# New York Scientific Data Summit 2021

## Data-Driven Discovery in Science and Industry

Organized by: Brookhaven National Laboratory, Carnegie Mellon University, Columbia University in the City of New York, Los Alamos National Laboratory, North Carolina Agricultural and Technical State University, Stony Brook University, Texas – The University of Texas at Austin

# AGENDA

## Tuesday, October 26 – Day 1

| | |
|---|---|
| **12:00 – 12:10PM** | *Opening & Welcome Remarks,* Kerstin Kleese van Dam (Brookhaven National Laboratory) |

### Session #1, DOE/National Security

### Session Chair: Francis J. Alexander (Brookhaven National Laboratory)

| | |
|---|---|
| **12:10 – 12:45 PM** | *Opportunities and Challenges for AI-Enhanced Decision Making in Nuclear Proliferation Detection,* Angela Sheffield (National Nuclear Security Administration, Department of Energy) |
| **12:45 – 1:20 PM** | *Large-scale, data-driven methods and applications developed as part of the Exascale Computing Project*, Lori Diachin (Lawrence Livermore National Laboratory) |
| **1:20 – 1:55PM** | *Applying HPC and AI/ML Capabilities for Stockpile Stewardship Mission*, Thuc Hoang (National Nuclear Security Administration Department of Energy |
| **1:55 – 2:30 PM** | *Decision Support for Wicked Problems,* Robert Lempert (RAND Corporation, and Frederick S. Pardee Center for Longer Range Global Policy and the Future Human Condition) |
| **2:30 – 2:45 PM** | *Virtual Coffee Break* |

### Session #2, Infrastructure and Natural Hazards

### Session Chair: Hubertus van Dam (Brookhaven National Laboratory)

| | |
|---|---|
| **2:45 – 3:20 PM** | *Behaviors of People and Systems: Decision Problems and Data Driven Analysis,* Ramayya Krishnan (Carnegie Mellon University) |
| **3:20 – 3:55 PM** | *Conflict, Coordination & Control: Do We Understand the Actual Rules Used to Balance Flooding, Energy, and Agricultural Tradeoffs in River Basins,* Patrick Reed (School of Civil and Environmental Engineering, Cornell) |
| **3:55 – 4:30 PM** | *Anomaly Detection for Identifying Fire Hazards in Power Distribution Systems,* Jhi-Young Joo (Lawrence Livermore National Laboratory) |
| **4:30 – 5:10 PM** | **Panel Discussion,** Moderator: Francis J. Alexander (Brookhaven National Laboratory)<br>**Panelists: Day 1 Speakers** |
| **5:10 – 5:40 PM** | *Poster Session & Exhibit Booths on vFairs* |

# AGENDA

## Wednesday, October 27 – Day 2

**11:00AM–12:00PM** *Poster Session & Exhibit Booths on vFairs*

### Session #3, Climate Modeling

### Session Chair: Meifeng Lin (Brookhaven National Laboratory)

**12:00 – 12:35 PM** *Harnessing Modern Computational Resources to Evaluate Interactions Between the Economy and Climate,* Kenneth Judd (Hoover Institution, Stanford University)

**12:35 – 1:10 PM** *Modeling sea-level rise and its uncertainties under climate change,* Minghua Zhang (State University of New York at Stony Brook)

**1:10 – 1:45 PM** *Data Assimilation and Its Connections with Uncertainty Quantification, Forecast and Machine Learning,* Nan Chen (University of Wisconsin-Madison)

**1:45 – 2:20 PM** *Tropical cyclone risk modeling,* Michael Tippett (Columbia University)

**2:20 – 2:35 PM** *Virtual Coffee Break*

### Session #4, Climate Environmental Impact

### Session Chair: Gregory Goins (North Carolina A&T State University)

**2:35 – 3:10 PM** *Microbial controls of climate-smart soil health management practices,* Arnab Bhowmik (North Carolina A&T State University)

**3:10 – 3:45 PM** *From model-based to data-driven remote sensing in environmental management*, Leila Hashemi-Beni (North Carolina A&T State University)

**3:45 – 4:20 PM** *Planning for Spatially Correlated Failures in Coupled Natural Gas and Power Transmission via Stochastic Optimization,* Seth Blumsack (The Pennsylvania State University)

**4:20 – 5:00 PM** **Panel Discussion***,* Moderator: Nathan Urban (Brookhaven National Laboratory)
**Panelists: Day 2 Speakers**

**5:00 – 5:30 PM** *Poster Session & Exhibit Booths on vFairs*

# AGENDA

## Thursday, October 28 – Day 3

**11:00 AM–12:00 PM** *Poster Session & Exhibit Booths on vFairs*

### Session #5, Decision Making Under Uncertainty

### Session Chair: Omar Ghattas (University of Texas at Austin)

**12:00 – 12:35 PM** ***Physics-Informed Deep Learning for Traffic State Estimation and Fundamental Diagram Discovery,*** Xuan (Sharon) Di (Columbia University)

**12:35 – 1:10 PM** ***Machine Learning and Game Theory for Societal Challenges,*** Fei Fang (Carnegie Mellon University)

**1:10 – 1:45 PM** ***Deep Generative Surrogate Modeling and Inversion in Subsurface Flows,*** Nicholas Zabaras (University of Notre Dame)

**1:45 – 2:20 PM** ***Propagation of Uncertainty from Data to Inference for Large-Scale Inverse Problems with Application to Ice Sheet Flow,*** Noemi Petra (University of California, Merced)

**2:20 – 2:35 PM** *Virtual Coffee Break*

### Session #6, Digital Twins

### Session Chair: Byung-Jun Yoon (Brookhaven National Laboratory)

**2:35 – 3:10 PM** ***A probabilistic graphical model foundation to enable predictive digital twins at scale,*** Michael Kapteyn (Oden Institute for Computational Engineering and Sciences)

**3:10 – 3:45 PM** ***Tracking and Forecasting Epidemics,*** Roni Rosenfeld (Carnegie Mellon University)

**3:45 – 4:20 PM** ***Challenges and Opportunities of Translating AI for Healthcare: Why Many AI Tools Do Not Show Impact in Combating COVID-19?!,*** May Wang (Georgia Institute of Technology and Emory University)

**4:20 – 5:00 PM** **Panel Discussion,** Moderator: Qiang Du (Columbia University)
**Panelists: Day 3 Speakers**

**5:00 – 5:30 PM** *Poster Session & Exhibit Booths on vFairs*

## Friday, October 29 – Day 4 Workshop:
## Challenges and Opportunities in Autonomous Scientific Facilities

**11:00 AM – 12:00 PM** *Poster Session & Exhibit Booths on vFairs*

### Session #7, Session Chair: Jerome Lauret (Brookhaven National Laboratory)

**12:00 – 12:35 PM** ***A Perspective on Autonomous Experimentation and Discovery,*** Marcus Noack (Lawrence Berkeley National Laboratory)

**12:35 – 1:10 PM** ***Edge ML Detectors: Integrating Machine Learning with Sensors,*** Sioan Zohar (Brookhaven National Laboratory)

**1:10 – 1:45 PM** ***Co-Design Methodologies for Edge Computing ASICs in Scientific Research Environments,*** Sandeep Miryala (Brookhaven National Laboratory)

**1:45 – 2:20 PM** ***ML/AI applications in astronomy,*** Anze Slosar (Brookhaven National Laboratory)

**2:20 – 2:35 PM** *Virtual Coffee Break*

### Session #8, Session Chair: David Brown (Brookhaven National Laboratory)

**2:35 – 3:10 PM** ***Nuclear data evaluation with Bayesian networks,*** Georg Schnabel (International Atomic Energy Agency (IAEA))

**3:10 – 3:45 PM** ***Automating Particle Accelerator Operations with Machine Learning,*** Jonathan Edelen (RadiaSoft)

**3:45 – 4:20 PM** ***Remote and on-the-fly: artificial intelligence driven science in laboratories and central facilities,*** Phillip Maffettone (Brookhaven National Laboratory)

**4:20 – 5:00 PM** ***Panel Discussion, Moderator:*** Kevin Yager (Brookhaven National Laboratory)
Panelists: Day 4 Speakers

*-Conference End-*

# DAY 1 TALKS

## TUESDAY, OCTOBER 26

**SPEAKER:**

**Angela M. Sheffield**
Eisenhower School, National Defense University
National Nuclear Security Administration, Department of Energy

*angela.sheffield@nnsa.doe.gov*

**TIME: 12:10 – 12:45PM**

**TITLE:** Opportunities and Challenges for AI-Enhanced Decision Making in Nuclear Proliferation Detection

**ABSTRACT:** Working with allies and partners around the globe, the U.S. employs nuclear nonproliferation measures and programs to control the spread of nuclear weapons-usable capabilities and reduce the threat of the nuclear weapons. Nuclear nonproliferation is underpinned by technologies and science-based approaches to detect nuclear proliferation activities and predict and characterize the capabilities of nuclear weapons programs based on observable signatures and indicators. Nuclear proliferation detection technologies provide insights to analysts and decision makers to inform nuclear nonproliferation programs and strategies for intervention. Modern advances in computing, artificial intelligence, and new data sources present new opportunities to enhance and expand nuclear proliferation detection.

Within the Department of Energy's National Nuclear Security Administration, the Office of Defense Nuclear Nonproliferation Research and Development (DNN R&D) is the leading U.S. government organization for the development of advanced technologies in support of the United States' nuclear nonproliferation and nuclear security goals. Leveraging the expertise of the Department of Energy National Laboratories and working with partners in industry and academia, DNN R&D drives advances in the science of artificial intelligence and works with partners across the U.S. government to accelerate adoption of AI-enabled technologies to enhance U.S. nuclear proliferation detection.

This presentation will provide a technical overview of decision making in nuclear proliferation detection and discuss principles, challenges, opportunities, and examples of successful approaches in designing intelligent systems to support and inform these high-consequence decisions.

**BIO:** Angela M. Sheffield is a student at the Eisenhower School for National Security and Resource Strategy at the National Defense University. She is on detail from the Department of Energy (DOE) National Nuclear Security Administration (NNSA), where she is the Senior Program Manager for Data Science and Artificial Intelligence (AI) at U.S. Office of Defense Nuclear Nonproliferation Research and Development (DNN R&D). An internationally recognized expert in nuclear nonproliferation and AI, Ms. Sheffield leads the U.S. government's premier program to develop the Next Generation of AI to transform national security and fulfill mission requirements across the U.S. government to prevent the spread of nuclear weapons.

Prior to joining DOE, Ms. Sheffield worked as a scientist at DOE's Pacific Northwest National Laboratory (PNNL), where she led project teams to develop predictive modeling and analytics methodologies to reduce the threat of terrorism and proliferation of weapons of mass destruction. Ms. Sheffield has dedicated her career to driving science and technology innovations to address the toughest challenges in national security. She joined PNNL after a distinguished career as an active-duty officer in the U.S. Air Force, where she specialized in the research and development and technical intelligence of U.S. and adversary weapon systems. Ms. Sheffield holds a B.S. in Economics from the United States Air Force Academy and an M.S. in Operations Research from Kansas State University.

**SPEAKER:**

**Lori Diachin**
Deputy Director, Exascale Computing Project
Lawrence Livermore National Laboratory

*diachin2@llnl.gov*

**TIME: 12:45 – 1:20PM**

**TITLE:** Large-scale, data-driven methods and applications developed as part of the Exascale Computing Project

**ABSTRACT:** The Department of Energy's (DOE's) Exascale Computing Project (ECP) is advancing the applications and software technologies necessary to ensure the first US-based exascale systems, expected to be delivered in 2021-2023, are being used to address critical national imperatives in discovery science, national security, and energy assurance. This talk will first provide a high-level overview of the ECP and then focus on the applications, algorithms, and software being developed to address large scale data-driven challenges. In particular, we will highlight several projects with ongoing work in the development of exascale machine learning methodologies, online data analysis and reduction methods, community interfaces and workflow tools that support these techniques at scale, and we conclude by highlighting their use in applications relevant to the DOE mission areas.

**BIO:** Lori Diachin is the Deputy Director for the U.S. Department of Energy's Exascale Computing Project (ECP). ECP is a collaborative Department of Energy effort supported by both the National Nuclear Security Administration and the Office of Science to accelerate the delivery of a capable exascale computing ecosystem for breakthroughs in scientific discovery, energy assurance, economic competitiveness, and national security.  Lori also serves as the Principal Deputy Associate Director (PDAD) in the Computing Directorate at Lawrence Livermore National Laboratory. The Computing Directorate has approximately 1200 staff serving the needs of the laboratory in areas ranging from high performance computing, computing sciences for the missions of LLNL, data science, and information technology for business and workforce enablement.

Lori Diachin has over 30 year experience in high performance computing and applied mathematics research where her areas of expertise include mesh quality improvement, mesh component software, numerical methods, and parallel computing. She is the co-author of over 50 technical journal articles, book chapters, and conference proceeding articles in these areas. Before joining LLNL, Lori was a computer scientist at Argonne National Laboratory and a Member of the Technical Staff at Sandia National Laboratory. Lori received her bachelor's degree in Mathematics from Edinboro University of Pennsylvania in 1988 and her Ph.D. in Applied Mathematics from University of Virginia in 1992. She joined Lawrence Livermore National Laboratory in 2003.

## DAY 1 TALKS

### TUESDAY, OCTOBER 26

**SPEAKER:**

Thuc Hoang
National Nuclear Security Administration
Department of Energy

*thuc.hoang@nnsa.doe.gov*

**TIME: 1:20 – 1:55PM**

**TITLE:** Applying HPC and AI/ML Capabilities for Stockpile Stewardship Mission

**ABSTRACT:** The Advanced Simulation and Computing (ASC) program is a cornerstone of the National Nuclear Security Administration (NNSA)'s Stockpile Stewardship Program (SSP), providing simulation capabilities and computational resources to support the annual stockpile assessment and certification process, study advanced nuclear weapons design and manufacturing processes, analyze accident scenarios and weapons aging, and provide the tools to enable stockpile Life Extension Programs (LEPs) and resolution of Significant Finding Investigations (SFIs). To carry out its mission, ASC continually focuses on maximizing the utilization and efficiency of its high-performance computing (HPC) and artificial intelligence (AI) systems deployed at the NNSA Laboratories as it strives to increase the NNSA science-based predictive capabilities.

**BIO:** Thuc Hoang is the Director of the Office of Advanced Simulation and Computing (ASC) and Institutional Research and Development Programs in the Office of Defense Programs, within the Department of Energy National Nuclear Security Administration (NNSA). The ASC program develops and deploys high-performance simulation capabilities and computational resources to support the NNSA annual stockpile assessment and certification process, and other nuclear security missions. Ms. Hoang manages ASC's research, development, acquisition, and operation of high-performance computing (HPC) systems, in addition to the NNSA Exascale Computing Initiative and future computing technology portfolio. She has served on proposal review panels and advisory committees for the National Science Foundation, Department of Defense and DOE Office of Science, as well as for some other international HPC programs. Ms. Hoang holds a Bachelor of Science in Electrical Engineering from Virginia Tech and a Master of Science in Electrical Engineering from Johns Hopkins University.

**SPEAKER:**

**Robert Lempert**
Principal Researcher
RAND Corporation and Director of the Frederick S. Pardee Center for Longer Range Global Policy and the Future Human Condition

*lempert@rand.org*

**TIME:** **1:55-2:30PM**

**TITLE:** Decision Support for Wicked Problems

**ABSTRACT:** Quantitative information is often necessary for good decisions. But providing effective quantitative analysis can be a particular challenge for so-called “wicked problems,” which are characterized by the presence of deep uncertainty, contested interests and values, unclear system boundaries, and often non-linear dynamics. In addition, as the ability to simulate complex systems improves, so too does the need for quantitative decision support methods that can make use of the unique types of information such simulations provide about a fast-changing, contingent, often hard-to-predict world. In recent years, decision making under deep uncertainty (DMDU) methods have enabled significant advances in decision support under such conditions. These approaches are made possible by advanced computational capabilities, data analysis, and visualizations methods and are specifically designed to help identify and adjudicate tradeoffs in the presence of deep uncertainty. Drawing on examples from climate change and other policy areas, this talk will survey the history and current application of DMDU and address directions for the future.

**BIO:** Robert Lempert is a principal researcher at the RAND Corporation and Director of the Frederick S. Pardee Center for Longer Range Global Policy and the Future Human Condition. His research focuses on risk management and decision-making under conditions of deep uncertainty. Dr. Lempert’s work aims to advance the state of art for organizations managing risk in today’s conditions of face-paced, transformative, and surprising change and helping organizations adopt these approaches to help make proper stewardship of the future more commonly practiced. Dr. Lempert is a Fellow of the American Physical Society, a member of the Council on Foreign Relations, a coordinating lead author for Working Group II of the United Nation’s Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report, a chapter lead for the Fourth US National Climate Assessment, chair of the peer review panel for California’s Fourth Climate Assessment, a member of California’s Climate-Safe Infrastructure Working Group, and has been a member of numerous study panels for the U.S. National Academies, including America’s Climate Choices and Informing Decisions in a Changing Climate. Dr. Lempert was the Inaugural EADS Distinguished Visitor in Energy and Environment at the American Academy in Berlin and the inaugural president of the Society for Decision Making Under Deep Uncertainty (http://www.deepuncertainty.org). A Professor of Policy Analysis in the Pardee RAND Graduate School, Dr. Lempert is an author of the book *Shaping the Next One Hundred Years: New Methods for Quantitative, Longer-Term Policy Analysis*.

# DAY 1 TALKS

## TUESDAY, OCTOBER 26

**SPEAKER:**

Ramayya Krishnan
W. W. Cooper and Ruth F. Cooper Professor of Management Science and Information Systems at Heinz College and the Department of Engineering and Public Policy
Carnegie Mellon University

*rk2x@cmu.edu*

**TIME:** **2:45-3:20PM**

**TITLE:** Behaviors of People and Systems: Decision Problems and Data Driven Analysis

**ABSTRACT:** This is a talk motivated by and set in the context of real world applications. Using prior and ongoing work on consumer behavior in the presence of social influence as well as on resilient supply chains, I will introduce relevant decision problems that arise in these contexts and motivate the need for a multi-disciplinary approach to address them. Specifically, I will highlight how the work brings together concepts and methods from social science, choice models from statistical analysis and optimization models from operations research. The opportunity to socialize ideas across these disciplinary boundaries is the objective of a new series of CCC-ACM SIGAI -INFORMS workshops which the speaker is helping organize with a group of colleagues.

**BIO:** Ramayya Krishnan is the W. W. Cooper and Ruth F. Cooper Professor of Management Science and Information Systems at Heinz College and the Department of Engineering and Public Policy at Carnegie Mellon University. He is Dean of the Heinz College of Information Systems and Public Policy. Krishnan's research has focused on the development of decision support tools to analyze, interpret, and act on consumer and social behavior in digital and networked platforms. He is an INFORMS Fellow and an elected member of the National Academy of Public Administration. He served as President of INFORMS (the Operations Research Society in 2019) and is currently co-organizing a series of workshops on AI and Operations Research that is being sponsored by CCC-ACM SIGAI and INFORMS. More information about him is available at Ramayya Krishnan | Carnegie Mellon University's Heinz College (cmu.edu)

**SPEAKER:**

Patrick Reed
Joseph C. Ford Professor of Engineering
School of Civil and Environmental Engineering, Cornell University

*patrick.reed@cornell.edu*

*http://reed.cee.cornell.edu*

**TIME:** **3:20-3:55PM**

**TITLE:** Conflict, Coordination & Control: Do We Understand the Actual Rules Used to Balance Flooding, Energy, and Agricultural Tradeoffs in River Basins?

**ABSTRACT:** Multi-reservoir systems require adaptive control policies capable of managing evolving hydroclimatic variability and human demands across a wide range of time scales. However, traditional operating rules are static, ignoring the potential for coordinated information sharing to reduce conflicts between multi-sectoral river basin demands. This study shows how recent advances in high-performance computing and multi-objective control enable the design of coordinated operating policies that continuously adapt as a function of evolving hydrologic information. The benefits of the proposed control innovations are demonstrated for the Red River basin of Vietnam, where four major reservoirs serve to protect the capital of Hanoi from flooding, while also supplying farmers with irrigable water supply and the surrounding region with electric power. Operating policies recently proposed by the Vietnamese government seek to improve coordination and adaptivity in the Red River using a conditional if/then/else rule system that triggers alternative control actions using information on current storage and recent hydrology. However, these simple, discontinuous rules fail to protect Hanoi to even the 100-yr flood. These findings accentuate the need to transition from static rule curves to dynamic operating policies in order to manage evolving hydroclimatic variability and socioeconomic change.

**BIO:** Dr. Reed's Decision Analytics for Complex Systems research group has a strong focus on the sustainability of Food-Energy-Water systems given conflicting demands from ecosystem services, expanding populations, and climate change. The tools developed in Dr. Reed's group bridge complexity science, risk management, economics, multiobjective decision making, artificial intelligence, and high performance computing. Engineering design and decision support software developed by Dr. Reed is being used broadly in academic, governmental, and industrial application areas with thousands of users globally. The management modeling tools developed by the Reed Research Group combine multiobjective optimization, high performance computing, and advanced spatiotemporal visualization and uncertainty modeling techniques.

## DAY 1 TALKS

### TUESDAY, OCTOBER 26

**SPEAKER:**

**Jhi-Young Joo**
Distribution Automation Lead Engineer
Lawrence Livermore National Laboratory

*joo3@llnl.gov*

**TITLE: 3:55-4:30PM**

**TITLE:** Anomaly Detection for Identifying Fire Hazards in Power Distribution Systems

**ABSTRACT:** Wildfires caused by electric equipment have become major concern for utilities in vulnerable regions, as witnessed by the scale and cost of recent disasters such as the 2018 Camp Fire and 2019 Kincade Fire. Part of the challenge in preventing such events is a lack of systems for monitoring equipment's condition. Even routine inspections of equipment may not suffice for detecting potential issues. In the meantime, high-speed, high-resolution sensing and measurement technologies are becoming more prevalent in power systems operations. These measurements can be used to detect unique signatures of equipment malfunction and anomalies such as arcing faults that can potentially cause outages and wildfires. This presentation will discuss the methods we used in identifying anomalies in synchrophasor measurements of three-phase voltage and current at a partner utility distribution substation. We used a combination of unsupervised and supervised learning methods to effectively identify anomalies including normal operation events, as well as potential arcing events. We will conclude with remarks on lessons learned from navigating actual utility operational and measurement data and current/future directions of the work.

**BIO:** Dr. Jhi-Young Joo is Distribution Automation Lead Engineer at Lawrence Livermore National Laboratory. Her current research is focused on data analytics for distribution system operation and planning. Before joining LLNL in 2018, she worked as a research scientist at Lawrence Berkeley National Laboratory, and as a tenure-track faculty member at Missouri University of Science and Technology. She received her Ph.D. from Carnegie Mellon University, and M.S. and B.S. from Seoul National University, all in Electrical and Computer Engineering.

**SPEAKER:**

**Kenneth Judd**
Paul H. Bauer Senior Fellow, Hoover Institution
Stanford University

*kennethjudd@mac.com*

**TIME:** **12:00– 12:35PM**

**TITLE:** Harnessing Modern Computational Resources to Evaluate Interactions Between the Economy and Climate.

**ABSTRACT:** Climate scientists use the most powerful computers to model climate dynamics but ignore important interactions between the climate and the economy. Economists use laptops to evaluate alternative climate change policies, using primitive models of both the economy and climate. The reality of climate change and economics is far more challenging than either problem. Climate modeling is an initial value problem. Economic problems are far more challenging in that today's decisions by economic actors depend what they expect to happen in the future forcing us to solve two-point boundary value problems. A serious examination of climate change and economic behavior requires combining the climate and economic systems, and solving complex stochastic two-point boundary value problems. It is infeasible to solve a system of equations that fully captures the complexity and uncertainty of both the climate and economic systems, but it is possible to examine models that treat both the climate and economic systems in a serious manner. DSICE is the framework that I have developed, and applied to evaluate the social cost of carbon, the feasibility of achieving a two-degree target, and evaluating the uncertainties we face over the current century. This is accomplished by utilizing computational machinery orders of magnitude greater than any modeling ever done before in economics.

**BIO:** Dr. Kenneth L. Judd is the Paul H. Bauer Senior Fellow at the Hoover Institution. He has worked on the economics of taxation, imperfect competition, microeconomic theory, and mathematical economics. In the past thirty years, he has focused on applying modern computational tools and methods to solving complex economic models. His book Numerical Methods in Economics was published by MIT Press in 1998, and is the leading graduate textbook in that subject. In the past ten years, he has built the most computationally intensive models used by economists to analyze climate change policies. He earned a BS in Computer Science, a BS and MS in Mathematics, and a PhD in Economics, all at the University of Wisconsin.

## DAY 2 TALKS

WEDNESDAY, OCTOBER 27

**SPEAKER:**

Minghua Zhang
Distinguished Professor of Atmospheric Sciences
State University of New York at Stony Brook

*minghua.zhang@stonybrook.edu*

**TIME: 12:35– 1:10PM**

**TITLE:** Modeling sea-level rise and its uncertainties under climate change

**ABSTRACT:** The presentation will start with methods to simulate global and regional sea-level rises in response to global warming. Sources and characterization of uncertainties will be described. Simulated sea-level rise along the New York coasts in the 21st century will be presented. The presentation will end with a case study to use simulated sea-level rise with consideration of uncertainties in decision making.

**BIO:** Minghua Zhang is Distinguished Professor of Atmospheric Sciences in the State University of New York at Stony Brook, where he previously served as Director of the Institute for Terrestrial and Planetary Atmospheres, and Dean of the School of Marine and Atmospheric Sciences. Minghua Zhang's research is about numerical modeling of climate and global climate change, in which he has published over 150 peer-reviewed articles. For his climate modeling research, he served for ten years as co-Chair of the Atmospheric Model Working Group of the Community Earth System Model that is administered at the National Center for Atmospheric Research (NCAR). Minghua has a Ph.D. from the Institute of Atmospheric Physics/Chinese Academy of Sciences. He is a fellow of the American Meteorological Society and AAAS. He currently serves as Editor-in-Chief of the Journal of Geophysical Research (JGR)–Atmospheres.

**SPEAKER:**

**Nan Chen**
Assistant Professor, Department of Mathematics
University of Wisconsin-Madison

*chennan@math.wisc.edu*

**TIME: 1:10 – 1:45PM**

**TITLE:** Data Assimilation and Its Connections with Uncertainty Quantification, Forecast and Machine Learning

**ABSTRACT:** Data assimilation combines model with observational data to improve the results. It has been widely used to improve the initialization of ensemble forecast for complex turbulent systems. In this talk, we will present some recent development of other important roles of data assimilation. We will start with incorporating information theory into nonlinear data assimilation for uncertainty quantification of state estimation. Then we will show that data assimilation can be used to provide an efficient algorithm for solving high-dimensional Fokker-Planck equations. Finally, we will illustrate the combination of data assimilation and machine learning for improving the forecast of turbulent systems in the presence of only short observational data and imperfect models. Applications in climate science will be presented when discussing these topics.

**BIO:** Nan Chen is an Assistant Professor at the Department of Mathematics, University of Wisconsin-Madison. He is also a faculty affiliate of the Institute for Foundations of Data Science. Nan received his PhD degree from the Courant Institute of Mathematical Sciences and the Center of Atmosphere and Ocean Science, New York University (NYU) in 2016. He worked as a postdoc research associate at NYU for two years before joining UW-Madison. Nan's research interests lie in the applied mathematics, geophysics, dynamical models, stochastic methods, numerical algorithms, machine learning techniques and general data science. He is also active in developing both dynamical and stochastic models and use these models to predict real-world phenomena related to atmosphere ocean science, climate and other complex systems such as the Madden-Julian Oscillation (MJO), the monsoon, the El Nino Southern Oscillation (ENSO) and the sea ice based on real observational data. His recent work also involves the development of new uncertainty quantification and stochastic methods to material science.

## DAY 2 TALKS

WEDNESDAY, OCTOBER 27

**SPEAKER:**

**Michael Tippett**
Associate Professor, Department of Applied Physics and Applied Mathematics
Columbia University

*mkt14@columbia.edu*

**TIME: 1:45 – 2:20PM**

**TITLE:** Tropical cyclone risk modeling

**ABSTRACT:** Tropical cyclones (TCs; including hurricanes and typhoons) are a leading cause of natural disasters and were responsible for an estimated USD78 billion in global economic losses in 2020. To know the TC risk at a particular location requires quantifying the full range of possible TC impacts, especially extreme ones, at that location. However, the relative shortness of the historical TC record means it can only hint at the risk from the most extreme storms. Also, observations of past TCs provide little indication of what to expect in a warming climate. Here I will describe how the open-source Columbia HAZard model (CHAZ) addresses these two issues with a combination of data and physics-based reasoning. First, CHAZ takes a Monte Carlo approach and is able to generate large numbers of synthetic TCs. Second, CHAZ generates TCs in a manner that is conditional on climate quantities, which allows it to estimate TC risk in both the current climate and in projected future climates. I will describe the structure of the basic CHAZ elements for TC genesis, track, and intensification, along with new models for symmetric and asymmetric wind fields, and future plans. I will briefly describe the central role that CHAZ is playing in the new Columbia World Project: Hurricane Risk Models for Vulnerable Populations.

**BIO:** Michael Tippett is an Associate Professor in the department of Applied Physics and Applied Mathematics at Columbia University. He analyzes data from numerical weather prediction models, meteorological observations, and storm reports to find patterns that can improve understanding, facilitate prediction, and manage risk. Tippett's climate research ranges from the detection and attribution of climate change in models and observations on centennial time-scales, to seasonal forecasts of the El Niño-Southern Oscillation (ENSO). On weather time-scales, he investigates how severe thunderstorms (those resulting in tornadoes, hail, or damaging wind) and tropical cyclones are related to climate, now and in the future.

**SPEAKER:**

**Arnab Bhowmik**
Assistant Professor of Soil Science and Soil Microbiology, Department of Natural Resources and Environmental Design
College of Agriculture and Environmental Sciences
North Carolina A&T State University (NC A&T)

*abhowmik@ncat.edu*

**TIME:** **2:35-3:10PM**

**TITLE:** Microbial controls of climate-smart soil health management practices

**ABSTRACT:** Climate-smart agricultural management practices have the potential to provide ecosystem services including soil carbon accrual and soil health enhancement. However, if not properly managed these agroecosystems create conditions for increased greenhouse gas emissions (e.g. nitrous oxide) thereby offsetting the benefits. Novel and potential biological soil health indicators were measured at field and lab scales in a wide range of agroecosystems in the United States. These included experimental plots affiliated with North Dakota State University, Washington State University, Pennsylvania State University and North Carolina A&T State University. Our studies evaluate the effect of soil health management practices like reduced tillage, incorporation of animal wastes, compost, biochar, cover cropping and crop rotations on soil microbial communities, their functions and greenhouse gas emissions. Results indicate that better understanding of the key microbial players involved in nitrogen cycling and their response to different soil conditions are crucial for optimizing soil health benefits and minimizing tradeoffs.

**BIO:** Arnab Bhowmik, PhD is an Assistant Professor of Soil Science and Soil Microbiology in the Department of Natural Resources and Environmental Design at the College of Agriculture and Environmental Sciences at North Carolina A&T State University (NC A&T). Dr. Bhowmik's Soil Sustainability Lab at NC A&T is interested in studying mechanisms that govern soil organic matter dynamics and the biological networks that provide ecosystem services in sustainable agroecosystems. Specifically, it comprises of interpreting soil microbial and biogeochemical processes that regulate nutrient cycling (mainly carbon and nitrogen), biological soil health indicators, hemp best management practices and greenhouse gas emissions from climate adaptive agricultural management practices.

Since earning his doctorate in 2016 from North Dakota State University, Dr. Bhowmik has worked as a post-doctoral fellow at The Pennsylvania State University to continue collaborating on USDA-NIFA funded projects to study biological aspects of soil for climate-smart agriculture. He has published his research in reputed journals and also won several awards including Academic Gold Medal during MS and Distinguished Graduate Student Award (Gamma Sigma Delta) during PhD. He is also an University Nominee for the ORAU Ralph E. Powe Junior Faculty Award and recipient of the 2020 University-wide Rookie Faculty Researcher of the Year Award.

## DAY 2 TALKS

WEDNESDAY, OCTOBER 27

**SPEAKER:**

Leila Hashemi-Beni
Assistant Professor of Geomatics and the Director of Geospatial Science and Remote Sensing Laboratory
Department of Built Environment at College of Science and Technology
North Carolina A&T State University

*lhashemibeni@ncat.edu*

**TIME: 3:10-3:45PM**

**TITLE:** From model-based to data-driven remote sensing in environmental management

**ABSTRACT:** The recent proliferation of remote sensing (RS) platforms (e.g., satellites, aircraft, and UAVs) equipped with advanced sensor technologies (e.g., optical, hyperspectral, SAR, LiDAR) has enabled systematic production of massive amounts of high spatial, spectral and temporal data. Extracting and combining the information contained in these rich multisource data enable novel views and creating a comprehensive and detailed knowledge basis of the environmental dynamics for both rapid changing event (i.e. Flood) and steady progressions (e.g., erosion monitoring). This presentation will discuss data driven methods in remote sensing approaches for environmental management. The methods have better performance than model-based methods to exploit spatial and temporal features and dependencies..

**BIO:** Dr. Leila Hashemi-Beni is an assistant professor of Geomatics and the director of Geospatial Science and Remote Sensing Laboratory at the Department of Built Environment at College of Science and Technology, North Carolina A&T State University. Her research experience and interests span the areas of multi temporal and multisource data fusion and image classification, 3D data modeling, UAV and satellite remote sensing and data analytics, automatic matching and change detection between various datasets and developing GIS and remote sensing methodologies for environmental management. She is currently working as a PI/Co-PI on many projects funded by the National Science Foundation (NSF), National Oceanic and Atmospheric Administration (NOAA), North Carolina Collaboratory, North Carolina Department of Transportation and Industry.

**SPEAKER:**

**Seth Blumsack**
Professor of Energy Policy and Economics, and Director, Center for Energy Law and Policy
Pennsylvania State University

*sab51@psu.edu*

**TIME:** **3:45-4:20PM**

**TITLE:** Planning for Spatially Correlated Failures in Coupled Natural Gas and Power Transmission via Stochastic Optimization

**ABSTRACT:** We develop and illustrate a method for the joint planning of natural gas and electric power systems that are subject to spatially correlated failures of the kind that would be expected to occur in the case of extreme weather events. Our approach utilizes a two-stage stochastic planning and operations framework for a jointly planned and operated gas and electric power transmission system. Computational tractability is achieved through convex relaxations of the natural gas flow equations and the use of a machine learning algorithm to reduce the set of possible contingencies. We illustrate the method using a small test system used previously in the literature to evaluate computational performance of joint gas-grid models. We find that planning for geographically correlated failures rather than just random failures reduces the level of unserved energy relative to planning for random (spatially uncorrelated failures). Planning for geographically correlated failures, however, does not eliminate the susceptibility of the joint gas-grid system to spatially uncorrelated failures.

**BIO:** Seth Blumsack is Professor of Energy Policy and Economics and International Affairs in the John and Willie Leone Family Department of Energy and Mineral Engineering and Director of the Center for Energy Law and Policy at The Pennsylvania State University. At Penn State he also serves as a faculty member in the Operations Research program; and is Co-Director of the Penn State Energy and Environmental Economics and Policy Initiative. He also holds a position on the External Faculty of the Santa Fe Institute and is an Adjunct Research Professor with the Carnegie Mellon Electricity Industry Center. He earned a B.A. in Mathematics and Economics from Reed College in 1998, an M.S. in Economics from Carnegie Mellon in 2003, and a Ph.D. in Engineering and Public Policy from Carnegie Mellon in 2006.

# DAY 3 TALKS

## THURSDAY, OCTOBER 28

**SPEAKER:**

**Xuan (Sharon) Di**
Associate Professor
Department of Civil Engineering and Engineering Mechanics
Smart Cities Center, Data Science Institute
Columbia University in the City of New York

*sharon.di@columbia.edu*

**TIME: 12:00-12:35PM**

**TITLE:** Physics-Informed Deep Learning for Traffic State Estimation and Fundamental Diagram Discovery

**ABSTRACT:** Physics-informed deep learning (PIDL), a paradigm hybridizing physics-based models and deep neural networks (DNN), has been booming in science and engineering fields for its robust predictive power and sample-efficient training. However, its application to transportation is still at a nascent stage. In this talk, I will introduce how we apply PIDL to a classical transportation problem – traffic state estimation. I will first demonstrate how to solve partial differential equations of traffic flow models by training DNN regularized by known physics. Then without prior knowledge of fundamental diagrams, a mathematical mapping from traffic density to velocity, I will talk about how PIDL can be leveraged to not only estimate traffic states, but discover the fundamental diagram as a surrogate machine learning model.

**BIO:** Xuan (Sharon) Di is an Associate Professor in the Department of Civil Engineering and Engineering Mechanics at Columbia University in the City of New York and serves on a committee for the Smart Cities Center in the Data Science Institute. Prior to joining Columbia, she was a Postdoctoral Research Fellow at the University of Michigan Transportation Research Institute (UMTRI). She received her Ph.D. degree from the Department of Civil, Environmental, and Geo-Engineering at the University of Minnesota, Twin Cities. Dr. Di received a number of awards including the NSF CAREER, Transportation Data Analytics Contest Winner from Transportation Research Board (TRB), the Dafermos Best Paper Award Honorable Mention from the TRB Network Modeling Committee, Outstanding Presentation Award from INFORMS, and the Best Paper Award and Best Graduate Student Scholarship from North-Central Section Institute of Transportation Engineers (ITE). She also serves as the reviewer for a number of journals, including Transportation Science, Transportation Research Part B/C/D, European Journal of Operational Research, Networks and Spatial Economics, IEEE ITS, and Transportation.

Dr. Di directs the DitecT (Data and innovative technology-driven Transportation) Lab @ Columbia University. Her research lies at the intersection of game theory, dynamic control, and machine learning. She is specialized in emerging transportation systems optimization and control, shared mobility modeling, and data-driven urban mobility analysis. Details about DitecT Lab and Prof. Sharon Di's research can be found in the following link: http://sharondi-columbia.wixsite.com/ditectlab.

**SPEAKER:**

**Fei Fang**
Leonardo Assistant Professor
Institute for Software Research, School of Computer Science
Carnegie Mellon University

*feifang@cmu.edu*

*http://feifang.info/*

**TIME: 12:35-1:10PM**

**TITLE:** Machine Learning and Game Theory for Societal Challenges

**ABSTRACT:** Many real-world challenges we face involve multiple self-interested decision-makers who interact with each other in an environment full of uncertainty. This talk covers our work on machine learning and game theory that can be used to tackle such challenges in security, environmental sustainability, and food security domains. Our work has been deployed by the US Coast Guard for protecting the Staten Island Ferry in New York City since April 2013. Our work has also led to the deployment of PAWS (Protection Assistant for Wildlife Security) in multiple conservation areas around the world, which provides predictive and prescriptive analysis for anti-poaching efforts. In addition, our recent work has been used by 412 Food Rescue, a non-profit volunteer-based food rescue platform to improve their operational efficiency.

**BIO:** Fei Fang is Leonardo Assistant Professor at the Institute for Software Research in the School of Computer Science at Carnegie Mellon University. Before joining CMU, she was a Postdoctoral Fellow at the Center for Research on Computation and Society (CRCS) at Harvard University, hosted by David Parkes and Barbara Grosz. She received her Ph.D. from the Department of Computer Science at the University of Southern California advised by Milind Tambe (now at Harvard). Her research lies in the field of artificial intelligence and multi-agent systems, focusing on integrating machine learning with game theory. Her work has been motivated by and applied to security, sustainability, and mobility domains, contributing to the theme of AI for Social Good. She is the recipient of the IJCAI-21 Computers and Thought Award. She was named to IEEE Intelligent Systems' "AI's 10 to Watch" list for 2020. Her work has won the Best Paper Runner-Up at AAAI'21, Distinguished Paper at IJCAI-ECAI'18, Innovative Application Award at IAAI'16, the Outstanding Paper Award in Computational Sustainability Track at IJCAI'15. She received an NSF CAREER Award in 2021. Her dissertation is selected as the runner-up for IFAAMAS-16 Victor Lesser Distinguished Dissertation Award, and is selected to be the winner of the William F. Ballhaus, Jr. Prize for Excellence in Graduate Engineering Research as well as the Best Dissertation Award in Computer Science at the University of Southern California.

## DAY 3 TALKS

### THURSDAY, OCTOBER 28

**SPEAKER:**

Nicholas Zabaras
Scientific Computing and Artificial Intelligence (SCAI) Laboratory
University of Notre Dame
Notre Dame, IN, USA

*zabaras@gmail.com*

*https://www.zabaras.com/*

**TIME: 1:10-1:45PM**

**TITLE:** Deep Generative Surrogate Modeling and Inversion in Subsurface Flows

**ABSTRACT:** Most attempts at the interface of physical modeling and machine learning (ML) employ computationally-generated data in order to drive the various statistical discovery objectives. This enables a direct transfer and application of ML tools and techniques, but removes valuable structure, symmetries and invariances that were present in the model. In order to rediscover this structure, if at all possible, ML tools would need copious amounts of data. Even when big data is available, it is important to ensure that predictions produced by ML models trained on this data satisfy these constraints. Purely data-based, modern ML tools as those based on deep neural networks (NNs) provide rich representations for learning complex nonlinear functions, but lack robustness and fail when higher-level abstractions implied by the physical structure are needed to make predictions. We will discuss some aspects in the development of multiscale deep learning approaches for surrogate modeling and UQ incorporating known physical constraints as prior knowledge. Examples will be provided in the development of multi-fidelity generative deep learning algorithms for building surrogate and uncertainty quantification models for fluid dynamical systems including subsurface flows. We will use such techniques together with generative and adversarial AI models for the identification of the heterogeneous conductivity field and reconstructing contaminant release history for subsurface remediation.

**BIO:** Prof. Nicholas Zabaras joined Notre Dame in 2016 as the Viola D. Hank Professor of Computational Science and Engineering after serving as Uncertainty Quantification Chair and founding director of the "Warwick Centre for Predictive Modeling (WCPM)" at the University of Warwick. Until 2020 he served as the Director of the interdisciplinary University of Notre Dame "Center for Informatics and Computational Science (CICS)". He was also the Hans Fisher Senior Fellow with the Institute for Advanced Study at the Technical University of Munich where he is currently serving as "TUM Ambassador". Prior to this, he spent 23 years serving in all academic ranks of the faculty at Cornell University. He received his Ph.D. in Theoretical and Applied Mechanics from Cornell, after which he started his academic career at the faculty of the University of Minnesota. Professor Zabaras' research focuses on the integration of Artificial Intelligence and Scientific Computing for the predictive modeling and control of complex systems.

**SPEAKER:**

Noemi Petra
Associate Professor of Applied Mathematics
School of Natural Sciences
University of California, Merced

*npetra@ucmerced.edu*

**TIME:** **1:45-2:20PM**

**TITLE:** Propagation of Uncertainty from Data to Inference for Large-Scale Inverse Problems with Application to Ice Sheet Flow

**ABSTRACT:** Model-based projections of complex systems will play a central role in prediction and decision-making, e.g., anticipate ice sheet contribution to sea level rise. However, models are typically subject to considerable uncertainties. Such uncertainties stem from unknown or uncertain coefficient fields, constitutive laws, source terms, geometries, and initial and/or boundary conditions in the model. While many of these parameters cannot be directly observed, they can be inferred from observations, such as those of ice surface velocities in ice sheets. This typically leads to a severely ill-posed inverse problem whose solution can be extremely challenging. In this talk, we consider the problem of inferring the basal sliding coefficient field for an uncertain Stokes ice sheet forward model from surface velocity measurements via the framework of Bayesian inversion.

Bayesian inversion facilitates the integration of data with complex physics-based models to quantify and reduce uncertainties in model predictions. This opens the door to more advanced capabilities for decision-making under uncertainty. However, the algorithmic developments for Bayesian inversion are subject to several computational challenges. Among those are high dimensionality of the inversion parameters, expensive to evaluate parameter-to-observable maps, and model uncertainty additional to the uncertainty in inversion parameters. In this talk we discuss solution methods that are scalable, exploit problem structures, and are robust with respect to model uncertainty.

**BIO:** Noemi Petra is an Associate Professor of Applied Mathematics in the School of Natural Sciences at the University of California, Merced. She is currently the faculty advisor of the UC Merced SIAM Student Chapter. Noemi earned her B.Sc. degree in Mathematics and Computer Science from Babes-Bolyai University, Romania, and her Ph.D. degree in Applied Mathematics from the University of Maryland, Baltimore County. Prior to joining the University of California, Merced she was the recipient of an ICES (Institute for Computational Engineering & Sciences) Postdoctoral Fellowship at the University of Texas at Austin. During Summers 2015 and 2016, she was a Visiting Faculty in the Mathematics and Computer Science Division at Argonne National Laboratory, funded by the Department of Energy (DOE) Visiting Faculty Program (VFP) and during 2017-2018 she served as the secretary of the SIAM Uncertainty Quantification Activity Group (SIAM UQ). As of 2017, Noemi is a recipient of an NSF CAREER grant award. Her research interests include large-scale inverse problems governed by differential equations, uncertainty quantification in inference and prediction, and optimal experimental design.

# DAY 3 TALKS

## THURSDAY, OCTOBER 28

**SPEAKER:**

**Michael Kapteyn**
Postdoctoral Fellow
Oden Institute for Computational Engineering and Sciences

*mkapteyn@utexas.edu*

**TIME: 2:35-3:10PM**

**TITLE:** A probabilistic graphical model foundation to enable predictive digital twins at scale

**ABSTRACT:** Predictive digital twins have the potential to underpin high-consequence decision-making in complex systems by enabling asset-specific predictive modeling. In this work we propose a probabilistic graphical model as a formal mathematical representation of a digital twin and its associated physical asset. We first define an abstraction of the asset-twin system, viewing it as a coupled pair of dynamical systems evolving over time through their respective state-spaces, and interacting via observed data and control inputs. We formally define the system using a probabilistic graphical model, enabling us to draw upon well-established theory and algorithms from Bayesian inference, dynamical systems, control theory, and scientific machine learning.

We demonstrate an application of the proposed foundation in which we create, calibrate, and deploy a structural digital twin of an unmanned aerial vehicle. The digital twin is imbued with predictive capabilities via physics-based reduced-order models. The probabilistic graphical model is used to dynamically update these models in response to sensor data, and to perform dynamic analysis, prediction, and decision-making, all with quantified uncertainty.

**BIO:** Michael Kapteyn is a postdoctoral fellow in the Oden Institute at UT Austin. He recently completed his PhD in Computational Science and Engineering from MIT, where he was supervised by Prof. Karen Willcox. He holds a Bachelor of Engineering from the University of Auckland, and an SM in Aeronautics and Astronautics from MIT. Michael's research is focused on developing the mathematical and computational foundations necessary to enable predictive digital twins at scale. His digital twin research has been published in Nature Computational Science, recognized by AIAA with best paper awards, and featured by numerous media outlets.

**SPEAKER:**

**Roni Rosenfeld**
Professor and Head, Machine Learning Department
School of Computer Science
Carnegie Mellon University

*roni@cs.cmu.edu*

**TIME:** **3:10-3:45PM**

**TITLE:** Tracking and Forecasting Epidemics

**ABSTRACT:** Epidemiological forecasting is critically needed for decision making by national and local governments, public health officials, healthcare institutions and the general public. The Delphi group at Carnegie Mellon University was founded in 2012 to advance the theory and technological capability of epidemiological forecasting, and to promote its role in decision making, both public and private. Our long term vision is to make epidemiological forecasting as useful and universally accepted as weather forecasting is today. I will describe some of the methods we developed over the past decade for forecasting flu, dengue and other epidemics, and the challenges we faced in adapting these methods to the Covid pandemic.

**BIO:** Roni Rosenfeld (B.Sc., mathematics and physics, 1985, Tel-Aviv University; M.Sc. 1991, Ph.D. 1994, computer science, Carnegie Mellon University) is head of the Machine Learning Department and professor of machine learning, language technologies, computer science and computational biology, in the School of Computer Science at Carnegie Mellon University, Pittsburgh, Pennsylvania. He also holds a courtesy appointment at the Heinz School of Public Policy at Carnegie Mellon, and an adjunct appointment at the University of Pittsburgh School of Medicine.

Rosenfeld has been teaching machine learning and statistical language modeling since 1997. He has taught thousands of undergraduate and graduate students, has been a mentor to four post-doctoral students and an advisor to about a dozen Ph.D. students and a score of Masters and undergraduate students.

Professor Rosenfeld’s current interests include tracking and forecasting epidemics, using speech and language technologies to aid international development, using machine learning for social good, and advancing data numeracy for all. He has also performed research in statistical language modeling, machine learning, speech recognition and viral evolution. He has published well over 100 scientific articles in academic journals and conferences.

Rosenfeld is a recipient of the Allen Newell Medal for Research Excellence and of the Spira Teaching Excellence Award.

http://www.cs.cmu.edu/~roni

# DAY 3 TALKS

## THURSDAY, OCTOBER 28

**SPEAKER:**

**May Wang**
The Wallace H. Coulter Distinguished Faculty Fellow
Professor and Director of Biomedical Big Data Initiative
Georgia Distinguished Cancer Scholar, Kavli Fellow
The Petit Institute Faculty Fellow, FAIMBE, FIAMBE
Board of Directors of American Board of AI in Medicine
Georgia Institute of Technology and Emory University

*maywang@gatech.edu*

**TIME:** **3:45-4:20PM**

**TITLE:** Challenges and Opportunities of Translating AI for Healthcare: Why Many AI Tools Do Not Show Impact in Combating COVID-19?!

**ABSTRACT:** Rapid advancements in biotechnologies such as wearable sensors, –omic (genomics, proteomics, metabolomics, lipidomics etc.), next generation sequencing, bio-nanotechnologies, and molecular imaging etc. accelerate the data explosion in biomedicine and health wellness. Multiple nations around the world have been seeking novel effective ways to make sense of "big data" with AI for evidence-based, outcome-driven, and affordable 5P (Patient-centric, Predictive, Preventive, Personalized, and Precise) health care. The goal is develop multi-modal and multi-scale (i.e. molecular, cellular, whole body, individual, and population) biomedical data analysis with AI for discovery, development, and delivery. Ultimately, the goal is to promote healthy aging, improve quality of patient care, and reduce healthcare cost. However, during the COVID-19 pandemic, many AI tools have been developed, but a few have shown true impact as shown in a recent MIT Technology Review on AI tools impact in combating COVID-19 pandemics. In this talk, I will provide a systematic review regarding the causes, followed by how to harness multi-modality data (e.g. imaging, genomics, physiological, and clinical EMR) by using HL7 Fast Health Informatics Resource (FHIR), and progresses made and opportunities in improving data quality to enable AI translation. Then I will discuss what needs to be done to make multi-modality data analysis outcome usable by clinicians through xAI and intelligent reality.

**BIO:** Dr. Wang is a Wallace H. Coulter Distinguished Faculty Fellow and a full professor in the Departments of Biomedical Eng. and Electrical and Computer Eng. at Georgia Institute of Technology and Emory University. Her research is Biomedical Big Data Analytics with a focus on Biomedical and Health Informatics (BHI) and Artificial Intelligence (AI) for predictive, personalized, and precision health (pHealth). She published over 260 peer-reviewed articles in referred journals and conference proceedings (12,800+ Google Scholar citations), and delivered over 240 invited and keynote lectures. Dr. Wang is the Director of Biomedical Big Data Initiative, a Kavli Fellow, a Georgia Distinguished Cancer Scholar, a Petit Institute Faculty Fellow, an AIMBE Fellow, an IAMBE Fellow, and a member of Board of Directors in American Board of AI in Medicine. She received BEng from Tsinghua University China, and MSCS, and PhD EE degrees from Georgia Institute of Technology. Dr. Wang is a recipient of Georgia Tech Outstanding Faculty Mentor Award for Undergraduate Research, and a recipient of Emory University MilliPub Award (for a high-impact paper that is cited over 1,000 times). Currently, she is Chair of Biomedical and Health Informatics Technical Community, and chair of ACM Special Interest Group in Bioinformatics. She serves as the Senior Editor for IEEE Journal of Biomedical and Health Informatics, an Associate Editor for both IEEE Transactions for BME and IEEE Reviews for BME, a standing panelist for NIH CDMA study section, a multi-year NSF Smart and Connect Health panelist, and a panelist for Brain Canada and multiple European countries. She helped grow the bioinformatics and health informatics technical communities in IEEE EMBS, ACM, and Gordon Research Conferences. In 2021, Dr. Wang is a member of Georgia Tech Provost's Emerging Leader's Program and is elected into IAMBE Executive Committee. During 2018-2020, Dr. Wang was Carol Ann and David Flanagan Distinguished Faculty Fellow at Georgia Tech. During 2017-2019, she served as Vice President of IEEE EMBS and AIMBE Bioinformatics Nomination Committee Chair. During 2015-2018, she was Georgia Tech Biomedical Informatics Program Co-Director in Atlanta Clinical and Translational Science Institute (ACTSI). During 2015-2016, she was IEEE Engineering in Medicine and Biology Society (EMBS) Distinguished Lecturer and an Emerging Area Editor for PNAS. Before 2016, Dr. Wang was Director of Bioinformatics and Biocomputing Core in a large NIH/NCI-sponsored U54 Center for Cancer Nanotechology Excellence, and Co-Director of Georgia-Tech Center of Bio-Imaging Mass Spectrometry for over 10 years. Dr. Wang's research has been supported by NIH, NSF, CDC, Georgia Research Alliance, Georgia Cancer Coalition, Shriners' Hospitals for Children, Children's Health Care of Atlanta, Enduring Heart Foundation, Coulter Foundation, Microsoft Research, HP, UCB, and Amazon.

**SPEAKER:**

**Marcus Noack**
Applied Mathematics
Research Scientist
Berkeley, CA

*MarcusNoack@lbl.gov*

**TIME:** **12:00 - 12:30PM**

**TITLE:** A Perspective on Autonomous Experimentation and Discovery

**ABSTRACT:** Modern instruments have the capability to perform increasingly complex experiments to advance scientific understanding across a range of inquiries, including the properties of materials, scientific mechanisms, biological substances, and engineering designs. As these instruments grow in sophistication, complexities, and capabilities, the measurements and output they produce are becoming increasingly multi-faceted. Stemming in part from technological and engineering advances, such as more spatially and temporarily resolved detectors, faster and larger compute capabilities, and robot-controlled experiments, the resulting output is more voluminous and coming faster than ever before. As scientific questions become more complicated and the underlying parameter space increases in dimensionality, the measurement outcomes depend on an increasing number of user-controllable inputs, such as synthesis, environmental, and processing parameters. Making sense of this requires a systematic way to efficiently ask questions and economically explore this high-dimensional space, probing areas of interest and avoiding time-consuming and costly experiments that may yield relatively little new insight. To this end, we are seeing an explosion in new techniques aimed at autonomous data acquisition. In these ``self-driving laboratories”, advanced mathematical ideas come into play to help steer and design experiments, coordinating the time and sequencing of complex measurements. Enabling and sharing the design of these laboratories of the future will increase efficiency, reduce cost, and greatly accelerate scientific understanding. On April 20-22, 2021, The Department of Energy’s Center for Advanced Mathematics for Energy Research Applications (CAMERA) held an open workshop titled ``Autonomous Discovery in Science and Engineering” to bring the community together and gather insights into this rapidly evolving field.

In this talk, I want to give my perspective on the state of the field of autonomous experimentation and its characterization within Machine Learning and Artificial Intelligence, using the information collected during the CAMERA workshop. Afterward, I will speak to our own efforts to develop a generally applicable tool for autonomous experimentation called gpCAM and show some examples of its use at large-scale experimental facilities.

**BIO:** Marcus Noack got his master’s degree in geophysics from the Friedrich-Schiller University in Jena, Germany. Working as a Ph.D. candidate at Simula Research Laboratory in Oslo, he was able to pursue his interests in the theory of wave propagation and mathematical function optimization. There, he leveraged his knowledge in theoretical and numerical physics and applied mathematics and connected it with High-Performance Computing to create efficient methods to model wave propagation and solve highly non-linear inverse problems. He graduated with a Ph.D. in applied mathematics from the University of Oslo. Starting at Lawrence Berkeley National Laboratory as a Post Doc, he was working on stochastic function approximation and autonomous experimentations; a line of work he is still pursuing now as a Research Scientists.

# DAY 4 TALKS

## FRIDAY, OCTOBER 29

**SPEAKER:**

Sioan Zohar
Brookhaven National Laboratory
Instrumentation Division

*szohar@bnl.gov*

**TIME:** **12:35 - 1:05PM**

**TITLE:** Edge ML Detectors: Integrating Machine Learning with Sensors

**ABSTRACT:** The Instrumentation Division (ID) at Brookhaven National Laboratory (BNL) develops and delivers cutting edge sensors that advance detection limits at scientific research facilities around the globe. As the recent data explosion pushes data center and cloud-based tech stacks close to their limits, new methodologies that migrate machine learning (ML) based feature extraction and event rejection closer to the sensor have shown great promise in reducing power consumption and data rates. In this talk, we review the ID's development in memristor based computing and progress with integrating machine learning directly onto readout ASICs bonded to the sensor substrate. In addition to the anticipated performance gains and data reduction, this approach offers several unique opportunities. From a deployment and operations perspective, sensors require intermittent recalibration. Packaging the sensor together with the ML algorithm offers an opportunity to autonomously identify and correct detector miscalibration. From a fundamental compute perspective, packaging ML directly with the detector provides the opportunity to engineer nanostructured sensors that exploit naturally occurring non-linearities as feature extractors.

**BIO:** Dr. Sioan Zohar graduated in 2010 from Columbia University with his PhD in applied physics where he investigated the high frequency dynamics of spintronic heterostructures. He continued this work during his postdoctoral research at Argonne National Laboratory's Advanced Photon Source (APS), where he developed novel x-ray imaging techniques. Following his postdoc, Dr. Zohar worked as a beam line physicist at the APS and Linac Coherent Light Source where he developed FPGA based beam stabilization and signal processing systems. Currently, Dr. Zohar works at Brookhaven National Laboratory where he researches machine learning models for deployment in FPGAs, ASICs, and reservoir systems.

**SPEAKER:**

Sandeep Miryala
Brookhaven National Laboratory
Instrumentation Division

*smiryala@bnl.gov*

**TIME: 1:10 - 1:40PM**

**TITLE:** Co-Design Methodologies for Edge Computing ASICs in Scientific Research Environments

**ABSTRACT:** We present our recent progress applying co-design methodologies on read-out architectures for optimized data-stream handling of edge-based Application Specific Integrated Circuits (ASICs). To date, the majority of ML based scientific data processing runs on GPUs, supercomputers, or dedicated AI accelerators. In this paradigm, FPGA and ASIC readout systems are agnostic to the end user's data requirements, and domain expertise is not considered in their design. Here, we present our progress using High Level Synthesis (HLS) tools to translate python based neural networks into hardware description languages. Initial RTL simulations show inference accuracy consistent with Tensorflow models. Such networks are typically too large for ASIC specific implementation. A potential solution using memristor crossbar arrays for in-memory computing is explored.

**BIO:** Sandeep Miryala was born in Hyderabad, India, in 1986. He received the Bachelor of Technology (B. Tech) degree in electronics and communications engineering from the Jawaharlal Nehru Technological University (JNTU), Hyderabad, India, in 2007, the Masters of Technology (M.Tech) degree in Semiconductor Devices and VLSI Technology from the Indian Institute of Technology (IIT), Roorkee, India, in 2010, and the Ph.D. in Informations and systems engineering from Politecnico Di Torino, Torino, Italy, in 2014.

From 2014-16, he joined the Electrical Engineering Department, Nikhef, Amsterdam, Netherlands, as a Technical Scientist designing digital integrated circuits for Vertex Locater chip of the CERN, LHCb experiment. From 2016-2019, he was an ASIC Design Engineer in Electrical Engineering Department of Fermilab and has contributed to the integrated circuit developments for inner tracker, high granular calorimeter, and endcap timing detectors of CERN, CMS experiment. He joined BNL as an ASIC design Engineer in 2019 and has been contributing to the DUNE far end SP-TPC detector developments namely COLDADC and COLDATA chips. These chips are selected for the production run and the final board level testing is ongoing.

He is a senior IEEE member and has been serving as a track chair and Technical Program Committee (TPC) member for major IEEE conferences and workshops (ICECS, ISCAS, ISVLSI, ICCD etc.). His current research interests include digital readout architectures, digital design for extreme environments (radiation and cryogenic), edge computing ASICs, Non-Volatile Memories.

## DAY 4 TALKS

### FRIDAY, OCTOBER 29

**SPEAKER:**

**Anže Slosar**
Group Leader for Cosmology & Astrophysics group
Brookhaven National Laboratory

*anze@bnl.gov*

**TIME: 1:45 - 2:15PM**

**TITLE:** ML/AI applications in astronomy

**ABSTRACT:** Optical astronomy takes images of the sky and so it looks like a natural space in which to employ the rich arsenal of ML/AI image analysis techniques that were developed for photographic image analysis over the past decade. However, both the basic properties of the images and the quantities that we are trying to extract are very different from the canonical problems of the traditional image analysis. I will discuss similarities and differences between astronomical images and ordinary photographs and present some solutions that are being developed to bridge the two worlds.

**BIO:** Dr. Anže Slosar graduated from Cambridge, followed by postdocs in Ljubljana (Slovenia), Oxford and an Inaugural Fellowship at Berkeley Center for Cosmological Physics before moving to a staff Scientist position at Brookhaven National Laboratory. He is a group leader for Cosmology & Astrophysics Group. He is interested in all experimental probes of cosmology and has worked on numerous topics including Lyman-alpha forest, galaxy clustering, primordial non-Gaussianity and recently on analysis of photometric galaxy surveys.

**SPEAKER:**

Georg Schnabel
Department of Nuclear Sciences and Applications
International Atomic Energy Agency

*G.Schnabel@iaea.org*


**TIME:** **2:35 - 3:05PM**

**TITLE:** Nuclear data evaluation with Bayesian networks


**ABSTRACT:** The advancement of nuclear technology, nuclear physics and related fields, such as nuclear astrophysics depends on precise and comprehensive knowledge of the properties of the atomic nuclei, such as cross sections and half-lives, which are referred to as nuclear data. Often the generation of comprehensive nuclear data relies not only on data from various experiments, e.g., at particle accelerators and nuclear reactors, but also on nuclear models calibrated to experimental data using Bayesian inference. Due to the complex nature of experiments, many potential error sources must be considered. Also nuclear models are often not perfectly able to reproduce the experimental data because they neglect physical processes or rely on approximations. Considering the many possible ways in which the pieces of information of various experiments may relate to each other and to predictions of nuclear models, a flexible modelling framework for estimation and uncertainty quantification is tremendously useful to accelerate the production of high-quality nuclear data. Bayesian networks are graphical models that provide such a flexible framework. This talk discusses the application of Bayesian networks in the context of nuclear data evaluation with examples. On the technical level, it covers the integration of a specific sparse Gaussian process construction into the Bayesian network framework and inference within a system of not precisely known functions linked


**BIO:** Georg Schnabel studied physics at the Technical University of Vienna where he received his doctorate in 2015. Subsequently he worked as a research scholar at CEA Saclay in France and Uppsala University in Sweden on uncertainty quantification methods for nuclear models and data before he joined in 2020 the Nuclear Data Section at the IAEA where his responsibilities include the coordination of nuclear data library projects and the development of scientific codes for data processing and statistical analysis.

# DAY 4 TALKS

## FRIDAY, OCTOBER 29

**SPEAKER:**

**Jonathan Edelen**
Senior Research Scientist, Group Leader
RadiaSoft

*jedelen@radiasoft.net*

**TIME: 3:10 - 3:40PM**

**TITLE:** Automating Particle Accelerator Operations with Machine Learning

**ABSTRACT:** In recent years machine learning has been identified as having the potential for significant impact on the modeling, operation, and control of particle accelerators. Machine learning and neural networks are attractive for this application due to their ability to model nonlinear behavior, interpolate on complicated surfaces, and adapt to system changes over time. The confluence of advances in machine learning coupled with easier access to compute resources and ML tools has led to a number of concentrated efforts at national laboratories and industry all aimed at improving the way we operate both large and small-scale facilities. This talk will take a journey through some of the recent efforts to bring ML into the control room. We will begin with an overview of particle accelerators and highlight the most promising applications of machine learning in this space. We will then provide detailed examples of how machine learning is being applied as 1: a means to speed up optimization with accelerator simulations 2: identify tuning errors and predict faults using anomaly detection and 3) for improving the automation of tuning tasks for user facilities.

**BIO:** Jonathan Edelen is an accelerator physicist with a broad range of experience across the field. He specializes in thermionic cathode guns and space-charge effects in low energy electron beams, RF control systems, and machine learning for anomaly detection. Jon earned a PhD in accelerator physics from Colorado State University. After completing his PhD, Jon was selected for the prestigious Bardeen Fellowship at Fermilab. While at Fermilab he worked on RF systems for the PIP-II Injector Test, beam dynamics and RF transient effects in the proposed PIP-II accelerator, and thermionic cathode RF guns at the Advanced Photon Source. Currently, Jon is focused on building advanced control algorithms for particle accelerators including solutions involving machine learning. He is also developing improved simulation tools for studying field enhanced thermionic emission in thermionic energy converters. Jonathan has published papers and presented at international conferences on a variety of topics in accelerators.

**SPEAKER:**

**Phillip Maffetone**
Assistant Computational Scientist, Data Science and Systems Integration Program
National Synchrotron Light Source II

*pmaffetto@bnl.gov*

**TIME:** **3:45 - 4:15PM**

**TITLE:** Remote and on-the-fly: artificial intelligence driven science in laboratories and central facilities

**ABSTRACT:** Scientific experiments at university laboratories, central facilities, and industrial settings are increasingly demanding of remote, high-throughput, and adaptive operation conditions. To accommodate such needs, new approaches must be developed that enable on-the-fly decision making for data intensive challenges, and automated solutions for collecting data.

This talk will outline a suite of advancements in autonomous experimentation for a diverse range of scientific problems. The discussion will span mobile robotics, Bayesian optimization and reinforcement learning for experiment planning, and deep and statistical learning for on-the-fly analysis and visualization of large datasets. At the core of this body of work is extended collaboration with domain experts and the leveraging of scalable, open-source infrastructure, including the Bluesky project. Rather than attempt to develop a one-size-fits-all solution for every experiment, the presentation will outline the pragmatic rationale for a federation of agents interfaced with streaming data.

**BIO:** Dr. Phil Maffettone is an Assistant Computational Scientist in the Data Science and Systems Integration (DSSI) program at the National Synchrotron Light Source II (NSLS-II). He previously developed the brain on the world's first mobile robotic scientist. His current research is focused on accessible and scalable autonomous experimentation around NSLS-II and beyond.

# NYSDS 2021 CONTRIBUTED POSTERS

**P1: Integrating Physical and Machine Learning Modelling to Improve Renewable Energy Forecast**

*Yangang Liu[a], Tao Zhang[a], Yunpeng Shan[a], Xin Zhou[a], Shinjae Yoo[a], Chenxiao Xu[b]*
*[a]Brookhaven National Laboratory, [b]Stony Brook University*

**P2: Combining Reinforcement Learning with Physics-Based Model Predictive Control for Nutrient Removal in Water Resource Recovery**

*Lav R. Varshney[a], Hersh V. Kshetry[a], and Nina Kshetry[a]*
*[a]Ensaras, Inc.*

**P3: Machine Learning to evaluate the information content of satellite data for fine particle estimates in India: a modeling testbed**

*Zhonghua Zheng[a], Arlene M. Fiore[a,b], Daniel M. Westervelt[a,c], George P. Milly[a], Jeff Goldsmith[d], Ruth S. DeFries[e], Alexandra N. Karambelas[a], and Gabriele Curci[f], Cynthia A. Randles[g], Antonio R. Paiva[g]*
*[a]Lamont-Doherty Earth Observatory, Columbia University, [b]Department of Earth and Environmental Sciences, Columbia University, [c]NASA Goddard Institute for Space Studies, [d]Department of Biostatistics, Columbia University, [e]Department of Ecology, Evolution, and Environmental Biology, Columbia University, [f]Department of Physical and Chemical Sciences, University of L'Aquila, [g]ExxonMobil Research and Engineering Company*

**P4: Wind and temperature measurements around a supertall building**

*Katia Lamer[a], Edward P. Lukea, Zackary Mages[b], Erin Leghart[b], Zeen Zhub and Andrew M. Vogelmann[a]*
*[a]Brookhaven National Laboratory, [b]Stony Brook University*

**P5: Modularity and Integration in Energy Grids: Data-Driven Definitions, Benefits, and Drawbacks of Energysheds**

*Lav R. Varshney[a,b], Le Xie[c], and Sanjoy K. Mitter[d]*
*[a]Brookhaven National Laboratory, [b]University of Illinois Urbana-Champaign, [c]Texas A&M University, [d]Massachusetts Institute of Technology*

**P6: Anomaly detection algorithms for an in-situ performance analysis framework**

*Julianne Starzee [a], Arnav Agrawal [a], Sandeep Mittal [b], Christopher Kelly [b], Shinjae Yoo [b], Wei Xu [b], Huub Van Dam [b], Line Pouchard [b], Kerstin Kleese Van Dam [b]*
*[a]Cornell University (Physics), [b]Brookhaven National Laboratory (CSI)*

**P7: Optimizing neural networks for next-gen AI**

*Andrew Deutsch[a], Yihui Ren[b], Sandeep Mittal[c]*
*[a]Department of Physics, Stevens Institute of Technology, [b]Mentor and [c]Co-mentor, Computational Science Initiative, Brookhaven National Lab*

**P8: Evaluation of PDF Parsing Tools for the Scientific Literature**

*Darrien Hunt [a] , Gilchan Park b, Line Pouchard [b]*
*[a]Computer Science, Hampton University, Hampton, VA 23669, [b]Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973*

**P9: Automated generation of particle pickers from solution scattering data**

*Thomas Flynn, Yuwei Lin, Qun Liu, Liguo Wang, Sean McSweeney*
*Brookhaven National Laboratory, Upton, NY*

**P10: Methods for Uncertainty Quantification for Deep Learning Digital Twins of Free Electron Laser Scientific Facilities**

*Lipi Gupta [a], Aashwin Mishra [b], Auralee Edelen [b]*
*[a] NERSC Lawrence Berkeley National Laboratory, [b] SLAC National Accelerator Laboratory*

**P11: Using Machine Learning for spin-group reclassification of neutron resonances**

*G.P.A. Nobre[a], D.A. Brown[a], S. Hollickb, S. Scoville[c], P. Fernandez[d], M. Fucci[e], R.-M. Crawford[f], S. Ruiz[g], A. Coles[a], M. Vorabbi[a]*
*[a]National Nuclear Data Center, Brookhaven National Laboratory, [b]Yale University Physics Department, [c]Rensselaer Polytechnic Institute Department of Physics, [d]University of Puerto Rico Mayagüez Campus, [e]University At Albany, Department of Physics, [f]Willamette University, Department of Physics, [g]Georgia Institute of Technology, School of Physics*

**P12: Data-driven Chaos Indicator for Nonlinear Dynamics and Applications on Storage Ring Lattice Design**

*Yongjun Li[a], Jinyu Wan[b], Allen Liu[c], Yi Jiao[b], Robert Rainer[a]*
*[a]Brookhaven National Laboratory, [b]Institute of High Energy Physics, [c]Purdue University*

**P13: Deep neural network methods for partial differential equations**

*Jiawei Sun[a], Vanessa López-Marrero[b], Nathan Urban[b]*
*[a]The Ohio State University, [b]Brookhaven National Laboratory*

**P14: Solving Inverse Problems with Physics-Informed Neural Networks**

*Maati McKinney[a], Vanessa López-Marrero[b]*
*[a]Harvard University, [b]Brookhaven National Laboratory*

**P15: A Physics-Informed Deep Learning Paradigm for Car-Following Models**

*Zhaobin Mo[a], Rongye Shi[b], Xuan Di[a,b]*
*[a]Department of Civil Engineering and Engineering Mechanics, Columbia Univerisity, [b]Data Science Institute, Columbia University*

**P16: Applying Bayesian Optimization to Achieve Optimum Cooling at the Low Energy RHIC Electron Cooling System**

*Yuan Gao[a], Weijian Lin[b], Kevin Brown[a,c], Xiaofeng Gu[a], Georg Hoffstaetter[a,b,d], John Morrisa, Sergei Seletskiy[a]*
*[a]Collider-Accelerator Department, Brookhaven National Laboratory, Upton, NY 11973, [b]Cornell Laboratory for Accelerator Based Sciences and Education, Cornell University, Ithaca, NY 14853, [c]Electrical and Computer Engineering Department, Stony Brook University, Stony Brook, NY 11794, [d]Electron-Ion Collider Department, Brookhaven National Laboratory, Upton, NY 11973*

**P17: Robust importance sampling for error estimation in the context of optimal Bayesian transfer learning**

*Omar Maddouri [1], Xiaoning Qian [1,2], Francis J. Alexander [2], Edward R. Dougherty [1], Byung-Jun Yoon [1,2]*
*[1]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA, [2]Computational Science Initiative, Brookhaven National Laboratory, Upton, NY 11973, USA*

**NYSDS 2021**
CONTRIBUTED POSTERS

### P18: Performance Optimization of High-Throughput Virtual Screening Pipelines

*Hyun-Myung Wooa, Xiaoning Qian[a,b], Li Tanb, Shantenu Jha[b,c], Francis J. Alexanderb, Edward R. Dougherty[a], Byung-Jun Yoon[a,b]*

*[a]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, [b]Computational Science Initiative, Brookhaven National laboratory, Upton, NY 11973, [c]Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ 08854*

### P19: TRIMER: Transcription Regulation Integrated with MEtabolic Regulation

*Puhua Niu[a], Maria J. Soto[b], Byung-Jun Yoon[a,c], Edward R. Dougherty[a], Francis J. Alexander[c], Ian Blabybd, Xiaoning Qian[a,c]*

*[a] Texas A&M University, College Station, TX, USA, [b] 2US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [c] Brookhaven National Laboratory, Upton, NY, USA, [d]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

### P20: scTenifoldKnk: an efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation

*Daniel Osorio[1], Yan Zhong[2], Guanxun Li[2], Qian Xu[1], Yongjian Yang[3], Yanan Tian[4], Robert S. Chapkin[5], Jianhua Z. Huang[2], James J. Cai[1,3,6],**

*[1]Department of Veterinary Integrative Biosciences, [2]Department of Statistics, [3]Department of Electrical and Computer Engineering, [4]Department of Veterinary Physiology and Pharmacology, [5]Department of Nutrition, [6]Interdisciplinary Program of Genetics, Texas A&M University, College Station, TX 77843, USA*

### P21: Adaptive Group Testing with Mismatched Models

*Mingzhou Fan [a], Byung-Jun Yoon [a,b], Francis J. Alexander [b], Edward R. Dougherty [a], Xiaoning Qian [a,b]*

*[a] Department of Electrical & Computer Engineering, Texas A&M University, [b] Computational Science Initiative, Brookhaven National Laboratory*

### P22: Reliability-Time-Aware Imputation for In-hospital Mortality Prediction

*Tao Zhang[a], Xin Daia, Yuewei Lin[a], Nicholas D'Imperio[a], Shinjae Yoo[a]*

*[a]Brookhaven National Laboratory*

### P23: A Case Study of the Mortality Prediction on MIMIC-II Dataset

*Yuewei Lin[a], Xin Dai[a], Tao Zhang[a], Nicholas D'Imperio[a], Shinjae Yoo[a]*

*[a]Brookhaven National Laboratory*