# **New Frontiers in Computing 2022**

# **Emerging trends in Applied Data Science**

https://r6.ieee.org/scv-cs/nfic-2022-applied-data-science/



# Saturday August 13<sup>th</sup>, 2022 4pm - 8pm PDT







## Message from Conference Chair, Gary Ni



The landscape of business today is in a constant state of change, and data science makes a significant impact on the industry. The emergence of data science as a field of study and practical application over the last century has led to the development of technologies such as deep learning, natural language processing, and computer vision. As human interaction with technology grows daily, the amount of data generated on a routine basis is the amount of data generated and consumed currently in scores of ZetaByte per month and growing at about 20% compounded on an annual basis. This data is found in raw form and is of immense value to business and research.

This conference is a collaboration between the North American Taiwanese Engineering & Science Association (NATEA) and IEEE Computer Society of Silicon Valley to bring forth the dreams and aspirations of uplifting communities for higher living standards. NFIC's vision is to leverage new frontiers in computing and communications by being at the forefront of emerging technologies. At this conference, we bring together researchers and practitioners from academia and industry to exchange their research ideas. In addition, you will hear from visionaries, practitioners, and industry experts about how this vision is being brought to life.

With our sights set in 2022, we look optimistically to the future and the opportunities and challenges ahead. I hope you enjoy the variety of speakers we have brought for you today. Thank you to the program committee, IEEE SCV Computer Society Chapter, speakers, and volunteers for all the hard work and efforts in making this event a success. And finally, thank YOU for taking the time to participate and engage with our community as we move towards more creative and connected communities.

## About this event

Organized by North America Taiwanese Engineering and Science Association (NATEA), Silicon Valley Chapter and IEEE SCV Computer Chapter, this conference is in its 24th year. The first conference was organized in the year 1999.

Time (PST)	Presentation Title	Speaker
4:00 – 4:10 PM	Opening Remarks	PC Chairs (NATEA and IEEE-CS)
4:10 – 4:55 PM	Keynote Speech: Explainable Al for medical imaging informatics (including Q&A)	Prof. KC Santosh, Chair, Department of Computer Science, University of South Dakota
4:55 – 5:40 PM	Machine Learning for Misinformation Containment: A Candid Assessment of the State of the Art	Dr. Vishnu S. Pendyala, San Jose State University
5:40 – 6:25 PM	Yolo4 and Its Applications (including Q&A)	Prof. Mark Liao, NTUT, Taipei, Taiwan
6:25 – 7:10 PM	<b>3-D Digital Pathology- Inspired AI for Precision Diagnosis</b> (including Q&A)	Dr. Yen-Yin Lin and Dr. Tun-Wen Pai, Taipei, Taiwan
7:10 – 7:45 PM	Trends in Securing AI Data Pipeline	Prakash Ramchandran, Executive Secretary, eOTF
7:45 – 8:00 PM	Closing Remarks	PC Chairs (NATEA and IEEE-CS)

#### Program Agenda

## Introduction to New Frontiers in Computing

By John Pang Yu, Ph.D. NFIC Founding Council Member

Keeping up to speed, especially in Silicon Valley, is crucial for today's engineers and scientists. Possessing adequate knowledge to make reasoned decisions regarding an emerging technology's importance plays an essential role in professional career success. The New Frontiers in Computing (NFIC) conference was initiated by two members each from NATEA and IEEE-CS Santa Clara Valley in 1999 and has been organized to provide engineers and scientists with means to achieve that goal.

Held annually since 1999, the New Frontiers in Computing Conference started providing computer and engineering professionals with enough technical information on a developing field to make informed decisions as to its role in their professional careers. It has achieved this for the last 23 years by presenting an inexpensive one-day or half-day conference on emerging technologies. The following topics were presented by recognized leaders and experts from both the research and the emerging application's communities.

Past NFIC Conferences:

- 2021: Emerging Technologies of Artificial Intelligence and Beyond
- 2020: Postponed due to COVID-19

2019: Network of the Future, Emerging Technologies for 5G and Beyond, Stanford University

- 2018: Accelerating Smart and Connected Communities, Santa Clara University
- 2017: FinTech: Hedging the Financial Storm!, Stanford University
- 2016: Cognitive Computing: to the Singularity and Beyond, Stanford University
- 2015: Smart Grid Ecosystem- The Public and The Personal, Stanford University
- 2014: The Future of Online Education, Stanford University
- 2013: Social Network Analysis: It's Who You Know, Stanford University
- 2012: Emerging Medical Computing: Health Care Up Close & Personalized, Stanford University
- 2011: Emerging Automotive Computing Engineering in Overdrive, Stanford University
- 2010: Cloud Computing and the Web, San Jose State University
- 2009: Handheld Devices, Stanford University
- 2008: Cloud Computing: The New Faces of Computing, Stanford University
- 2007: Multi-Core Processors, Stanford University
- 2006: The World with RFID, Biltmore Hotel, Santa Clara

2005: Sensor Networks - The New Environment, Stanford University

2004: Semiconductors to Nanotechnology - The Coming Convergence, Stanford University

2003: Emerging Issues in Security, Mobility and Privacy, Stanford University 2002: Bioinformatics, Stanford University

2001: Nanotechnologies, Stanford University 2000: Internet Protocol Telephony, Stanford University

1999: Systems on a Chip, Stanford University

# Introduction of NFIC Conference Organizers: NATEA and IEEE-CS Santa Clara Valley

#### About NATEA

The North America Taiwanese Engineering & Science Association (formerly North America Taiwanese Engineers' Association: NATEA - http://www.nates.org) was founded on March 2, 1991 in Silicon Valley, California. Since then, NATEA has grown into twelve (12) Regional Chapters across North America including two in Canada and 10 in major cities of the US, with more than 3,500 members. Among them, the NATEA Silicon Valley Chapter is the largest one with more than 1500 members, and thousands of supporters engaging in social media interaction. Many members serve in leadingedge technical and managerial positions, and are founders of some of the most successful high-tech companies.

Since 2018, the NATEA has transformed to NATEA 2.0 empowering digital transformation for providing a new and enhanced platform for attracting young professionals to participate in leadership pipeline.

Most noticeably, the NATEA mission has extended to caring humanity in addition to science and technology with encouragement received from its Honorary Advisory Council Member Dr. Yuan Tseh Lee who is a Nobel Prize Laureate in Chemistry and former President of Taiwan Academia Sinica.

#### About the Silicon Valley Chapter of IEEE Computer Society

The Silicon Valley Chapter of IEEE Computer Society is the largest chapter in the Silicon Valley Section of IEEE. The section itself is one of the largest, if not largest in the world. With over 3,600 subscribers in its mailing list, 1,300+ paid members, and a strong following of over 12,400 on Twitter alone, the chapter is uniquely positioned by virtue of its location and available expertise.

The chapter operates as part of the global <u>IEEE Computer Society</u>, the world's leading organization of computing professionals. IEEE CS was founded in 1946, and is the largest of the <u>IEEE's</u> 38 societies. The Computer Society is dedicated to advancing the theory and application of computing and information technology.

The Silicon Valley Chapter emphasizes all aspects of computing to the local members and welcomes visitors. It organizes monthly Technical Meetings, which give opportunity for professional networking and where invited speakers from academia and industry share their vision on the computing domain. The Chapter also puts on two conferences each year, ICADS (International Conference on Applied Data Science), and coorganizes the NFIC (New Frontiers in Computing) conference with NATEA (North America Taiwanese Engineering & Science Association) and the IEEE Student Chapter at Stanford University. The chapter started hosting an AMA (Ask me Anything) series featuring world renowned experts such as Pedro Domingos and Bjorn Schuller. The Chapter is a fully volunteer driven organization, where the volunteers benefit by developing their professional network, experience and satisfaction. Current Chapter Chair is Vishnu S. Pendyala, San Jose State University.

The chapter started its quarterly flagship publication, Feedforward and an awards program in 2022. For more details, please visit the chapter website at <u>https://r6.ieee.org/scv-cs/</u>.

## **Synopses and Biographies**

# 4:10 – 4:55 PM Introductions and Keynote: Active learning and explainable AI for medical imaging informatics – infectious disease outbreak

#### **Professor KC Santosh**

#### Synopsis:

When we consider AI for healthcare, infectious disease outbreaks are no exception. The talk will begin with machine learning models that help in not only predicting but also detecting abnormalities due to infectious diseases such as Pneumonia, TB, and Covid-19. Prof. KC Santosh will open my talk with infectious disease prediction models and unexploited data, where we will learn that predictive analytical tools are close to garbage-in garbage-out (at least for Covid19). He will then cover multimodal learning and representation based on both shallow learning (handcrafted features) as well as deep learning (deep features) that typically apply on medical imaging tools. Like in computer vision, Prof. KC will open an obvious question, how big data is big in addition to common techniques: data augmentation and transfer learning. With all these facts, as most of the models are limited to education and training, he will end the talk with the statement "ML innovation should not be limited to building models." What we need is #ExplainablableAI in #ActiveLearning framework.

#### Speaker:



Professor KC Santosh is the Chair of the Department of Computer Science (CS) at the University of South Dakota (USD). Prior to that, he worked as a research fellow at the U.S. National Library of Medicine (NLM), National Institutes of Health (NIH). He worked as a postdoctoral research scientist at the LORIA research center, Université de Lorraine in direct collaboration with industrial partner ITESOFT, France. He also served as a research scientist at the INRIA Nancy Grand Est research center (France), where he received his PhD in Computer Science – Artificial Intelligence. His research projects, primarily in Applied AI, are funded (of more than \$2m) by multiple agencies, such as SDCRGP, Department of Education, National Science Foundation, and Asian Office of Aerospace Research and Development. He has demonstrated expertise (with 10 books, 220+ research articles, and 20+ journal edited issues, as of Dec. 2021) in artificial intelligence, machine learning, pattern recognition, computer vision, image processing, and data mining with applications such as medical imaging informatics, document

imaging, biometrics, forensics, and speech analysis. He completed leadership and training programs for Deans/Chairs (organized by the Councils of Colleges of Arts & Sciences (U.S. 21)) and PELI – President's Executive Leadership Institute (USD 21). He is highly motivated/interested in academic leadership. To name a few, Prof. Santosh is the proud recipient of the Cutler Award for Teaching and Research Excellence (USD 2021), the President's Research Excellence Award (USD 2019) and the Ignite Award from the U.S. Department of Health & Human Services (HHS 2014).

# 4:55 – 5:40 PM Machine Learning for Misinformation Containment: A Candid Assessment of the State of the Art

#### Dr. Vishnu S. Pendyala

#### Synopsis:

Misinformation containment has been proven to be NP-hard more than a decade ago. It is undoubtedly a complex problem to solve and appropriately attracted plenty of attention from the research community. A wide variety of machine learning algorithms such as support vector machines and logistic regression, ensemble techniques like random forest and Adaboost, deep learning frameworks such as LSTM and GAN, language models like BOW / TF-IDF and BERT, and many more have been tried out in the attempts to solve the problem. In terms of feature engineering as well, no stone has been left unturned. Manual feature extraction, graph embeddings, and other approaches to representational learning have all been tried. Not just supervised and unsupervised learning, but various other types of learning such as few-shot learning, meta learning, transfer learning, self-supervised learning, semi-supervised learning, reinforcement learning, and active learning have been explored extensively for the problem. Despite the voluminous research literature purporting to solve the problem using machine learning methods, misinformation containment is largely unsolved and is in fact growing by the day. It is therefore pertinent to understand this huge disconnect between what is claimed in the literature and the actual reality. The talk will provide insights into the current state-of-the-art solutions and analyze why they are not helping enough. The talk will present some future directions that in the speaker's opinion hold the promise and explain why there is hope.

#### Speaker:



Dr. Vishnu S. Pendyala is a faculty member of the Department of Applied Data Science at San Jose State University and the chair of IEEE Computer Society, Silicon Valley Chapter. He has over two decades of experience with software industry leaders like Cisco and Synopsys in the Silicon Valley, USA. Dr. Pendyala served on the Board of Directors, Silicon Valley Engineering Council during 2018-2019. During his recent 3year term as an ACM Distinguished speaker and before that as a researcher and industry expert, he gave numerous (50+) invited talks. He holds MBA in Finance and PhD, MS, and BE degrees in Computer Engineering from US and Indian universities. Dr. Pendyala taught a one-week course sponsored by the Ministry of Human Resource Development (MHRD), Government of India, under the GIAN program in 2017 to Computer Science faculty from all over the country and delivered the keynote in a similar program sponsored by AICTE, Government of India in 2022. Dr. Pendyala's book, "Veracity of Big Data: Machine Learning and Other Approaches to Verifying Truthfulness" made it to several libraries, including those of MIT, Stanford, CMU, and internationally.

#### 5:40 – 6:25 PM Yolo4 and Its Applications

#### Dr. Mark Liao

#### Synopsis:

YOLOv4 has been ranked first in the world object detection competition for two and a half months. It defeated the R&D teams of international companies such as Google, Amazon, Facebook, Microsoft, and Qualcomm. The birth of YOLOv4 is actually closely related to the project "Development of Smart Transportation System". This project is funded by the Ministry of Science and Technology, and it has led to a four-year cooperative relationship between the Academia Sinica and the listed company Elan Electronics. This speech will explain in detail the beginning and end of the implementation of this smart transportation project, and how to develop YOLOv4, the fastest and most accurate object detector in the world during the execution of the project.

#### Speaker:



Dr. Mark Liao received his Ph.D. degree in electrical engineering from Northwestern University in 1990. In July 1991, he joined the Institute of Information Science, Academia Sinica, Taiwan, and currently is a Distinguished Research Fellow and Director. He has worked in the fields of multimedia information processing, computer vision, pattern recognition, multimedia protection, and artificial intelligence for more than 30 years. He was appointed an Honorary Chair Professor of National Chiao-Tung University from 2016 to 2019. He received the Young Investigators' Award from Academia Sinica in 1998; the Distinguished Research Award from the National Science Council in 2003, 2010, and 2013; the Academia Sinica Investigator Award in 2010; the TECO Award from the TECO Foundation in 2016, and the 64th Academic Award from the Ministry of Education in 2020. His professional activities include: President, Image Processing and Pattern Recognition Society of Taiwan (2006-08); Editorial Board Member, ACM Computing Surveys (2018 – present), IEEE Signal Processing Magazine (2010-13); Associate Editor, IEEE Transactions on Image Processing (2009-13), IEEE Transactions on Information Forensics and Security (2009-12) and IEEE Transactions on Multimedia (1998-2001). He has been a Fellow of the IEEE since 2013.

# 6:25 – 7:10 PM 3D Digital Pathology Inspired AI for Precision Diagnosis

#### Dr. Tun-Wen Pai and Dr. Yen-Yin Lin

#### Synopsis:

Digital medical images for training AI models have made a major impact on precision diagnosis, among which digital pathology transforming glass slide stains to whole slide images (WSIs) facilitates systematic analysis of tissue morphology and biomarker distribution. Integrating thick tissue staining, 3D image scanning, software and deep learning algorithms, we have retrieved hundreds more high resolution digital images with spatial features from each clinical biopsy to develop novel AI models for precision diagnosis of morphology variation, tumor recognition, and biomarker expression. Our breakthrough technologies will support matching the right patients to the right treatment, contributing to precision medicine.

#### Co- Author:



Dr. Tun-Wen Pai earned his Ph.D. in E&CE from Duke University, Durham, NC; MS in E&CF from John Hopkins University, Baltimore, MD USA. Presently, Dr. Pai is the Chairman of Computer Science and Information Engineering, National Taipei University of Technology, Taipei, Taiwan. Dr. Pai has published a number of papers which are available upon request.

#### Speaker:



Dr. Yen-Yin Lin earned his Ph.D. and MS degrees in Electrical Engineering, and BS degree in Nuclear Science from National Tsinghua University, Hsinchu, Taiwan. Presently, Dr. Lin is the CEO of JelloX biotech Inc., Hsinchu, Taiwan. His previous work experience includes: Chief Executive Officer, MOST Industrial Value Creation Program to develop AI and image inspection system for precision anatomy research, Brain research center at National Tsinghua University in Hsinchu, Taiwan; Visiting Research Fellow at Stanford University in California; R&D Consultant at Microtech Instruments in Oregon; Visiting Research Assistant at Brookhaven National Laboratory in New York. Dr. Lin's publications include 47 papers in prestigious journals, more than 80 conference papers (3 invited speeches in the first tier/+10000 participants conference, 1 worldwide webinar hosted by the Optical Society of America) and 16 issued USA or Taiwan patents. Dr. Lin received the following honors: 2022 Boehringer Ingelheim Grass Roots Award; 2021 The 20th Business Startup Award-MOEA in Taiwan; 2021 Gold Award-Entrepreneur Warrior Competition; Future Technology Award-MOST in Taiwan 2021 and others. In 2007 Dr. Lin received the outstanding Ph.D. dissertation award from the Optical Engineering Society of the Republic of China in Taiwan. He originated several novel laser systems, such as, novel electro-optic laser Q-switch (USA Patent No. US20110075688 A1), guasi-phase-matching PDT/PDD laser sources (TW Patent No.196412), cascaded quasi-phase-matching QPM laser source (TW Patent No.I225948).

### 7:10 – 7:45 PM Trends in Securing AI Data Pipeline

#### Prakash Ramchandran

#### Synopsis:

ML & DL are part of AI. ML models apply labels to data or collect samples dynamically and curate them. Data needs be clean and relevant to building or optimizing models and delivering the analysis and inferences that lead to improving the accuracy of results with greater optimization and making them more human-like intelligent yet applying MLOps, automation and efficient delivery for multiple domains and use cases. In the data pipeline from various sources they can be compromised and hence need to secure them and the speaker here will lead you through the causes and biases that can be controlled through use of best practices and share with you the trends in securing data pipelines in AI.

#### Speaker:



Prakash Ramchandran is leading Emerging Open Tech Foundation (eOTF) as Co-Founder and Secretary since 2020. The objectives of eOTF are helping the Indian subcontinent in its Digital and Operational Transformation. He is co-chair of INGR Edge Service working group focused on Next Generation Networking beyond 5G since 2019 from Silicon Valley. He has been a part of NFIC from IEEE CS SCV chapter and a constant promoter of IEEE & NATEA partnership for several years. With 40+ years of ICT industry experience in the US, EU, India and Asia-Pacific, he has led several startups and managed technically and innovated as ISP, ASP, CSP companies for decades. His insight into Data Analytics and AI is equally formidable and has several presentations in Open Source forums globally. He holds a Master's degree in EE from IIT Bombay.

# **Conference Planning Committee**

NFIC 2022 Planning Committee Members			
Gary Ni John Pang Yu Pat Fasang Rockwell Hsu Ray <mark>Sun</mark> Howard Ho	NATEA NATEA NATEA NATEA NATEA NATEA		
Vishnu S. Pendyala Prakash Ramchandran Kevin Cameron Paul Wesling John Delaney	IEEE-CS IEEE-CS IEEE - CS IEEE-CS IEEE-CS		
Pat Fasang Editor-in-Chief John Pang Yu Editor Vishnu Pendyala Editor	Program Book Program Book Program Book		
Paul Wesling	Webmaster		