registries and for type-matching; they also develop a polymorphic-type inference over compositions of such signatures. In the article entitled 'Scientific workflow management systems and workflow patterns', Shiroor et al. argue that control flow constructs are necessary in scientific workflow systems and present an approach to incorporate a set of common control flow patterns in the Kepler system. In recognition of the importance of scientific workflows for silico experiments, Mattoso et al. present a first reference model for the scientific experiment life cycle in the article entitled 'Towards supporting the life cycle of large scale scientific experiments'. Finally, in the article entitled 'Scientific workflow interoperability framework', Alqaoud et al. present a framework to address scientific workflow interoperability, an important issue in the area of scientific workflows (Lin et al., 2009).

This special issue augments the snapshot of representative work in scientific workflows collected in Zhao et al. (2009), and we hope this special issue can stimulate and inspire further research in this field. As a recent paper (Bell et al., 2009) concludes, "In the future, the rapidity with which any given discipline advances is likely to depend on how well the community acquires the necessary expertise in database, *workflow management*, visualisation, and cloud computing technologies…". We are looking forward to an era in which scientific workflows are becoming a fundamental technology that interplays with other key technologies to revolutionise the way that people do science in the 21st century (Zhao et al., 2008).

References

Bell, G., Hey, T. and Szalay, A. (2009) 'Beyond the data deluge', *Science*, Vol. 323, No. 5919, pp.1297–1298, March.

Deelman, E., Gil, Y. and Zemankova, M. (2006) *NSF Workshop on the Challenges of Scientific Workflows*, Arlington, VA, May.

Lin, C., Lu, S., Fei, X., Chebotko, A., Pai, D., Lai, Z., Fotouhi, F. and Hua, J. (2009) 'A reference architecture for scientific workflow management systems and the VIEW SOA solution', *IEEE Transactions on Services Computing (TSC)*, Vol. 2, No. 1, pp.79–92.

Zhao, Y., Raicu, I. and Foster, I. (2008) 'Scientific workflow systems for 21st century e-science, new bottle or new wine?', *Proc. of the Second IEEE International Workshop on Scientific Workflows (SWF 2008)*, pp.467–471.

Zhao, Z., Belloum, A. and Bubak, M. (2009) 'Special section on workflow systems and applications in e-science', *Future Generation Computer System*, Vol. 25, No. 5, pp.525–527.