UC San Diego
SCHOOL OF COMPUTING, INFORMATION AND DATA SCIENCES

Explainable & Secure Computing/AI Symposium Agenda

February 13th | 8:30 am-2:00 pm

Franklin Antonio Hall 4202, UC San Diego
3180 Voigt Dr, La Jolla, CA 92093

8:30 – 9:00 am: Check-in & Coffee

9:00 – 9:30 am: Welcome Remarks & AI-Health Overview
Erik Mjoen, Director, Industry Relations, SCIDS

9:30 – 11:00 am: Session 1 – Talks

Towards Trustworthy and Responsible AI: Automated Interpretability, Adversarial Robustness, and AI safety
Lily Weng, Assistant Professor, HDSI

Securing the Data Layer: Access Control and Privacy in Research Data Management
Subhasis Dasgupta, Researcher, San Diego Supercomputer Center

From Data to Decisions: Secure and Explainable AI Across the Property Lifecycle
Kien Trinh, Sr. Principal Machine Learning Scientist, Cotality
Swetha Varadharajan, Principal Data Scientist, Cotality

11:00 – 11:10 am: Break

11:10 – 12:10 pm: Session 2 – Panel

Moderator: Yu-Xiang Wang, Associate Professor, HDSI

Lily Weng, Assistant Professor, HDSI
Subhasis Dasgupta, Researcher, San Diego Supercomputer Center
Sarah Aerni, VP of Technology/AI, Intuit
Barry Kunst, VP of Marketing, Solix Technologies

12:10 – 12:40 pm: Discussion: Collectively defining emerging opportunities

12:40 – 1:00 pm: Closing Remarks

1:00 – 2:00 pm: Lunch

Speaker Bio & Talk Description

Lily Weng

Lily Weng is an Assistant Professor in the Halıcıoğlu Data Science Institute at UC San Diego with an affiliation in the CSE department. She received her PhD in Electrical Engineering and Computer Science (EECS) from MIT in August 2020, and her Bachelor and Master degree both in Electrical Engineering at National Taiwan University. Prior to UCSD, she spent 1 year in MIT-IBM Watson AI Lab and several research internships in Google DeepMind, IBM Research and Mitsubishi Electric Research Lab. Her research interest is in machine learning and deep learning, with primary focus on Trustworthy AI. Her vision is to make the next generation AI systems and deep learning algorithms more robust, reliable, explainable, trustworthy and safer. Her work has been recognized and supported by multiple NSF awards, ARL award, Intel Rising Star Faculty Award, Hellman Fellowship, and Nvidia Academic award. For more details, please see https://lilywenglab.github.io/.

Towards Trustworthy and Responsible AI: Automated Interpretability, Adversarial Robustness, and AI safety

Deep learning models have become remarkably powerful – but often operate as black boxes. In this talk, I will share how my lab is making these systems more transparent, reliable, and trustworthy. I'll highlight three research directions to bring interpretability into deep learning: (1) automated tools [1-4] that reveal what neural networks learn internally at scale; (2) interpretable model designs [5-8] that make model's decision process more understandable and controllable; and (3) evaluation frameworks [9-12] that quantify interpretability and enable trust. I'll also touch on our recent work [13-16] in robust learning and jailbreak attacks for safer AI deployment. Together, these efforts aim to move modern AI beyond accuracy – toward systems we can truly understand, align, and trust. For more details, please visit https://lilywenglab.github.io/.

Subhasis Dasgupta

Dr. Dasgupta is a researcher at the San Diego Supercomputer Center, known for his innovative work in data management and security. He developed the AWESOME polystore system, which has greatly improved the way complex and varied data are integrated and analyzed. His expertise has been key in important projects like the National Data Platform (NDP) and the NOURISH project, showing his dedication to solving big societal challenges through data science. With a strong background in access control and a range of important publications, patents, and books, Dr. Dasgupta has made significant contributions to secure data processing systems. His work connects scientific research with technological advancements, which are crucial for public health and national research efforts. Additionally, Dr. Dasgupta was one of the founders of the cloud management company Kaavo Inc. Dr. Dasgupta has also worked on projects like DER Security and Quantum Data Hub and was a key member of the China Data Lab, where he set up a facility for integrating information across China to understand Chinese policy, which was well appreciated by Social scientists and Legal Scholars. He is currently involved in cutting-edge projects like COVID-19 monitoring, interdialytic hypertension, green energy, etc. He also advises various labs in the medical and engineering schools in the USA, UK, and India.

Securing the Data Layer: Access Control and Privacy in Research Data Management

Modern AI and machine learning systems increasingly rely on large, diverse research datasets. However, questions about data access, conditions, and privacy are often addressed only after models are developed. This talk highlights the potential for a secure data architecture and explores how it could be implemented in a research lab environment. Drawing on practical experience with access control and anonymization techniques in research data platforms, I will discuss how fine-grained data access, privacy-preserving views, and policy-aware governance affect downstream AI pipelines. The presentation will cover key challenges, design principles, and lessons learned for securing research data pipelines that support responsible, explainable, and collaborative AI.

Swetha Maithreyi Varadharajan

Swetha Maithreyi Varadharajan is a Machine Learning Scientist at Cotality, with expertise in Generative AI, statistical modeling, and large-scale machine learning systems within the property valuation and real estate domain. As a technical lead, she is committed to mentoring talent, shaping ML best practices, and enabling teams to translate research breakthroughs into real-world impact. Swetha holds a Master's degree in Data Science from the University of California, San Diego.

Kien Trinh

Dr. Kien Trinh is a Machine Learning Scientist at Cotality, where he leverages data science and Generative AI to create advanced solutions for property valuation, pricing, and enriched property data. He holds a PhD in Physics from the University of Southern California. Kien's professional background spans diverse domains, including text mining, location-based mobile advertising, and real estate analytics.

From Data to Decisions: Secure and Explainable AI Across the Property Lifecycle

This presentation highlights how Cotality applies secure and explainable AI to high-impact challenges across the property ecosystem. We outline the diverse markets served—enterprise data, real estate, mortgage, and insurance—where trusted analytics rely on large-scale, high-quality datasets and deep scientific expertise.

Through real-world use cases including fraud detection, property resiliency scoring, and high-volume document extraction, we show how confidence scoring, explainability, and rigorous validation enable reliable decision-making in complex environments. The session provides a clear view of how secure, transparent AI systems can be designed and deployed responsibly at scale.