2022 International Workshop on **Trustworthy AI**

Monday, August 1, 2022 | 9:00 PM (EDT)

Tuesday, August 2, 2022 | 9:00 AM (Taiwan Time)

Online

Advisor INDUSTRIAL DEVELOPMENT BUREAU, MINISTRY OF ECONOMIC AFFAIRS 經濟部工業局

Organizer ITRI Industrial Technology Research Institute

Co-organizers FOXCONN 鴻海科技集團 | AI Labs.tw 台灣人工智慧實驗室 | fido ALLIANCE | 台灣雲端物聯網產業協會 Cloud Computing & IoT Association in Taiwan | 台灣資通產業標準協會 Taiwan Association of Information and Communication Standards

Workshop Agenda		
09:00 - 09:20	**Welcome and Opening Remarks**	
	• **Pei-Zen Chang**, Executive Vice President, ITRI	
	• **Feng Tang (Audrey Tang)**, Minister without Portfolio, Executive Yuan, R.O.C. (Taiwan)	
	• **Jang-Hwa Leu**, Director General, Industrial Development Bureau, Ministry of Economic Affairs, R.O.C. (Taiwan)	
	• **Yuh-Jye Lee**, Deputy Executive Secretary, Office of Science & Technology, Executive Yuan, R.O.C. (Taiwan)	
	Keynote Session	Session Chair: **Dr. Yuh-Jye Lee**
09:20 - 09:45	AI Risk Management	• **Elham Tabassi**, ITL Chief of Staff, NIST
09:45 - 10:10	Trustworthy AI	• **Jeannette M. Wing**, Executive Vice President for Research, Columbia University
10:10 – 10:25	Trustworthy Federated Learning	• **Ethan Tu**, Founder of Taiwan AI Labs
10:25 – 10:35	**Break**	
	Featured Session	Session Chair: **Ms. Karen Chang**, Chair of FIDO Taiwan Regional Forum
10:35 - 10:50	Feature Disentanglement for AI Ethics	• **Shang-Wei Chou**, Principal Engineer, Egis Technology Inc.
10:50 – 11:05	Current Status and Future Perspective of AI Development in Taiwan	• **Kai-Lung Hua**, Deputy General Director of ICL, ITRI
	Panel Discussion	Moderator: **Jane Yung-jen Hsu**, Professor of Computer Science and Information Engineering, National Taiwan University
11:05 – 12:00	Panelists:	• **Yuh-Jye Lee**, Deputy Executive Secretary, Office of Science & Technology, Executive Yuan, R.O.C. (Taiwan) • **Elham Tabassi**, ITL Chief of Staff, NIST • **Wei-Bin Lee,** CEO of Hon Hai Research Institute • **Shang-Wei Chou**, Principal Engineer, Egis Technology Inc. • **Tzeng-Yow Lin**, Chief Executive of National Measurement Laboratory and General Director of

The organizer reserves the rights to change the workshop format, agenda or any parts of the event if necessary.

Welcome and Opening Remarks

General Chair

Pei-Zen Chang

Executive Vice President, ITRI

Advisor

Feng Tang (Audrey Tang)

Minister without Portfolio, Executive Yuan, R.O.C. (Taiwan)

Advisor

Jang-Hwa Leu

Director General, Industrial Development Bureau, Ministry of Economic Affairs, R.O.C. (Taiwan)

Guest

Yuh-Jye Lee

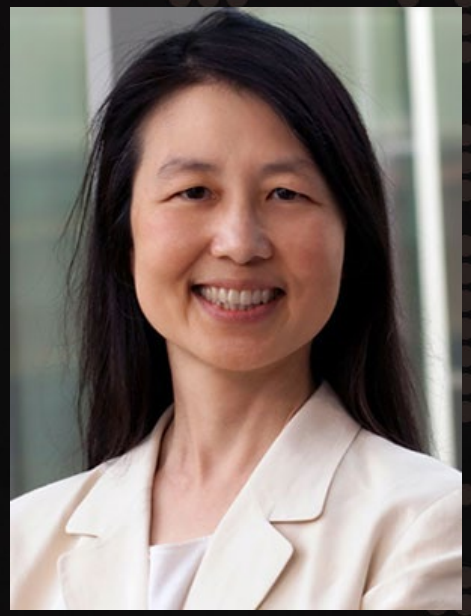
Deputy Executive Secretary, Office of Science & Technology, Executive Yuan, R.O.C. (Taiwan)

Elham Tabassi

ITL Chief of Staff,
NIST

Elham Tabassi is the Chief of Staff in the Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST). She leads NIST Trustworthy and Responsible AI program that aims to cultivate trust in the design, development, and use of AI technologies by improving measurement science, standards, and related tools in ways that enhance economic security and improve quality of life. She has been working on various machine learning and computer vision research projects with applications in biometrics evaluation and standards since she joined NIST in 1999. She is a member of the National AI Resource Research Task Force, a senior member of IEEE, and a fellow of Washington Academy of Sciences.

Jeannette M. Wing



Executive Vice President for Research, Columbia University

Jeannette M. Wing is the Executive Vice President for Research at Columbia University and Professor of Computer Science. In her EVPR role, she has overall responsibility for the University's research enterprise at all New York locations and internationally. The New York locations include the Morningside and Manhattanville campuses, Columbia University Irving Medical Center, Lamont-Doherty Earth Observatory, and Nevis Laboratories. She joined Columbia in 2017 as the inaugural Avanessians Director of the Data Science Institute.

Prior to Columbia, Dr. Wing was Corporate Vice President of Microsoft Research, served on the faculty and as department head in computer science at Carnegie Mellon University, and served as Assistant Director for Computer and Information Science and Engineering at the National Science Foundation.

Dr. Wing's research contributions have been in the areas of trustworthy AI, security and privacy, specification and verification, concurrent and distributed systems, programming languages, and software engineering. Her 2006 seminal essay, titled "Computational Thinking," is credited with helping to establish the centrality of computer science to problem-solving in fields where previously it had not been embraced, and thereby influencing K-12 and university curricula worldwide.

She is a Fellow of the American Academy of Arts and Sciences, American Association for the Advancement of Science, the Association for Computing Machinery (ACM), and the Institute of Electrical and Electronic Engineers. She received distinguished service awards from the ACM and the Computing Research Association and an honorary doctorate degree from Linköping University, Sweden. She earned her bachelor's, master's, and doctoral degrees in computer science, all from the Massachusetts Institute of Technology.

Ethan Tu

Founder of
Taiwan AI Labs

Ethan Tu is most well-known as the founder of PTT, Taiwan's largest and most popular online bulletin board system. In the US, he worked at the National Institute of Health (NIH) to develop cancer detection systems and human genetic research. He also worked as Microsoft's Director of Research and Development for AI in the Asia-Pacific Region for 11 years and was a key figure in development of Microsoft's Cortana.

In 2017, Ethan founded Taiwan AI Labs, Asia's first open AI research organization specializing in next-generation AI solutions. AI Labs has shown significant results in medical AI/imaging applications, which contributed to the detection of Covid-19 from chest x-rays during the pandemic. He also started The Taiwan AI Federated Learning Alliance to encourage the pooling of data samples for better AI advancement while protecting privacy rights. Ethan also leads teams to develop AI solutions for drones, Smart Cities, speech and facial recognition, and the fighting of fake news.

Shang-Wei Chou

Principal Engineer,
Egis Technology Inc.

- Education
  - Chung Yuan Christian University, Taoyuan City PhD in Electronic Engineering
- Experience
  - Principal Engineer. Egis Technology Inc. (2021-)
    - Mixed-mode SRAM-based CIM architecture development
    - Self-learning algorithm development for noise suppression
  - CEO/GM, Silicon Jazz Co., Ltd. (Startup 2017-2021)
    - Response for company's general management included strategy, finance and technology.
    - Defining technology roadmap and associated foundries strategy
    - Building overall company operation and product development.
    - In charge of mixed-mode AI IC/IP marketing and sales
- Skill Summary
  - Sound experience of mixed-mode Computing-in-Memory Circuits for Deep Learning and AI chips
  - CMOS and memory product and circuit design
  - Semiconductor CMOS and memory device development and applications

Kai-Lung Hua

Deputy General Director of ICL, ITRI

Dr. Kai-Lung Hua is the Deputy General Director of Information and Communications Research Laboratories (ICL) at Industrial Technology Research Institute (ITRI). Prior to joining ITRI, Dr. Hua held the positions of Professor of the Department of Computer Science and Information Engineering, Director of the Artificial Intelligence Research Center, and Dean of the Office of Industry-Academia Collaboration at National Taiwan University of Science and Technology. He received the Ph.D. degree from the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA, in 2010.

He is a Member of Eta Kappa Nu and Phi Tau Phi. His current research interests include digital image and video processing, computer vision, and machine learning. He has published more than 150 papers in international journals and conferences. He was the recipient of the MediaTek Doctoral Fellowship and several research awards, including the 2020/2019 Outstanding Research Award of Taiwan Tech, 2018 Young Scholar Award of Taiwan Tech, Top Performance Award of 2017 ACM Multimedia Grand Challenges, Top 10 % Paper Award of 2015 IEEE International Workshop on Multimedia Signal Processing, the Second Award of the 2014 ACM Multimedia Grand Challenge, the Best Paper Award of the 2013 IEEE International Symposium on Consumer Electronics, and the Best Poster Paper Award of the 2012 International Conference on 3D Systems and Applications. His research has been funded by many government agencies and industrial companies such as Qualcomm, Nokia, and Intel.

Panel Discussion

Moderator

Jane Yung-jen Hsu

Professor of Computer Science and Information Engineering, National Taiwan University

Panelist

Yuh-Jye Lee

Deputy Executive Secretary, Office of Science & Technology, Executive Yuan, R.O.C. (Taiwan)

Panelist

Elham Tabassi

ITL Chief of Staff, NIST

Panelist

Wei-Bin Lee

CEO of Hon Hai Research Institute

Panelist

Shang-Wei Chou

Principal Engineer, Egis Technology Inc.

Panelist

Tzeng-Yow Lin

Chief Executive of National Measurement Laboratory and General Director of CMS/ITRI

AI Risk Management

Elham Tabassi
ITL Chief of Staff, NIST

Abstract

AI systems sometimes do not operate as intended because they are making inferences from patterns observed in data rather than a true understanding of what causes those patterns. Ensuring that these inferences are helpful and not harmful in particular use cases – especially when inferences are rapidly scaled and amplified – is fundamental to trustworthy AI. While answers to the question of what makes an AI technology trustworthy differ, there are certain key characteristics which support trustworthiness, including accuracy, explainability and interpretability, privacy, reliability, robustness, safety, security (resilience) and mitigation of harmful bias. There also are key guiding principles to take into account such as accountability, fairness, and equity. Cultivating trust and communication about how to understand and manage the risks of AI systems will help create opportunities for innovation and realize the full potential of this technology.

This presentation overviews NIST's effort in developing a framework to better manage risks to individuals, organizations, and society associated with AI. The NIST Artificial Intelligence Risk Management Framework (AI RMF or Framework) is intended for voluntary use and to improve the ability to incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems.

Trustworthy AI

Jeannette M. Wing
Executive Vice President for Research, Columbia University

Abstract

Recent years have seen an astounding growth in deployment of AI systems in critical domains such as autonomous vehicles, criminal justice, healthcare, hiring, housing, human resource management, law enforcement, and public safety, where decisions taken by AI agents directly impact human lives. Consequently, there is an increasing concern if these decisions can be trusted to be correct, reliable, fair, and safe, especially under adversarial attacks. How then can we deliver on the promise of the benefits of AI but address these scenarios that have life-critical consequences for people and society? In short, how can we achieve trustworthy AI?

Under the umbrella of trustworthy computing, there is a long-established framework employing formal methods and verification techniques for ensuring trust properties like reliability, security, and privacy of traditional software and hardware systems. Just as for trustworthy computing, formal verification could be an effective approach for building trust in AI-based systems. However, the set of properties needs to be extended beyond reliability, security, and privacy to include fairness, robustness, probabilistic accuracy under uncertainty, and other properties yet to be identified and defined. Further, there is a need for new property specifications and verification techniques to handle new kinds of artifacts, e.g., data distributions, probabilistic programs, and machine learning based models that may learn and adapt automatically over time. This talk will pose a new research agenda, from a formal methods perspective, for us to increase trust in AI systems.

Feature Disentanglement for AI Ethics

Abstract

Shang-Wei Chou,
Principal Engineer, Egis Technology Inc.

AI is bringing convenience to a whole new level. It has a role in all aspects of life from the home to shopping and to the workplace. With the accumulation of vast amounts of data, AI may know users better than they know themselves. However, AI heavily relies on human knowledge and experience in the areas. It is an unexplainable black box for many people. Because lack of transparency, resulting in a lack of understanding about how systems can work. Users raise concerns about opacity and threats to individual privacy and autonomy. In fact, privacy issues could be solved by technical solutions. For example, through technologies such as federated learning or encryption, data can be provided while maintaining privacy.

In this speech, Egis will introduce the Feature Disentanglement to protect privacy. The feature and content can be extracted separately from the data by this technology. Users don’t need to upload the feature information to the cloud but only content. The feature data as long as stored in their own devices. We use Feature Disentanglement to implement noise prediction and suppression as an example. The working principle of this noise suppression is not to block the noise, but using sound waves to against it. A key point is we only suppress the noise with specific frequency but not all sounds in the environment. Through the technology of Feature Disentanglement, the frequency spectrum features and amplitude are separated to model training at engineering side. When the product is delivered to users, the adversarial sound of specific frequency can be predicted successfully by noise amplitude.

Current Status and Future Perspective of AI Development in Taiwan

Kai-Lung Hua
Deputy General Director of ICL, ITRI

Abstract

The Executive Yuan of Taiwan rolled out the Taiwan AI Action Plan (2018-2021) to increase Taiwan's technological advantages, and prioritize innovation and real-world implementation, thereby injecting new momentum into our industries. This Action Plan consists of five major components: AI talent program, AI pilot project, AI international innovation hub, test fields and regulatory co-creation, and AI for industrial innovation. The first half of this speech will elaborate on the current achievements of AI for industrial innovation, sharing visible results from different business categories such as manufacturing, medicine, innovative business, and autonomous mobility.

AI technology innovates and simultaneously presents unprecedented challenges. SMEs, compared to large enterprises, lack the resources to recruit talent and acquire data for model training, thus limiting their development. Given this, Taiwan's government is enacting the next Taiwan AI Action Plan (2022-2026), which is the focus of the second half of the speech, to nurture trustworthy AI technology. Taiwan aims to develop one-stop AI solutions, to accelerate the development and implementation of AI applications. The government also intends to build testing and assessment centers for AI technologies to ensure the quality and reliability of AI products while expanding the current market. The ultimate goal for such AI technological advancement is to realize Taiwan's 2030 vision for innovation, inclusion, and sustainability.