PROGRAM SCHEDULE

Saturday August 24

Time	Event
7:30 - 8:30 AM	**Registration & Check-in**
8:30 - 11:20 AM	**Morning Keynotes**
8:30 AM	Welcome, **Sheree Wen**, New Era AI Chair
8:35 AM	Keynote/welcome, **Tom Coughlin**, IEEE 2024 President
9:00 AM	Welcome, **Tim Lee**, IEEE-USA 2025 President-elect
9:05 AM	Grand Keynote, **Ramesh Menon**, Chief Artificial Intelligence Officer, US DIA
9:30 AM	Grand Keynote, **Dr. Peter Lee**, President, Microsoft Research
9:55 AM	Grand Keynote, **Dr. Bill Dally**, Chief Scientist, Nvidia
10:20- 10:35 AM	**Break (15 min)**
10:35 AM	Grand Keynote, **Dr. Blaise Agüera y Arcas**, CTO of Technology & Society, Google
11:00 AM	Grand Keynote, **Honorable Senator Maria Cantwell**
11:20 AM - 12:00 PM	**Panel - Perpetual and Boundless AI**
	Honorable Senator Maria Cantwell, leader
	Dr. Peter Lee, President, Microsoft Research, panelist
	Dr. Kristin Lauter, Sr. Director FAIR Labs, Meta, panelist
	Dr. Chirag Shah, Co-Founder RAISE, UW, panelist
	Dr. Blaise Agüera y Arcas, CTO of Technology & Society, Google, panelist
	Dr. Sheree Wen, President,UNESCO Washington, moderator
12:00 - 1:00 PM	**Lunch & Research Breakthrough Showcase**
1:00 - 4:15 PM	**Afternoon Keynotes**
1:00 PM	**Dr. Chirag Shah**, Co-Founder RAISE, UW
1:30 PM	**Dr. Stefano Soatto**, VP, AWS AI Labs and Professor, UCLA
2:00 PM	**Dr. Kristin Lauter**, Sr. Director FAIR Labs, Meta
2:30 PM	**Dr. Juan Lavista Ferres**, CVP, Chief Data Scientist, AI For Good Lab
3:00 - 3:15 PM	**Break (15 min)**
3:15 PM	**Prof. Steve Brunton**, NSF AI Institute in Dynamic Systems
3:45 PM	**Prof. Blake Hannaford**, ECE, ME and Surgery, UW
4:15 - 4:30 PM	**Break (15 min)**
4:30 - 6:00 PM	**Lightning Talks**
6:00 – 7:30 PM	**After Party Networking Reception & Lightning Talk Awards**

12:00 – 1:00 PM Research Breakthrough Showcase

Topic	Presenter
Building AI solutions that fit the jigsaw puzzle	**Sourabh Kulhare**, Global Health Labs
AI and Accessibility: Risks and Opportunities	**Dr. Jennifer Mankoff**, Paul G. Allen School of CS
Real-to-Sim-to-Real: A Scalable Data Diet for Robot Learning	**Dr. Abhishek Gupta**, Paul G. Allen School of CS

4:30 – 6:00 PM Lightning Talks

Lightning Talk	Presenter
Data Quality in Intelligent Systems	**Sreyashi Das**, Senior Data Engineer at Netflix
PartyRock: Unleashing the Power of Generative AI in a Playground	**Karthikeyan Gopal**, Senior Software Development Engineer, Amazon
Building Trustworthy AI – Guardrails for a Responsible Future	**Ashish Vaidya**, Principal Software Engineer, Amazon
Artificial Intelligence for Sales and Marketing	**Harshita Asnani**, Applied Data Scientist, Amazon
Revolutionizing Organizational Security with Microsoft Copilot's AI Capabilities	**Ankit Masrani**, Principal Software Engineer, Microsoft Security
SLIs for ML Models	**Sunandan Barman**, Production Engineer, Meta Platforms
Speeding Up Product Innovation: Leveraging Generative AI for Rapid Product Development and GTM Strategies	**Ram Joshi**, Staff Engineer at The Climate Corporation, and **Gunjan Paliwal**, Sr. Product Development & Marketing Manager
Bridging the Chasm: How Software Innovation is Outpacing Hardware in the AI Revolution	**Dr. Harisankar Sadasivan**, Member of Technical Staff SDE, AMD
Building Product Strategy for Modern Times	**Seema Bansal**, Product Manager, Microsoft

SPONSORS

EYNOTE SPEAKERS

Morning Keynotes

Welcome, **Sheree Wen**, New Era AI Chair

Keynote/welcome, **Tom Coughlin**, IEEE 2024 President

Welcome, **Tim Lee**, IEEE-USA 2025 President-elect

Grand Keynote, **Ramesh Menon**, Chief Artificial Intelligence Officer, US DIA

Grand Keynote, **Dr. Peter Lee**, President, Microsoft Research

Grand Keynote, **Dr. Bill Dally**, Chief Scientist, Nvidia

Grand Keynote, **Dr. Blaise Agüera y Arcas**, CTO of Technology & Society, Google

Grand Keynote, Honorable Senator Maria Cantwell

Dr. Sheree Wen Organizer and Chair, New Era AI

Bio:

After getting her Ph.D. from UC Berkeley, **Sheree Wen** worked at IBM, making vital contributions to the first MIPS High-speed mainframe computer. Dr. Wen then built companies with factories in the United States and Asian countries which developed, manufactured, and distributed OEM brands of IBM, AT&T, Unisys, and Wen products worldwide. Her companies also serviced and supplied products to Federal, State, and local governments. Sheree is also deeply committed to the public good. She served as a United States National Commissioner to UNESCO and is the Founder and President of UNESCO Washington State. Sheree has chaired and served on various commissions, committees, and councils in state, regional and local governments. She has worked closely with IEEE, serving as the chair of the Humanitarian Partnership Committee, the Communication and Cybersecurity Policy Committee, and Region 6 Government Relations and Women in Engineering. Currently, Sheree serves on various national and local boards. Sheree was honored as the young scientist of the year and Gold medal by AIME, Award of Excellence by United States Small Business Administrators, top women entrepreneurs by The New York Times, and featured in 20th Century Asian Professional Women by Japan Asahi News.

Thomas Coughlin, FIEEE, IEEE 2024 President

Topic: Keynote/welcome

Bio:

Tom Coughlin is the President of Coughlin Associates, Inc., San Jose, California (www.storagevisions.com). He is an authority on data storage products including storage devices, network attached storage, storage area networks, and storage applications. He does technical consulting and writes market and technology

reports on digital storage and its applications. Tom is the author of Digital Storage in Consumer Electronics: The Essential Guide published in March 2008 by Newnes press (a division of Elsevier). Dr. Coughlin held senior engineering and management positions at several digital storage component and systems companies. His responsibilities include technical evaluations, engineering consulting, project management, market and technology reports, technical article writing, expert witness work and conference organizing. He is the founder and organizer of the annual Storage Visions Conference (www.storagevisions.com) and the Creative Storage Conference (www.creativestorage.org).

Tom is also a fellow of the IEEE, Member of IEEE-HKN, Chair of the 2011 IEEE Sections Congress, past director of IEEE Region 6 and Past-President of IEEE-USA. He is one of the founding members of the SNIA Solid State Storage Initiative (now part of the Compute, Memory and Storage Initiative), an active member of SMPTE and has been active in IDEMA.

Tim Lee, IEEE-USA 2025 President-elect

Topic: Welcome

Bio:

Tim Lee is currently the 2024 IEEE USA President Elect. He is the past IEEE Region 6 (Western US) Director and member of IEEE Board of Directors (2021-2022). He is a Boeing Technical Fellow in advanced microelectronics and packaging in Los Angeles, California. Tim's interests in IEEE activities that promotes the use of technology that benefit humanity. Currently, Tim is Vice-Chair of the IEEE Future Networks Technical Community. He is the Co-Chair of the IEEE Heterogenous Integration Roadmap TWGs for mm-waves and Aerospace & Defense. He was the IMS2020 General Chair. Previously he was the Chair of the IEEE Internet Inclusion track for the IEEE Internet Initiative, the Chair for IEEE SIGHT and the Chair of the IEEE Humanitarian Activities Committee Projects.

In 2023, Tim's priorities for the IEEE include leveraging the US CHIPS and Science Act of 2022 programs to promote semiconductor workforce development, revitalization of IEEE membership in the US, industrial engagement, and the support of Diversity, Equity and Inclusion. He is leading the expansion of the IEEE MOVE program for emergency communications for disaster relief to Western US regions. To volunteer for MOVE, visit https://move.ieeeusa.org/ or contact me.

NEW ERA AI SUMMIT PROGRAM

Ramesh Menon, Chief Artificial Intelligence Officer, US DIA

Topic: Emerging Technologies & National Security

Abstract:

Learn about some of the emerging technology trends in national security space and you can help the US Government.

Bio:

Ramesh Menon is a Senior Executive at US Department of Defense leading next generation capabilities, talent and cultural transformation working in collaboration with our global allies and partners.

His specialties include: Thought leadership, Management Consulting, Cyber Security, Cloud Computing, Artificial Intelligence, Quantum Technologies, Digital Transformation and Organizational Design.

Dr. Peter Lee, President, Microsoft Research

Topic: The Emergence of General AI for Medicine

Abstract:

Large language models such as GPT-4 have emerged as powerful new tools for information work, including aspects of healthcare delivery, healthcare management, and advancing medical science. GPT-4, in particular, shows impressive capabilities that go beyond direct medical knowledge, to include even support for patient engagement and medical research. This talk presents highlights of an intensive, year-long study exploring the benefits, limitations, and risks of GPT-4 in medicine. Our fundamental conclusion is that GPT-4 may represent the most important advance in computing technology, with potentially transformative impact on healthcare, and despite having received no specialized training in medicine. Examples will be shown of how the general intelligence of GPT-4 can be used, with implications for the current and future practice of medicine.

Bio:

Dr. Peter Lee is President, Microsoft Research. He leads Microsoft Research and incubates new research-powered products and lines of business in areas such as artificial intelligence, computing foundations, health, and life sciences. Before joining Microsoft in 2010, he was at DARPA, where he established a new technology office that created operational capabilities in machine learning, data science, and computational social science. Prior to that, he was a professor and the head of the computer science department at Carnegie Mellon University. Dr. Lee is a member of the National Academy of Medicine and serves on the boards of the Allen Institute for Artificial Intelligence, the Brotman Baty Institute for Precision Medicine, and the Kaiser Permanente Bernard J. Tyson School of Medicine. He served on President Obama's Commission on Enhancing National Cybersecurity. He has testified before both the US House Science and Technology Committee and the US Senate Commerce Committee. With Carey Goldberg

and Dr.Isaac Kohane, he is the coauthor of the best-selling book, "The AI Revolution in Medicine: GPT-4 and Beyond." In2024, Peter Lee was named by Time magazine as one of the 100 most influential people in health and life sciences.

Bill Dally, Chief Scientist at Nvidia Research

Topic: Reflections on Deep Learning Hardware

Abstract:

The current resurgence of artificial intelligence, including generative AI like ChatGPT, is due to advances in deep learning. Systems based on deep learning now exceed human capability in speech recognition, object classification, and playing games like Go. Deep learning has been enabled by powerful, efficient computing hardware. The algorithms used have been around since the 1980s, but it has only been in the last decade - when powerful GPUs became available to train networks - that the technology has become practical. Advances in DL are now gated by hardware performance. Demand for training operations has increased by 10 million times in the last decade and is currently growing by 16x per year. In the last decade, GPU performance for deep learning has increased more than 1000x. Much of this gain was due to improvements in data representation. This talk will review this history and discuss future improvements.

Bio:

Dr. Bill Dally joined NVIDIA in January 2009 as chief scientist, after spending 12 years at Stanford University, where he was chairman of the computer science department. Dally and his Stanford team developed the system architecture, network architecture, signaling, routing and synchronization technology that is found in most large parallel computers today.

Dally was previously at the Massachusetts Institute of Technology from 1986 to 1997, where he and his team built the J-Machine and the M-Machine, experimental parallel computer systems that pioneered the separation of mechanism from programming models and demonstrated very low overhead synchronization and communication mechanisms. From 1983 to 1986, he was at California Institute of Technology (CalTech), where he designed the MOSSIM Simulation Engine and the Torus Routing chip, which pioneered "wormhole" routing and virtual-channel flow control.

He is a member of the National Academy of Engineering, a Fellow of the American Academy of Arts & Sciences, a Fellow of the IEEE and the ACM, and has received the ACM Eckert-Mauchly Award, the IEEE Seymour Cray Award, and the ACM Maurice Wilkes award. He has published over 250 papers, holds over 120 issued patents, and is an author of four textbooks. Dally received a bachelor's degree in Electrical Engineering from Virginia Tech, a master's in Electrical Engineering from Stanford University and a Ph.D. in Computer Science from CalTech. He was a cofounder of Velio Communications and Stream Processors. See https://www.linkedin.com/in/billdally/

Blaise Agüera y Arcas, CTO of Technology & Society, Google

Topic: Evolution, Intelligence, and Symbiosis

Abstract:
This talk will combine studies of abiogenesis (the transition from non-life to life), the "predictive brain" or active inference model of intelligence, and the concept of multiscale symbiotic prediction. I will argue that life and intelligence can be understood within this common framework, and that symbiotic cooperation is the key to scaling complexity and building more capable AI systems.
Bio:

Blaise Agüera y Arcas is Google's CTO of Technology & Society. He leads an organization working on basic research in AI, especially the foundations of neural computing, active inference, evolution, and sociality. In his tenure at Google he has led the design of augmentative, privacy-first, and collectively beneficial applications and he is the inventor of Federated Learning, an approach to training neural networks in a distributed setting that avoids sharing user data. Blaise also founded the Artists and Machine Intelligence program, and has been an active participant in cross-disciplinary dialogs about AI and ethics, fairness and bias, policy, and risk. Until 2014 he was a Distinguished Engineer at Microsoft. Outside the tech world, Blaise has worked on computational humanities projects including the digital reconstruction of Sergei Prokudin-Gorskii's color photography at the Library of Congress, and the use of computer vision techniques to shed new light on Gutenberg's printing technology. In 2018 and 2019 he taught the course "Intelligent Machinery, Identity, and Ethics" at the University of Washington, placing computing and AI in a broader historical and philosophical context. He has authored numerous papers, essays, op eds, and book chapters, as well two books: a novella, *Ubi Sunt*, and an interdisciplinary nonfiction work, *Who Are We Now?*. His upcoming book, *What Is Intelligence?*, will be published by MIT Press in 2025.

PANEL - PERPETUAL AND BOUNDLESS AI

Honorable Senator Maria Cantwell, leader

Dr. Peter Lee, President, Microsoft Research, panelist

Dr. Kristin Lauter, Sr. Director FAIR Labs, Meta, panelist

Dr. Chirag Shah, Co-Founder RAISE, UW, panelist

Dr. Blaise Agüera y Arcas, CTO of Technology & Society, Google, panelist

Dr. Sheree Wen, President,UNESCO Washington, moderator

Honorable Senator Maria Cantwell
Chair, Committee on Commerce, Science and Transportation

Sen. Maria Cantwell introduced the *Future of AI Innovation Act of 2024* which promotes strong partnerships between government, business, civil society, and academia to advance AI research; the bipartisan *Small Business Artificial Intelligence Training and Toolkit Act* of 2024 that would authorize to create and distribute artificial intelligence (AI) training resources and tools to help small businesses leverage AI in their operations and *AI Education Act of 2024* to expand scholarship and professional development opportunities to study artificial intelligence and quantum. These bills will open doors to AI for students at all levels, and upskill our workforce to drive American tech innovation, entrepreneurship and progress in solving the toughest global challenges and create jobs.

Afternoon Keynotes

Dr. Chirag Shah, Co-Founder RAISE, UW

Dr. Stefano Soatto, VP, AWS AI Labs and Professor, UCLA

Dr. Kristin Lauter, Sr. Director FAIR Labs, Meta

Dr. Juan Lavista Ferres, CVP, Chief Data Scientist, AI For Good Lab

Prof. Steve Brunton, NSF AI Institute in Dynamic Systems

Prof. Blake Hannaford, ECE, ME and Surgery, UW

Dr. Chirag Shah, Co-Founder RAISE, UW

Topic: Trust, But Verify: Verification and Validation of AI Systems

Abstract:

As AI systems penetrate more and more aspects of our lives, it becomes ever so important to ask how much can we trust these systems? From self-driving cars to healthcare decisions and from text generation to automated decision-making, we are increasingly relying on AI's capabilities. Blindly using these systems can be risky in many situations, but we also don't want to miss out on new capabilities that AI provides. So how do we use AI responsibly? I will argue in this talk that we need verification and validation as a way to ensure we can trust the systems we are using. I will show how to do this using human-in-the-loop for auditing and validating LLMs and their generation. The result is a human-AI collaborative mechanism that leads to responsibly leveraging the benefits of AI with manageable risks.

Bio:

Dr. Chirag Shah is Professor in Information School (iSchool) at University of Washington (UW) in Seattle. He is also Adjunct Professor with Paul G. Allen School of Computer Science & Engineering as well as Human Centered Design & Engineering (HCDE). He is the Founding Director for InfoSeeking Lab and Founding Co-Director of Center for Responsibility in AI Systems & Experiences (RAISE). His research involves building and studying intelligent information access systems, focusing on task-oriented search, proactive recommendations, and conversational systems. He is deeply engaged in work with generative AI, specifically in information access using large language models (LLMs). In addition to creating AI-driven information access systems that provide more personalized reactive and proactive recommendations, he is also focusing on making such systems transparent, fair, and free of biases. He is a Senior Member of IEEE and a Distinguished Member of ACM.

NEW ERA AI SUMMIT PROGRAM

Stefano Soatto, Vice President, AWS and Professor, UCLA

Topic: Where LLM Scaling Laws End

Abstract:

I will share the long-term vision driving the development of next-generation foundation models at AWS. Rather than just scaling current models, we start from the defining characteristics of optimal representations of sequential data, which pertain to any downstream task. These were established in the Seventies in the context of Stochastic Realization Theory. When the mechanisms generating data are non-stationary, as in business, finance, climate, and language data, optimal representations require effectively unbounded complexity, unthinkable in the Seventies but not today. So, if we scale computation, where does it end? Solomonoff answered that question in 1964, showing that optimal inference conspicuosly involves no "intelligence", no "insight", no "meaning", just cycling over programs; yet it performs optimal transduction. It does not even involve learning, just test-time computation over unbounded memory. Are scaling laws pointing us in that direction? Is too much of a good thing (scaling) not a good thing? I will discuss how current LLMs are effectively operating in the transductive setting, albeit suboptimally (through "in-context learning", which notably involves no learning), and have effectively unbounded memory through RAG. I will finally point to analytical tools that can be used to study the controllability, observability, and stability of the resulting systems, and the implications for their use and interaction with humans.

Bio:

Stefano Soatto is a Professor of Computer Science at the University of California, Los Angeles and a Vice President at Amazon Web Services, where he leads the AI Labs. He received his Ph.D. in Control and Dynamical Systems from the California Institute of Technology in 1996. Prior to joining UCLA he was Associate Professor of Biomedical and Electrical Engineering at Washington University in St. Louis, Assistant Professor of Mathematics at the University of Udine, and Postdoctoral Scholar in Applied Science at Harvard University. Before discovering the joy of engineering at the University of Padova under the guidance of Giorgio Picci, Soatto studied classics, participated in the Certamen Ciceronianum, co-founded the Jazz Fusion quintet Primigenia, skied competitively and rowed single-scull for the Italian National Rowing Team. Many broken bones later, he now considers a daily run around the block an achievement.

Soatto received the Siemens Prize with the Best Paper Award at CVPR in 1998 (with the late Roger Brockett), the Marr Prize at ICCV 1999 (with Jana Kosecka, Yi Ma, and Shankar Sastry), the Best Paper Award at ICRA 2015 (with Konstantine Tsotsos and Joshua Hernandez). He is a Fellow of the IEEE of the ACM.

NEW ERA AI SUMMIT PROGRAM

Dr. Kristin Lauter, Sr. Director FAIR Labs, Meta

Topic: Fundamental AI Research at Meta

Abstract:

This talk will highlight cutting-edge AI research at Meta, focusing on our open-source approach. AI shows great potential to help solve hard problems in science to benefit humanity. As we continue to build the foundations of general artificial intelligence, we also invest in developing algorithms to measure and improve AI safety and fairness. Responsible AI means that our systems need to PASS THE FIRE test: Privacy-Accessibility-Safety-Security Transparency-Health-Ethics Fairness-Inclusion-Robustness-Explainability. This talk will highlight some of my own research on Privacy and Safety.

Bio:
See https://ai.meta.com/people/786716476205590/kristin-e-lauter/

Dr. Juan Lavista Ferres, CVP, Chief Data Scientist, AI For Good Lab

Topic: AI For Good in the Era of GPT

Abstract:

As artificial intelligence (AI) technologies have matured and been deployed in production settings for over three decades, the advent of large language models and other forms of generative AI represents a significant paradigm shift. This talk will explore how these cutting-edge technologies are not just enhancing existing applications but also pioneering new ways to address some of the world's most pressing challenges. We will delve into the transformative potential of generative AI across various sectors, including healthcare, environmental sustainability, and accessibility, highlighting specific examples where AI has driven positive change and innovation. By exploring the intersection of technology and social good, this presentation aims to shed light on how generative AI can continue to be a formidable ally in our pursuit of a better future.

Bio:

Juan M. Lavista Ferres is the Corporate Vice President and Chief Data Scientist of the AI for Good Lab at Microsoft. Leading a team of dedicated data scientists and researchers in the domains of AI, Machine Learning, and statistical modeling, he partners with domain experts, researchers, and organizations worldwide to create a collaborative ecosystem that drives progress toward addressing some of the world's most pressing challenges.

Juan started the Microsoft AI for Good Lab efforts related to SIDS (Sudden Infant Death Syndrome), and his work was published in top academic Journals, such as Pediatrics. This initial success set the stage for him and the team to widen their focus to a multitude of global challenges such as digital literacy, equity, climate change, disaster response, conservation, health, cybersecurity, and more. The outcomes of these endeavors found resonance not only in respected

research journals but also in acclaimed news outlets, including Geekwire, BBC, NPR, Fast Company, Fortune Magazine, New York Times, CNN, and more, totaling over 100 global news platforms.

A sought-after speaker, Juan has been featured at Strata, IEEE, TedX, Cornell University, and UC Berkeley, among others. His influence extends across borders, as he's shared insights in the US, Canada, Switzerland, Argentina, Colombia, Costa Rica, New Zealand, Africa, and Uruguay. In addition, Juan's editorial leadership shapes the Microsoft Journal of Applied Research (MSJAR), helping to define AI and Data Science within Microsoft.

Prior to his current role, Juan's journey encompassed impactful roles in Microsoft's Experimentation Platform (EXP) and Bing Data Mining team, as well as co-founding Alerts.com. Drawing from a six-year tenure at the InterAmerican Development Bank, he applied data science to understand poverty reduction and equality in Latin America and the Caribbean. Juan dives in to address global challenges armed with a computer science degree from the Catholic University in Uruguay, a graduate degree in Data Mining and Machine Learning from Johns Hopkins University and a PhD in AI on Healthcare from Vrije Universiteit of Amsterdam. Juan resides in Kirkland, WA, with his wife and three children.

Prof. Steve Brunton, NSF AI Institute in Dynamic Systems

Topic: Machine Learning for Scientific Discovery, with Examples in Fluid Mechanics

Abstract:

This work will discuss several key challenges and opportunities in the use of machine learning for nonlinear system identification. In particular, I will describe how machine learning may be used to develop accurate and efficient nonlinear dynamical systems models for complex natural and engineered systems. I will emphasize the need for interpretable and generalizable data-driven models, such as the sparse identification of nonlinear dynamics (SINDy) algorithm, which identifies a minimal dynamical system model that balances model complexity with accuracy, avoiding overfitting. I will also introduce several key benchmark problems in dynamical systems and fluid dynamics that provide a diversity of metrics to assess modern system identification techniques. Because fluid dynamics is central to transportation, health, and defense systems, we will emphasize the importance of machine learning solutions that are interpretable, explainable, generalizable, and that respect known physics.

Bio:

Steven L. Brunton is a Professor of Mechanical Engineering at the University of Washington. He is also Adjunct Professor of Applied Mathematics, Aeronautics and astronautics, and Computer science, and he is also a Data Science Fellow at the eScience Institute. He is Director of the AI Center for Dynamics and Control (ACDC) at UW and is Associate Director for the NSF AI Institute in Dynamic Systems. Steve received the B.S. in mathematics from Caltech in 2006 and the Ph.D. in mechanical and aerospace engineering from Princeton in 2012. His research combines machine learning with dynamical systems to model and control systems in fluid dynamics, biolocomotion, optics, energy systems, and manufacturing. He received the Army and Air Force Young Investigator Program (YIP) awards and the Presidential Early Career Award for Scientists and Engineers (PECASE). Steve is also passionate about teaching math to engineers as co-author of four textbooks and through his popular YouTube channel, under the moniker "eigensteve". See https://dynamicsai.org/leadership/brunton/

Blake Hannaford, Professor of Electrical & Computer Engineering, Adjunct Professor of Mechanical Engineering and Surgery at the University of Washington, Seattle

Topic: Surgery via Robotics, AI and Internet

Abstract:

The science and art of surgery is rapidly incorporating today's most disruptive technologies to provide improved care and better access to care for patients. Robotics (today treating millions of patients a year) significantly reduces the invasiveness of pelvic surgeries such as hysterectomy and prostatectomy for faster and more comfortable recoveries. AI will increasingly enable semi-automated procedures and robotic assistants to surgeons. Experience with telemedicine during the 2020 Covid pandemic shutdowns has paved the way for Internet enabled remote surgery, potentially increasing access to surgical and other care for rural patients. This talk will review some research by interdisciplinary researchers from Engineering and Medicine at the University of Washington Biorobotics Laboratory into this future.

Bio:

Blake Hannaford received the B.S. degree in Engineering and Applied Science from Yale University in 1977, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of California, Berkeley. From 1986 to 1989 he worked on the remote control of robot manipulators in the Man-Machine Systems Group in the Automated Systems Section of the NASA Jet Propulsion Laboratory, Caltech and supervised that group from 1988 to 1989.

Since September 1989, he has been at the University of Washington in Seattle, where he is Professor of Electrical & Computer Engineering. He was named IEEE Fellow in 2005. He was at Google-X / Google Life Sciences / Verily from April 2014 to December 2015. He developed the Master of Science in Technology Innovation Robotics Degree Option at the GIX program in Bellevue WA from 2918 to 2020. And was a Hardware Research Manager at Meta Reality Labs Research for a year in 2023. He is currently doing full time teaching and research at UW ECE.

Research Breakthrough Showcase Speakers

Topic	Presenter
Building AI solutions that fit the jigsaw puzzle	**Sourabh Kulhare**, Global Health Labs
AI and Accessibility: Risks and Opportunities	**Dr. Jennifer Mankoff**, Paul G. Allen School of CS
Real-to-Sim-to-Real: A Scalable Data Diet for Robot Learning	**Dr. Abhishek Gupta**, Paul G. Allen School of CS

Sourabh Kulhare, Global Health Labs

Topic: Building AI solutions that fit the jigsaw puzzle

Abstract:

AI has potential for tremendous impact on health care, especially for the poorest people in our world community, who are most in need of better care. But AI solutions, however well-intentioned, are guaranteed to fail at the clinic if they don't match the specific constraints of the medical problem, because an AI solution is only one piece in a complex jigsaw puzzle. Therefore, as a necessary condition for success, we need to expand the traditional definition of "AI work" to include both understanding the non-AI pieces of the puzzle, and also translating these constraints into forces that will properly shape our AI solution.

Bio:

Sourabh Kulhare is an experienced Machine Learning Research Engineer at Global Health Labs working in multidisciplinary AI research groups. He focuses on developing efficient AI systems for low-cost healthcare applications. His research interests encompass deep learning architectures, sequence modeling, domain adaptation, generative modeling, and object detection. Sourabh holds a Master's degree (M.S) in Computer Engineering from the Rochester Institute of Technology, specializing in video summarization and natural language processing.
Linked In : https://www.linkedin.com/in/skrealworld/
Google Scholar :
https://scholar.google.com/citations?user=hTZnzOQAAAAJ&hl=en

Dr. Jennifer Mankoff, Center for Research and Education on Accessible Technology and Experiences (CREATE), University of Washington

Topic: AI and Accessibility: Risks and Opportunities

Abstract:

I will discuss how AI is impacting people with disabilities, both directly and indirectly. Among the positives, we are seeing AI use to solve problems that traditional accessibility tools either fall short of or do not even address. At the same time, AI has significant and deep biases against people with disabilities, and these biases may arise in situations that are out of the control of those affected, such as resume screening. I will discuss both as well as future needs and opportunities.

Bio:

Jennifer Mankoff directs the Center for Research and Education on Accessible Technology and Experiences (CREATE) at the University of Washington. CREATE's mission is to make technology accessible, and make the world accessible through technology. Mankoff's own research uses technologies including generative AI, data science, 3D printing, and computational knitting to solve accessibility problems. She strives to bring both structural and personal perspectives to her work. Jennifer received her PhD at Georgia Tech, advised by Gregory Abowd and Scott Hudson, and her B.A. from Oberlin College. She has identified as disabled since graduate school.

Dr. Abhishek Gupta, Paul G. Allen School, University of Washington

Topic: Real-to-Sim-to-Real: A Scalable Data Diet for Robot Learning

Abstract:

Robotic automation, powered by machine learning driven methods, has the potential to build systems that change the future of work, daily life and society at large by acting intelligently in human centric environments. As with most modern machine learning methods, a key component in building such a robotic system the availability of data, abundant, diverse and high quality. In domains of nature language or computer vision, data of this form has scaled passively with internet scale, since people naturally interact through the medium of language or images. In contrast, robots are hardly deployed in human-centric settings and certainly are not collecting internet scale data passively.

The key question I will ask is – how can we develop a data diet for robotic learning that scales passively? In particular, I will discuss how simulation, despite being fundamentally inaccurate, can provide a scalable source of data for robotic learning. We will discuss a class of real-to-sim-to-real methods that are able to construct simulation content on the fly from cheap real-world data, enabling scalable robust robot training. In doing so, I hope to shed some light on the unique challenge that data acquisition plays in robot learning and discuss how developing truly open-world robotic learning systems will necessitate a careful consideration of data quality and quantity.

Bio:

Abhishek Gupta is an assistant professor in computer science and engineering at the Paul G. Allen School at the University of Washington since 2022. He lead the Washington Embodied Intelligence and Robotics Development lab focusing on robot learning and reinforcement learning. Previously, he was a postdoctoral scholar at MIT, collaborating with Russ Tedrake and Pulkit Agarwal. Prior to that he received his Ph.D. and B.S degrees from UC Berkeley, working with Sergey Levine and Pieter Abbeel. Abhishek is the recipient of Toyota Research Institute Young Investigator award and an Amazon Science Hub award, along with award nominations at several top conferences and workshops. His research interests lie in scalable reinforcement learning methods for robot learning, in particular methods for continual adaptation in the real world.

Lightning Talks

Lightning Talk	Presenter
Data Quality in Intelligent Systems	**Sreyashi Das**, Senior Data Engineer at Netflix
PartyRock: Unleashing the Power of Generative AI in a Playground	**Karthikeyan Gopal**, Senior Software Development Engineer, Amazon
Building Trustworthy AI – Guardrails for a Responsible Future	**Ashish Vaidya**, Principal Software Engineer, Amazon
Artificial Intelligence for Sales and Marketing	**Harshita Asnani**, Applied Data Scientist, Amazon
Revolutionizing Organizational Security with Microsoft Copilot's AI Capabilities	**Ankit Masrani**, Principal Software Engineer, Microsoft Security
SLIs for ML Models	**Sunandan Barman**, Production Engineer, Meta Platforms
Speeding Up Product Innovation: Leveraging Generative AI for Rapid Product Development and GTM Strategies	**Ram Joshi**, Staff Engineer at The Climate Corporation, and **Gunjan Paliwal**, Sr. Product Development & Marketing Manager
Bridging the Chasm: How Software Innovation is Outpacing Hardware in the AI Revolution	**Dr. Harisankar Sadasivan**, Member of Technical Staff SDE, AMD
Building Product Strategy for Modern Times	**Seema Bansal**, Product Manager, Microsoft

Sreyashi Das, Senior Data Engineer at Netflix

Lightning Talk: Data Quality in Intelligent Systems

Abstract:

Data quality is a critical component in the development and deployment of intelligent systems, as it directly influences the accuracy, reliability, and trustworthiness of the insights and decisions derived from such systems. In this paper, we explore the multifaceted dimensions of data quality—including accuracy, completeness, consistency, timeliness, and relevance—and their impact on the performance of intelligent systems. We discuss the challenges and methodologies associated with ensuring high data quality in the context of rapidly evolving datasets and complex data environments. Furthermore, we examine the role of advanced data engineering techniques, such as data cleaning, validation, and enrichment, in maintaining and enhancing data quality. Through empirical studies and case analyses, we illustrate the profound effect that data quality has on machine learning models, predictive analytics, and real-time decision-making processes. Our findings underscore the necessity for robust data quality frameworks and best practices to harness the full potential of intelligent systems. By addressing these critical issues, we aim to provide actionable insights and strategies for researchers, practitioners, and organizations striving to optimize their intelligent systems for accuracy, reliability, and overall effectiveness.

Bio:

Sreyashi Das is a Senior Data Engineer at Netflix with a demonstrated history of working in media entertainment and consumer electronics industries. She excels in the design and implementation of both

streaming and batch data movement, along with analytical solutions. At Netflix she has developed new data products which provide the foundation for metric development and analysis insights for the Studio and Creative Production team. These products drive data health, launch timeliness, resource availability, and cost optimization. She has also designed a data extraction framework specifically for the animation team. Before her role at Netflix, her expertise was primarily in data warehousing and self-serve business intelligence. Sreyashi loves building high-quality data models.

Karthikeyan Gopal, Senior Software Development Engineer, Amazon

Lightning Talk: PartyRock: Unleashing the Power of Generative AI in a Playground

Abstract:

PartyRock is an exciting platform that empowers anyone to build AI-powered applications using the power of foundation models from Amazon Bedrock. In this lightning talk, I'll showcase how PartyRock's easy-to-use app builder allows users to quickly turn their ideas into working prototypes, all while learning about the latest advancements in generative AI. I'll walk through the key features of PartyRock, including how to create apps from scratch, remix existing apps, and leverage pre-built prompts to kickstart the creative process. By the end, attendees will see how PartyRock makes it fun and accessible for both technical and non-technical users to harness the potential of large language models and other generative AI technologies.

Bio:

Karthikeyan Gopal is a tenured senior software developer (~10 YOE) with extensive experience in building highly scalable, distributed, and multi-tenant systems. I have deep expertise in Java, SQL, RDBMS, and analytical queries, as well as proficiency in design patterns, caching, performance tuning, and system design. I also have deep expertise in Analytics, Payments and compliance domains. At Amazon, I have launched multiple payment products across different countries like Brazil, Singapore, Australia and Mexico. I have also led compliance projects, defining the technical strategy, and building components from the ground up for the emerging countries needs for Amazon. I've also served as a Bar Raiser, interviewing over 100 candidates, and have developed an AI powered "Vision" tool to help debug complex workflows across multiple services for all teams in Payments. Prior to Amazon, I was a Member Technical Staff at Zoho, where I significantly improved the performance of Zoho Analytics by 20x through content caching, precomputation, and in-memory database optimization over a period of 5 years. I hold a B.Tech in Computer Science and have a passion for problem-solving, as demonstrated by my participation and wins in various hackathons and coding events.

Ashish Vaidya, Principal Software Engineer, Amazon

Lightning Talk: Building Trustworthy AI – Guardrails for a responsible future

Abstract:

As AI rapidly transforms our world, ensuring its responsible development is of paramount importance. This talk explores the frameworks and practices to guide the creation of a responsible AI and introducing necessary guardrails to prevent its misuse. We'll delve into various key principles like fairness, transparency, and accountability, and how to keep these in focus throughout the AI development lifecycle. Discover how the guardrails can mitigate risks.

Bio:

Ashish is a Principal Software Engineer at Amazon. He has over 16 years of industry experience, and possesses a proven track record of designing, developing, and delivering high-quality solutions across a wide range of technologies. He is an expert in employing cloud technologies to build high scale, high availability, and low latency distributed systems, and improving their performance. In his current role, he is focused on designing and developing the future of Alexa – Amazon's Voice AI, by employing the cutting edge developments in the world of AI.

Beyond his work at Amazon, he actively shares his knowledge through tech talks, has co-authored a software career guide, and has four patents granted by USPTO with two more under review.

Harshita Asnani, Applied Data Scientist, Amazon

Lightning Talk: Artificial Intelligence for Sales and Marketing

Abstract:

Sales and Marketing are essential for success of every business, irrespective of size or industry. They are important for increasing revenue generation, lead generation, customer acquisition, product awareness etc. Artificial Intelligence can empower Sales and Marketing divisions activities by enabling pattern descriptions, better predictions, forecasts content generation etc. Through this talk we aim to introduce ways in which companies can leverage the power of Artificial Intelligence to improve Sales and Marketing activities.

Bio:

Harshita Asnani is an experienced Research Assistant and Software Development professional with a demonstrated history of working in the higher education industry and IT service provider industry. Skilled in R, Python (Programming Language), Data Modeling, Data Visualization, and Machine Learning. Strong research professional with a Master of Science – MS focused in Applied Data Science from Syracuse University.

NEW ERA AI SUMMIT PROGRAM

Sunandan Barman, Production Engineer, Meta Platforms

Lightning Talk: SLIs for ML Models

Abstract:

As machine learning (ML) models become increasingly critical to business operations, ensuring their reliability and performance is paramount. This talk explores the challenges of implementing Service Level Objectives (SLOs) for ML models and presents best practices for defining and monitoring custom SLOs that account for ML model specifics.

Using a case study, we demonstrate how implementing SLOs can lead to significant improvements in model accuracy and inference latency. Attendees will learn how to create effective SLOs for their ML models and improve the overall reliability and performance of their machine learning services. By attending this talk, you will gain a clear understanding of how to ensure the reliability and performance of your ML models, and how to define and monitor SLOs to achieve this goal.

Bio:

Sunandan Barman is a highly experienced backend developer with more than a decade of expertise in crafting scalable and distributed systems capable of handling immense loads. He has delivered large-scale projects consistently is proficient in setting direction for cross-functional teams on multi-year hyper-growth paths.

Linkedin : https://www.linkedin.com/in/sunandanbarman/

Ankit Masrani, Principal Software Engineer, Microsoft Security

Lightning Talk: Revolutionizing Organizational Security with Microsoft Copilot's AI Capabilities

Abstract:

Microsoft Copilot for Security is the first generative AI solution that empowers security and IT teams to protect at the speed and scale of AI, while remaining compliant to responsible AI principles.This talk will explore the unique advantages of Microsoft Copilot for Security, including its use of a specialised language model, seamless integration with specific security signals, and interoperability with other Microsoft security products. We will then delve into the practical applications such as investigating security vulnerabilities, reverse engineering malicious scripts and providing guided responses for expedited remediation. Join us to discover how Microsoft Copilot for Security can revolutionise your security operations and enhance your organisation's defences.

Bio:

Ankit Masrani is a Principal Software Engineer at Microsoft with over a decade of experience in building scalable and secure distributed systems. He has been instrumental in the development of various software products from the ground up across various global organizations like Amazon and Microsoft, earning recognition and patents for innovative contributions. Ankit is part of the Microsoft hiring panel and actively

mentors new and junior engineers, fostering their growth and success. He is a IEEE Senior Member, participates as an industry expert judge in hackathons and technology awards and contributes to research articles.

Ram Joshi (Staff Engineer at The Climate Corporation), Gunjan Paliwal (Product Development & Marketing Leader)

Lightning Talk: Speeding Up Product Innovation - Leveraging Generative AI for Rapid Product Development and GTM Strategies

Abstract:

In today's fast-paced market, speeding up product innovation and reducing time-to-market is crucial for staying ahead. This 10-minute lightning talk will highlight how generative AI can transform product development and go-to-market (GTM) strategies. We'll briefly explore how generative AI accelerates various stages—from ideation and design to development and launch—enhancing efficiency and creativity. Attendees will gain quick insights into some practical applications, learn about real-world examples of successful AI integration, and discover actionable strategies to overcome common challenges. This talk will provide you with a focused overview of leveraging generative AI for faster product innovation and effective product launches.

Bios:

Ram Joshi is an experienced Staff Engineer at The Climate Corporation in Seattle. With over a decade of expertise in software development across the stack, Ram currently specializes in mobile and web application development, and his previously worked in distributed systems engineering. He holds a Master's degree in Information Management and Systems, with a focus on Information Retrieval, Artificial Intelligence, and Web Architecture from the University of California, Berkeley.

Ram is committed to education and has taught the INFO 153 Web Architecture and Information Management course at UC Berkeley. His has expertise in various tools and technologies such as Python, Java, Javascript, C and Swift. With his blend of industry experience and academic background, Ram is always excited to tackle new and interesting challenges in software engineering.

Gunjan Paliwal is a seasoned, data-driven Product Marketing and Development Leader with over 10 years of experience at industry leaders such as Meta, Newell, Sears, and Microsoft. She specializes in AdTech, Marketplace, E-commerce, and B2B.
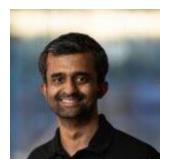
Dr. Harisankar Sadasivan, Member of Technical Staff SDE, AMD | Faculty (part-time), UW Seattle | Sr Member, IEEE

Lightning Talk: Bridging the Chasm: How Software Innovation is Outpacing Hardware in the AI Revolution

Abstract:

In the relentless march of AI, we face an unprecedented challenge: compute demand is skyrocketing, leaving Moore's Law and even cutting-edge heterogeneous hardware designs in the dust. This isn't just a gap—it's a chasm, yawning wider by orders of magnitude with each AI breakthrough. Enter the unsung hero of this technological arms race: software optimization. In this lightning talk, we'll dive into the heart of this computational battleground, exploring the meteoric rise of AI's computational hunger and its implications for the tech landscape. We'll examine why traditional hardware solutions are falling short, how they're evolving to keep pace, and the critical role of aggressive software optimizations in bridging the ever-widening performance gap. By dissecting emerging patterns in AI workload characteristics that are reshaping our approach to system design, we'll unveil the symbiotic evolution of hardware and software stacks in response to these unprecedented demands. Join us for a whirlwind tour of the cutting edge, where software ingenuity is rewriting the rules of the game in our quest to satiate AI's insatiable appetite for compute power.

Bio:

Dr. Hari Sadasivan is a Member of Technical Staff SDE at AMD, a part-time faculty member at the University of Washington, Seattle, and an IEEE Senior Member. Hari focuses on systems for AI and life-sciences. Hari drives the AMD Center of Excellence in AI at the Allen School of CSE and is the co-founder of the AMD-omics focus group which collaborates with research institutes world over. Hari holds a PhD in CSE from the University of Michigan Ann Arbor and is a peer-reviewer for several IEEE/ACM venues.

Seema Bansal, Product Manager, Microsoft

Lightning Talk: Building Product Strategy for Modern Times

Abstract:

As Product leaders, we have been there where the product team puts processes, roadmaps, and best practices in place for how to deliver products that users love. But at the same time, the product team gets pulled into different directions due to organizational changes or due to recent market competition developments. As product lead, when such changes happen, teams transact – how to ship in half the time, with a limited headcount, and across the breadth of platforms in a competitive market. This was the task at hand for Seema Bansal, when developing features for Microsoft Teams' 300 million monthly active users, across Android, iOS, Mac, Windows, and web apps in 181 countries and 19 languages. In this talk, Seema Bansal, Product Manager at Microsoft, will share scalable strategies that can be adopted to design and launch features in record time with group intelligence, empathy and quick decision-making, while keeping user delight and company culture in focus. She'll call out potential hurdles, pinpoint wasteful efforts that can be avoided, and approaches she took to maximize yield for her team. The session will cover examples from her product management career, including how she managed ever-changing user

expectations during the pandemic and integration of large language models into the product. Attendees will gain insights into delivering features quickly without compromising on quality.

Bio:

Seema Bansal, PM Microsoft

Seema Bansal is a Microsoft Product Manager with over 8 years of international PM experience. She is currently building features for 300M users of Microsoft Teams. Before working on Teams, she was managing security product in Cloud and AI within Microsoft Azure. Her computer science engineering & visual design academic background matched with her entrepreneurial endeavors has framedhow she approaches building and scaling products. It's through her empathy-driven product building approach, curiosity, and incredible mentors that have helped her grow into a seasoned product leader.
See https://www.linkedin.com/in/seema-bansal/