NJIT

INSTITUTE FOR DATA SCIENCE

DATA SCIENCE SUMMIT

ABOUT THE SUMMIT

The NJIT and Data Science community is invited to join the Institute for Data Science in its inaugural open summit with key industry professionals and NJIT research faculty with topics on Data Science and its implementation.

ON TUESDAY, 8 NOVEMBER 2022

101 HUDSON ST. IN JERSEY CITY

START AT 1:30 PM - 6 PM

GUILING "GRACE" WANG
Distinguished Professor, NJIT

REZA CURTMOLA
Professor, NJIT

SENJUTI BASU ROY
Associate Professor, NJIT

KATHY MEIER-HELLSTERN
Principal Engineer, Google Research

MANOJ KUMAR
Program Director, IBM

SHERRY MARCUS
Director, Data Science, Amazon

https://datascience.njit.edu I email: info.datascience@njit.edu I RSVP

Data Science Summit

Securing Software Supply Chains with in-toto

Reza Curtmola, Ph.D.
Professor
The New Jersey Institute of Technology

The security of software supply chains is a topic that has been largely overlooked over the past few years, despite numerous recent incidents which show that attacks can happen at any point in this chain, including the most recent one involving SolarWinds. We have developed in-toto, a novel framework that provides insights about processes that occurred in the various steps of the software supply chain. in-toto is the first security mechanism that protects software from the point when the developer commits the code until the end user installs it. in-toto is now a Linux Foundation project and has been deployed into several real-world open source and commercial systems.

Dr. Reza Curtmola is a Professor in the Department of Computer Science at NJIT and the Director of the NJIT Cybersecurity Research Center. He received his Ph.D. degree in Computer Science in 2007 from The Johns Hopkins University and spent one year as a postdoctoral research associate at Purdue University. His research focuses on web security and privacy, security of the software supply chain, security of cloud services, and applied cryptography. He is the recipient of the NSF CAREER award and has participated in several other projects funded by NSF and DARPA. Dr. Curtmola has published over 70 scientific articles under the umbrella of cybersecurity.

Data Science Summit

Optimization Opportunities in Human-in-the-loop AI

Senjuti Basu Roy, Ph.D.
Associate Professor
The New Jersey Institute of Technology

An emerging trend is to leverage an under-explored and richly heterogeneous pool of human knowledge inside machine algorithms, a practice popularly termed as human-in-the-loop (HIL) AI. A wide variety of applications, starting from query processing to text translation, feature engineering, or even human decision making in complex uncertain environments stand to benefit from such synergistic human-AI collaboration. This talk will discuss our ongoing projects, recent research results, and impacts that investigate a variety of optimization opportunities considering the roles and responsibilities of three key stakeholders inside HIL AI processes - humans (workers), machines (AI algorithms), and platforms (online infrastructure where the work takes place). Following that, the talk will briefly discuss how this ongoing research is well aligned in the context of the future-of-work.

Senjuti Basu Roy is the Panasonic Chair in Sustainability and an Associate Professor in the Department of Computer Science at the New Jersey Institute of Technology. Her research focus lies on the intersection of data management, data exploration, and AI, especially enabling human-AI analytics in scale. Senjuti has published more than 75 research papers in high impact data management and data mining conferences and journals. She is the tutorial co-chair of VLDB 2023, The Web Conference 2022, has served as the Mentorship co-chair of SIGMOD 2018, PhD workshop co-chair of VLDB 2018, and has been involved in organizing several international workshops and meetings. She is a recipient of the NSF CAREER Award, a PECASE nominee, and one of the 100 invited early career engineers to attend the National Academy of Engineering's 2021 US Frontiers of Engineering Symposium.

Data Science Summit

Responsible AI, from Principles to Practice

Kathy Meier-Hellstern, PhD
Principal Engineer and Director
Google Research

Google's AI Principles, which were released in 2018, are key to ensuring the way we develop artificial intelligence is aligned with Google's mission and core values, and in the best interest of humanity. However, significant work is required to interpret the principles and apply them. This talk describes Google's technical approach to interpret the principles for the Large Models we are developing in Research.

Kathy is a Principal Engineer and Director in Google Research, serving as the Responsible AI Tech Lead for Google's large language and multimodal models. Her research mission is to simplify RAI adoption, fueled by research breakthroughs and state-of-the-art technology. Kathy was previously a Principal Site Reliability Engineer at Google, focused on improving the end-to-end client experience in YouTube and Ads. Before joining Google, Kathy was Assistant Vice President of Optimization, Reliability & Customer Analytics in AT&T Labs, responsible for delivering enhanced analytic tools and software for AT&T's Next Generation networks. Kathy is an AT&T Fellow, and holds a Ph.D. and Master's degree in Operations Research from the University of Delaware.

Data Science Summit

Data Science and Computing Environment of the Future

Manoj Kumar, Ph.D.
Program Director
IBM T.J. Watson Research Center

Data Science is becoming the new driver of HPC workloads. However, its computational characteristics are very different than those of the conventional drivers of HPC, the scientific and engineering workloads. Memory access patterns, memory size, floating point precision required, etc., are some examples of the workload differences. Furthermore, the slow-down of Moor's Law, end of Dennard scaling to be more accurate, has ushered in rapid advances in 3-D packaging technologies to address the computing requirements of data-science. Finally, the cloud-based delivery of computing cycles, provided by heterogeneous data-science-optimized ASICs, coupled with the security requirements for the training data and the model parameters for data-science applications, reshape security-functionality provided in computing hardware.

In this talk, we will present our view on how HPC systems are likely to evolve to address the computing requirements of data-science, leveraging the innovations in packaging and semiconductor technology, incorporating data-sciences optimized hardware made possible by much improved design and fabrication logistics for ASICS, while providing the desired levels of data privacy and program protection.

Manoj leads the research program to develop hardware and algorithms for emerging analytics applications and secure computing requirements. Large graph analytics is a key part of the analytics application focus. Acceleration of Post Quantum Cryptography (PQC), Elliptic Curve Arithmetic, and cryptographic hash algorithms being the focus of secure computing research. Dr. Kumar has served in administrative roles as director of IBM's India Research Laboratory in New Delhi from 2000 to 2003, and as director of the Corporate Technology Evaluation department 2003-2005, coordinating the studies presented to the CEO on the major hardware, software, and services technology activities of the corporation.

Manoj received his B. Tech from IIT Kanpur in 1979, and M.S. and Ph.D. degrees from Rice University, Houston, Texas, USA, in 1981 and 1984 respectively, all in electrical engineering.