

DSI's Newest Center To Host the 2019 New York Scientific Data Summit



Columbia University is co-organizing and hosting the 2019 New York Scientific Data Summit, a two-and-a-half day symposium to explore data-driven discovery and innovation in science and industry. The summit, scheduled for June 12-14 in Columbia's Davis Auditorium, and co-organized by Brookhaven National Laboratory, will gather leading scientists and researchers from academia, national laboratories and industry who are working in diverse areas such as physics, astronomy, biomedical informatics and climate science, as well as large-scale computer systems and data science.



Registration for the event, which is open to the public, is available at the symposium web site.

The summit aims to become the leading scientific data event in the region, uniting researchers from the New York area as well as from across the nation and abroad. The event will be an opportunity for them to make connections and forge collaborations on large-scale initiatives and proposals relating to data-intensive science.

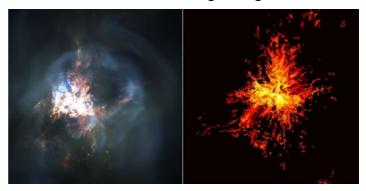
Participating speakers at the symposium are attending from six regional universities (Columbia, Rutgers, Yale, Princeton, UPenn, and Stony Brook), four national laboratories (Brookhaven, Argonne, Lawrence Berkeley, Oak Ridge), several leading institutes (NASA Goddard Institute for Space Studies, Flatiron Institute), a major regional research laboratory (D.E. Shaw Research), and

other prominent universities (UT Austin, TU Dresden, Cambridge, KAUST). More than 20 research posters will also be on display during the symposium. Click <u>here</u> to view the full program.

The summit is the fifth in a symposia series established by Brookhaven National Laboratory (BNL) and led by its <u>Computational Science Initiative</u> (CSI), which integrates computer science, applied mathematics, computational science, and scientific expertise and investments at the laboratory. This year, the summit is co-organized by Columbia's <u>Computing Systems for Data-Driven Science</u> Center, based in the <u>Data Science Institute</u> (DSI). The new center is a nexus at DSI for research in large-scale computer systems design, data analytics, and applications to problems in science, engineering and medicine. The center houses some of the top researchers in these fields, ten of whom will give research talks during the symposium.

"Invaluable collaborations among top researchers in data-driven science will come out of this summit," said <u>Steven Nowick</u>, a professor of Computer Science at Columbia who founded and chairs the <u>Computing Systems for Data-Driven Science</u> Center and is co-organizing the event.

"We are all awash in data, and we need to contend with the massive scale of the data being generated in all fields," added Nowick. "How do we develop and apply platforms to evaluate millions of medical records or petabytes of big data emanating from physics, astronomy and climate change research? And how do we visualize, analyze and interpret this data? It's a daunting and exciting endeavor,



pulling togethers researchers who exploit and design large-scale computing systems and algorithms, and who apply AI and machine learning, to enhance data-driven research."

The symposium will be organized into six topics:

- Streaming Data Analysis and Large-Scale Simulation
- Scalable Algorithms and Computer Systems for Scientific Applications
- Large-Scale Image Analysis and Mapping
- Focus Topic #1: Biomedical Informatics
- Focus Topic #2: Earth and Climate Science
- Focus Topic #3: Computational Astrophysics and Cosmology

The five keynote speakers are a diverse group of industrial and academic leaders working on the application and development of computer systems for data-driven science. They are:

<u>David Keyes</u>, professor of applied mathematics and computational science at King Abdullah University of Science and Technology (KAUST), and director of its Extreme Computing Research Center, and also an adjunct professor at Columbia University. He works at the algorithmic interface between parallel computing and numerical analysis of partial differential equations, with a focus on emerging computer architectures and their applications to critical scientific problems.

Mark Moraes, head of engineering at D.E. Shaw Research, a biochemistry research company based in New York City. The company has developed an advanced supercomputer, and supporting software and technologies for molecular dynamics simulation, with application to drug discovery.

<u>Gavin Schmidt</u>, a climate scientist and director of the NASA Goddard Institute for Space Studies and the Center for Climate Systems Research at Columbia.



Schmidt, also an adjunct researcher at <u>the Earth Institute</u>, develops advanced models that evaluate climate change.

<u>Rick Stevens</u>, a computer scientist at University of Chicago, and associate lab director for computing, environment and life sciences at Argonne National Laboratory. He develops tools and techniques to enable computational scientists to solve large-scale problems on high-performance computers.

<u>Karen Willcox</u>, University of Texas, Austin. Willcox is director of the Institute for Computational Engineering and Sciences and a professor of Aerospace Engineering and Engineering Mechanics at the University of Texas. Her research has produced computational methods for the design of next-generation engineered systems. Her methods are widely used in aircraft system design and environmental policy decision-making.

And <u>Frank Alexander</u>, the deputy director of Brookhaven National Laboratory's Computational Science Initiative, said he and his colleagues are excited to continue the New York Scientific Data Summit this year in partnership with Columbia's Data Science Institute.

"This partnership extends the long tradition of collaboration between our two institutions," said Alexander. "We are looking forward to a great slate of talks from a strong lineup of speakers from the New York area and beyond."

-By Robert Florida