

Realizing Opportunities for Advanced and Automated Workflows in Scientific Research

Committee

Daniel E. Atkins, III

Chair

DANIEL E. ATKINS (Chair) (NAE) is Emeritus W.K. Kellogg Professor of Information and Professor of Electrical Engineering and Computer Science at the University of Michigan (UM), Ann Arbor. The first phase of his career focused on computer architecture including high-speed arithmetic methods now widely used in modern computers, as well as the design and construction of application-specific experimental machines. The second phase of his career focused on pioneering interdisciplinary research on cyber-enabled distributed knowledge communities including collaboratories and digital libraries applied to both scientific research and education. He has served as dean of the school of engineering, founding dean of the School of Information, and associate vice president for research at UM, as well as the inaugural director of the Office of Cyberinfrastructure at the National Science Foundation (NSF). He chaired the Blue Ribbon Panel on Research Cyberinfrastructure for the NSF that became an international roadmap for initiatives on cyber-enabled research in the digital age. He has chaired or served on many advisory board for government, academia, philanthropy, and industry. Professor Atkins is a member of the National Academy of Engineering. He earned a Ph.D. in computer science and an M.S. in electrical engineering from the University of Illinois, Urbana-Champaign, and a B.S.E.E. from Bucknell University.

Ilkay Altintas

Member

ILKAY ALTINTAS is the Chief Data Science Officer at the San Diego Supercomputer Center (SDSC), University of California San Diego (UCSD), where she is also the founder and director for the Workflows for Data Science Center of Excellence. Since joining SDSC in 2001, she has translated her technical and management skills to scientific computing and data science as both a principal investigator and leader. She is a co-initiator of the popular open-source Kepler Scientific Workflow System, and the co-author of publications related to computational data science at the intersection of workflows, provenance, distributed computing, big data, reproducibility, and software modeling in many different application areas. She is also a well-known Massive Open Online Course (MOOC) instructor in the field of “big” data science, and reached out to hundreds of thousands of learners across any populated part of our continent. Among the awards she has received are the 2015 Institute of Electrical and Electronics Engineers (IEEE) Technical Committee on Scalable Computing (TCSC) Award for Excellence in Scalable Computing for Early Career Researchers and the 2017 Association for Computing Machinery's (ACM's) Special Interest Group on High Performance Computing (SIGHPC) Emerging Woman Leader in Technical Computing Award. Dr. Altintas received her Ph.D. degree from the University of Amsterdam in the Netherlands, with an emphasis on provenance of workflow-driven collaborative science and she is currently an assistant research scientist at UCSD.

Shreyas Cholia

Member

SHREYAS CHOLIA is Group Leader for the Usable Software Systems Group in the Data Science and Technology department at Lawrence Berkeley National Laboratory (LBNL), focused on usability aspects of computational and data analysis systems. He is particularly interested in how web interfaces and tools can facilitate large-scale scientific computing workflows. He is currently working on various projects that integrate Jupyter Notebooks with distributed and high performance scientific computing environments. He joined LBNL's Computational Research Division (CRD) in 2015, having worked for over a decade at the National Energy Research Scientific Computing Center (NERSC), where he led the science-gateway, web, and grid efforts. Prior to his appointment at LBNL, he was a developer and consultant at IBM. He has a B.A. in Computer Science and Cognitive Sciences from Rice University.

Mercè Crosas

Member

MERCÈ CROSAS is Harvard University's Research Data Officer, with the Office of Vice Provost for Research (OVPR), and Chief Data Science and Technology Officer at Harvard's Institute for Quantitative Social Science (IQSS). In the last ten years, Dr. Crosas has been Principal Investigator (PI) and co-PI of multiple research grants and collaborations related to data privacy, data provenance, research reproducibility, and data sharing in social science, biomedicine, and astronomy. She is part of numerous committees and working groups focused on research data management, data citation, and data standards, and is a co-author of the FAIR (Findable, Accessible, Interoperable, Reusable) data principles as well as the Joint Declaration of Data Citation Principles. Before re-joining Harvard in 2004, Dr. Crosas worked for six years in the educational software and biotech industries, initially as a software developer, and subsequently as director of the software development team. She contributed to the development of lab information management systems (LIMS) for SNP discovery and genotyping and mass spectrometry. Before that, she spent six years at the Harvard-Smithsonian Center for Astrophysics, first as a pre-doctoral fellow for her Ph.D. in astrophysics from Rice University, and later as a post-doctoral fellow, researcher, and software engineer with the Radioastronomy division. She earned a B.S. in physics from the Universitat de Barcelona, Spain.

Alfred O. Hero, III

Member

ALFRED HERO is the John H. Holland Distinguished University Professor of Electrical Engineering and Computer Science and the R. Jamison and Betty Williams Professor of Engineering at the University of Michigan. His research is on data science and developing theory and algorithms for data collection, analysis and visualization that use statistical machine learning and distributed optimization. These are being applied to network data analysis, personalized health, multi-modality information fusion, data-driven physical simulation, materials science, dynamic social media, and database indexing and retrieval. Dr. Hero has held visiting positions at Massachusetts Institute of Technology, Boston University, Lucent Bell Laboratories (Murray Hill), Ford Motor Company in addition to the University of Nice, the École Normale Supérieure de Lyon, and Telecom-ParisTech in France. Dr. Hero was President of the Institute of Electrical and Electronics Engineers' (IEEE's) Signal Processing Society (2006-2008) and was on the Board of Directors of the IEEE (2009-2011) where he served as Director of Division IX (Signals and Applications). He is also a member of the Big Data Special Interest Group of the IEEE Signal Processing Society. Dr. Hero received the B.S. (summa cum laude) from Boston University (1980) and the Ph.D. from Princeton University (1984), both in Electrical Engineering.

Rebecca Lawrence

Member

REBECCA LAWRENCE is Managing Director of F1000 Group. She was responsible for the launch of the open research publishing platform F1000Research in January 2013, and has subsequently led the initiative behind the recent launch of Wellcome Open Research, Gates Open Research, and many other funder- and institution-based publishing platforms. She is a member of the High-Level Advisory Group for the European Commission's Open Science Policy Platform (OSPP), chairing their work on next-generation indicators and their integrated advice: OSPP-REC. She has been a co-Chair of a number of working groups focusing on data and peer review, for organisations including the Research Data Alliance (RDA) and ORCID. She is also an Advisory Board member for the data policy and standards initiative, FAIRsharing, and for DORA (the San Francisco Declaration on Research Assessment). She has worked in Scientific, Technical and Medical (STM) publishing for almost 20 years for several publishers including Elsevier where she built and ran the Drug Discovery Group. She originally trained and qualified as a pharmacist, and holds a Ph.D. in cardiovascular pharmacology from University of Nottingham.

Bradley A. Malin

Member

BRADLEY A. MALIN (NAM) is the vice chair for research and professor of biomedical informatics at Vanderbilt University. He is also a professor of biostatistics, a professor of computer science, and is affiliated faculty in the Center for Biomedical Ethics and Society. He co-directs the Health Data Science (HEADS) Center, the Center for Genetic Privacy and Identity in Community Settings (GetPreCiSe)—a National Institutes of Health (NIH) Center of Excellence in Ethical, Legal, and Social Implications Research (CEER), and the Big Biomedical Data Science Ph.D. program. He is also the director of the Health Information Privacy Laboratory (HIPLab), which was established to address the growing need for data privacy research and development for the health information technology sector. Dr. Malin's research is in big health data analytics and the infrastructure necessary to support such investigations. He has made specific contributions to a number of health-related areas, including distributed data processing methods for medical record linkage and predictive modeling, intelligent auditing technologies to protect electronic medical records from misuse in the context of primary care, and algorithms to formally anonymize patient information disseminated for secondary research purposes. He is an elected fellow of the National Academy of Medicine and American College of Medical Informatics and was honored as a recipient of the Presidential Early Career Award for Scientists and Engineers (PECASE). Dr. Malin completed his education at Carnegie Mellon University, where he received a bachelor's in biological sciences, a master's in machine learning, a master's in public policy and management, and a doctorate in computer science (with a focus on databases and software systems).

Lara Mangravite

Member

LARA MANGRAVITE is president of Sage Bionetworks. This organization is focused on the development and implementation of practices for large-scale collaborative biomedical research. Sage Bionetworks' work is centered on new approaches to scientific process that use open systems to enable community-based research regarding complex biomedical problems. Previously, Dr. Mangravite served as Director of the Systems Biology research group at Sage Bionetworks where she focused on the application of collaborative approaches to advance understanding of disease biology and treatment outcomes at a systems level with the overriding goal of improving clinical care. Dr. Mangravite obtained a B.S. in physics from the Pennsylvania State University and a Ph.D. in pharmaceutical chemistry from the University of California, San Francisco. She completed a postdoctoral fellowship in cardiovascular pharmacogenomics at the Children's Hospital Oakland Research Institute.

Tapio Schneider

Member

TAPIO SCHNEIDER is Theodore Y. Wu Professor of Environmental Science and Engineering at Caltech and Senior Research Scientist at the Jet Propulsion Laboratory. His research is focused on understanding atmosphere dynamics on Earth and other planets; turbulence in atmosphere and oceans; and climate change and climate modeling. Previously, Dr. Schneider served as professor of climate dynamics at Swiss Federal Institute of Technology Zurich from 2013-2016, and associate research scientist at New York University's Courant Institute of Mathematical Sciences from 2000-2002. Dr. Schneider received his M.Sc. (1997) and Ph.D. (2001) in atmospheric and oceanic sciences from Princeton University. He was a Visiting Graduate Student (Physics) at the University of Washington, Seattle from 1994-1995, and studied mathematics and physics (Vordiplom 1993) at Albert-Ludwigs-Universität Freiburg, Germany.

Thomas S. Arrison

Staff Officer