Federal University of Santa Catarina
Software/Hardware Integration Lab

Towards Trustworthy Generative AI

Fredrik Heintz

Call for Participation

Date: Nov 13, 2024   Time: 14:00   Place: Room 105, INE, CTC, UFSC

AI development is fast. Especially generative AI has created a lot of excitement and is expected to have a major impact on our economy and society. At the same time we want AI to be trustworthy. This talk will discuss some of the challenges and opportunities related to making GenAI trustworthy. Special emphasis will be placed on the need to integrate reasoning and learning to combine the power of data-driven methods with the formal guarantees and general solutions provided by reasoning-based approaches. One example is Neurosymbolic AI. The successful realization of Trustworthy GenAI will be paramount for addressing the major challenges we as a global society is facing. The talk is partly based on the ongoing work in the Horizon Europe project TrustLLM, which has the goal of developing more factual and trustworthy large language models, and the completed Horizon Europe ICT-48 network of research excellence TAILOR which had the goal of developing the scientific foundations for Trustworthy AI through integrating learning, optimization and reasoning.

Fredrik Heintz is a Professor of Computer Science at Linköping University, where he leads the Division of Artificial Intelligence and Integrated Computer Systems (AIICS) and the Reasoning and Learning lab (ReaL). His research focus is artificial intelligence especially Trustworthy AI and the intersection between machine reasoning and machine learning. Director of the Wallenberg AI and Transformative Technologies Education Development Program (WASP-ED), Director of the Graduate School of the Wallenberg AI, Autonomous Systems and Software Program (WASP), Coordinator of the TrustLLM project, and Vice President for AI Research Adra the AI, Data, and Robotics partnership. Member of the Swedish AI Commission. Fellow of the Royal Swedish Academy of Engineering Sciences (IVA).

Phone: +55 48 3721-9516   e-mail: lisha@lisha.ufsc.br   Web: https://lisha.ufsc.br/