Computation has been described as the “third leg” of science, along with theory and experimentation. Certainly, modern information systems are vital to managing and processing the huge amounts of data produced by simulations and experiments. However, existing tools are only now beginning to catch up with the needs of today’s scientists, and much more needs to be done to support the computational needs of tomorrow’s scientists. In particular, scientists still need effective tools to deal with massive data sets that may be geographically scattered, to apply multiple complex and interacting transformations to the data, and to ensure the quality and repeatability of their computations. The IEEE Workshop on Workflow and Data Flow for Scientific Applications brings together researchers who are exploring how to build the next generation of information systems to address these needs.

The papers presented in this workshop demonstrate the ability of computer scientists and natural scientists to work together to create computer systems that support scientific exploration. The systems and techniques described in these papers range from general toolkits that make it easier for a wide range of scientists to produce and manage workflows and datafows, to specific case studies of scientific applications which leverage workflow support in innovative ways. Our hope is that by bringing researchers together to discuss these papers, we can help generate further discussions and collaborations aimed at applying workflow and data flow technology to meet scientists’ needs.

We would like to thank our Program Committee, Ian Foster, Michael Franklin, James Frew, Reagan Moore, Kian-Lee Tan and Marvin Theimer, as well as our external reviewers, Mike Wilde, Yong Zhao, and Rocky Dunlap, for reviewing the submitted papers and helping us to select a high-quality program. We would also like to thank the authors who submitted papers for helping to make the workshop a success.

Brian F. Cooper and Roger Barga
Program Co-Chairs