Title: **Special Session on Reliable, Robust and Secure Machine Learning Algorithms**

Organizers:
- Monowar Bhuyan, Umeå University, Sweden
- Xuan-Son Vu, Umeå University, Sweden
- Shashi Gowda, CEO, Devr Inc., USA
- Erik Elmroth, Umeå University, Sweden
- Harry Nguyen, University College Cork, Ireland

PC Members:
- Florian Pokorny, KTH, Sweden
- Hasin Afzal Ahmed, Gauhati University, India
- Mukesh Prasad, UTS, Australia
- Bidyut Patra, NIT Rourkela, India
- Md Zakirul Alam Bhuiyan, Fordham University, USA
- Thanh-Son Nguyen, A*Star, Singapore
- Vu-Linh Nguyen, Eindhoven University of Technology, Netherlands
- Khuong Nguyen, FPT AI Center, Vietnam
- Alex To, University of Melbourne, Australia
- Anh Duong Trinh, TU Dublin, Ireland
- Yang Liu, HIT, China
- Thao Minh Le, Deakin University, Australia
- Debashis Das, IIT Jodhpur, India
- Sriparna Saha, IIT Patna, India
- Nabendu Chaki, University of Calcutta, India
- Alp Yurtsever, Umeå University, Sweden
- Mahshid Helali Moghadam, Scania, Sweden

Description:
The wider adoption of machine learning (ML) and artificial intelligence (AI) make several applications successful across societies such as healthcare, finance, robotics, transportation and industry operations by inducing intelligence in real-time [1-2]. Designing, developing and deploying reliable, robust, and secure ML algorithms are desirable for building trustworthy systems that offer trusted services to users with high-stakes decision-making [2-4]. For instance, AI-assisted robotic surgery, automated financial trading, autonomous driving and many more modern applications are vulnerable to concept drifts, dataset shifts, misspecifications, misconfiguration of model parameters, perturbations, and adversarial attacks beyond human or even machine comprehension level, thereby posing dangerous threats to various stakeholders at different levels. Moreover, building trustworthy AI systems requires lots of research efforts in addressing different mechanisms and approaches that could enhance user and public trust. To name a few, the following topics are known to be topics of interest in trustworthy and secure AI, but are not limited to: (i) bias and fairness, (ii) explainability, (iii) robust mitigation of adversarial attacks, (iv) improved privacy and security

in model building, (v) being decent, (vi) model attribution and (vii) scalability of the model under adversarial settings [1-5]. All of these topics are important and need to be addressed.

This special session aims to draw together state-of-the-art machine learning (ML) advances to address challenges for ensuring reliability, security and privacy in trustworthy systems. The challenges in different learning paradigms include, but are not limited to (i) robust learning, (ii) adversarial learning, (iii) stochastic, deterministic and non-deterministic learning, and (iv) secure and private learning. Nonetheless, all aspects of learning algorithms that can deal with reliable, robust and secure issues are the focus of the special session. It will focus on the robustness, performance guarantee, consistency, transparency and safety of AI, which is vital to ensure reliability. The special session will attract analytics experts from academics and industries to build trustworthy AI systems by developing and assessing theoretical and empirical methods, practical applications, and new ideas and identifying directions for future studies. Original contributions and comparative studies among different methods are welcome with an unbiased literature review.

Topics:

Topics of the special session include (reliable/robust/secure learning methods), including but not limited to:

- Robustness of machine learning/deep learning/reinforcement learning algorithms and trustworthy systems in general.
- Confidence, consistency, and uncertainty in model predictions for reliability beyond robustness.
- Transparent AI concepts in data collection, model development, deployment and explainability.
- Adversarial attacks - evasion, poisoning, extraction, inference, and hybrid.
- New solutions to make a system robust and secure to novel or potentially adversarial inputs; to handle model misspecification, corrupted training data, addressing concept drifts, dataset shifts, and missing/manipulated data instances.
- Theoretical and empirical analysis of reliable/robust/secure ML methods.
- Comparative studies with competing methods without reliable/robust certified properties.
- Applications of reliable/robust machine learning algorithms in domains such as healthcare, biomedical, finance, computer vision, natural language processing, big data, and all other relevant areas.
- Unique societal and legal challenges are facing reliability for trustworthy AI systems.
- Secure learning from data having high missing values, incompleteness, and noise
- Private learning from sensitive and protected data

Expected number of submissions:
60 articles

Website:
Under construction, we have been conducting this special session in ICONIP since 2021, this year as well https://www.reliableml.cs.umu.se/

Important dates:

- 10th June 2023 – Paper submission deadline
- 31st July 2023 – Paper acceptance notification
- 20th August 2023 - Deadline for final camera-ready submission
- 20th - 23rd November 2023 – Changsha, China

Why was this session needed?

We had a very successful special session last two years at the ICONIP conference on this topic and found its role is even more important because of several real-time systems coming up every day that work across societies. Our session was the most attractive special session in ICONIP 2021 based on the below statistics and participants' feedback. We had one invited talk every year. However, we received a bit lower response in last year version. Accordingly, we are planning to have two invited talks this year and make them popular and successful. As mentioned at the beginning, increased adoption of AI and ML need more to develop and deploy to ensure the trustworthiness of real-time systems at scale and precision.

Outcome and statistics - ICONIP 2021 special session

Number of submissions (it was diverse across topics)	**47**
Number of papers accepted (similar here when coming into acceptance)	**19**
Number of participation (great discussions among participants)	**40 (approx.)**

Short bio of the organizers:

Monowar Bhuyan (*Member, IEEE*) received his PhD in computer science and engineering from Tezpur University, Assam, India in 2014. He is currently an Assistant Professor at the Department of Computing Science, Umeå University, Sweden and one of the research group leaders at the Autonomous Distributed Systems Lab. Before this, he worked with the Nara Institute of Science and Technology, Japan, Umeå University, Assam Kaziranga University, India, and Tezpur University, India, from January 2009 to December 2019. He has published over sixty papers in the leading international journals and conference proceedings and has written a book with Springer. His experience leading/co-leading research projects is over 20 MSEK, including national and EU grants. His research areas include machine learning, anomaly detection, security and privacy, and distributed systems.

Xuan-Son Vu (*Member, IEEE*) received a B.S. degree in information systems from the University of Engineering and Technology, VNU, Hanoi, Vietnam in 2011 and an M.Sc. degree from Kyungpook National University, South Korea, in 2014. He earned his PhD. in computer science from Umeå University, Sweden in 2020 with a focus on privacy-preserving machine learning with big data. Before joining Umeå University, from 2015 to 2016, he was a full-time member of UKPLab, TU Darmstadt, Germany. He is currently a *postdoctoral fellow* at Umeå University, Sweden, where he focuses on research in robust machine learning. His research interest has been primarily focused on knowledge - both acquiring knowledge from

multimodal data and using structured knowledge to power downstream applications. For the former, he has worked on ontology, information retrieval and natural language processing. For the latter, multimodal knowledge learning is one of the topics he enjoys researching the most. Dr Xuan-Son Vu's awards and honours include the 3rd place for best paper awards at CICLING 2019, best student paper award and best inquisitive mind award at CICLING 2018. He also received the best paper award at the 40th Conference of the Korea Information Processing Society in 2014.

Shashi Gowda is CEO and co-founder of Devr and a member of its board of directors. Shashi earned his bachelor's degree in bio-medical engineering from Boston University. A seasoned technology leader and entrepreneur, Shashi brings over 20 years of experience in the telecommunications, big data and emerging tech industries. Devr provides tools for enterprises to design and orchestrate data privacy, enabling rich open ecosystems for data monetization with continuous compliance. Shashi champions advances in digital transformation, data privacy and continuous compliance using emerging technologies such as Blockchain and AI. He is passionate about bringing technologies into highly regulated industries which enable them to innovate with data, in an era of growing privacy concerns, regulatory complexity, and high costs of non-compliance. Before founding Devr, Shashi was CTO of Virtual Control and supported the sale of its AI and ML solutions to Agilent Technologies. Prior to that, Shashi held various technology design and leadership roles, at both corporates and at early-stage companies, and has worked across Asia, Europe, and the US.

Harry D. Nguyen is an Assistant Professor in the School of Computing Science, University College Cork, Ireland. He holds a PhD in Information Systems and Analytics from the National University of Singapore. His research interests include health optimisation, big data analytics, deep learning and mobile human-computer interaction. His work has been published in the International Conference on Information Systems (ICIS), Communications of the Association for Information Systems, Journal of Decision Systems, and Health and Quality of Life Outcomes. He regularly serves as a program board member and/or reviewer for international conferences and journals, including the International Conference on Human-Computer Interaction (HCII), Design Science Research in Information Systems and Technology (DESRIST) and Health Systems.

Erik Elmroth (*Member, IEEE***)** is a Full Professor in Computing Science at Umeå University, Sweden. He has been Head and Deputy Head of the Department of Computing Science for 13 years and deputy director for a national supercomputer center for another 13 years. He has established the Umeå University research on distributed systems, see http://www.cloudresearch.org. His experience from management and executive groups in large-scale research projects includes highlights such as the 550M EURO Wallenberg AI, Autonomous Systems and Software Program and the Strategic Research Area eSSENCE. He has been a member of the Swedish Research Council's committee for research infrastructure and chair of its expert panel on eScience as well as Chair of the Board of the Swedish National Infrastructure for Computing. He has developed two international research strategies for the Nordic Council of Ministers. International experiences include a year at NERSC, Lawrence Berkeley National Laboratory, University of California, Berkeley, and one semester at the Massachusetts Institute of Technology (MIT), Cambridge, MA. Elmroth is a

lifetime member of the Swedish Royal Academy of Engineering Sciences and vice-chair of its division for Information Technology.

Other Information:
Papers submitted to this Special Session are reviewed according to the same rules as the submissions to the regular sessions of ICONIP 2023. Authors who submit papers to this session are invited to mention them in the form during the submission. Submissions to regular and special sessions follow identical format, instructions, deadlines and procedures of the other papers. Click here for information on paper submission, please, for further information and news, refer to the ICONIP's website: http://iconip2023.org/index.html

References:

1. Xiong, Pulei, et al. "Towards a robust and trustworthy machine learning system development: An engineering perspective." *Journal of Information Security and Applications*, Vol. 65, pp. 103121, 2022.
2. Barmer, Hollen; Dzombak, Rachel; Gaston, Matthew; Heim, Eric; Palat, Vijaykumar; Redner, Frank; et al. (2021): Robust and Secure AI. *Carnegie Mellon University. Report*. https://doi.org/10.1184/R1/16560252.v1
3. Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C., ... & Song, D. (2018). Robust physical-world attacks on deep learning visual classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* pp. 1625-1634, 2018.
4. M. Shafique *et al*., "Robust Machine Learning Systems: Challenges, Current Trends, Perspectives, and the Road Ahead," in *IEEE Design & Test*, vol. 37, no. 2, pp. 30-57, April 2020, doi: 10.1109/MDAT.2020.2971217.
5. Yang, Yuting, et al. "Quantifying Robustness to Adversarial Word Substitutions." *arXiv preprint arXiv:2201.03829,* 2022.
6. Balcan, Maria-Florina, et al. "Robustly-reliable learners under poisoning attacks." *Conference on Learning Theory*. PMLR, 2022.