Title: Beyond Technicalities: Positive Normative Choices and the Ethical Dimensions of Responsible AI System Design

Abstract: Artificial Intelligence (AI) systems are increasingly permeating various sectors of our lives, from healthcare to finance, transportation to education. As these systems become more pervasive, the imperative for their responsible and trustworthy design and deployment escalates.

However, AI development is fraught with complexity. There is a vast array of quality measures for AI, including accuracy, robustness, explainability, and fairness. The relationships between these measures are intricate, often involving trade-offs. For instance, increasing a model's accuracy might reduce its explainability, or enhancing its robustness might compromise its fairness. Furthermore, there is no universal hierarchy of these quality measures. The relative importance of different measures can vary depending on the specific context in which the AI model is used. This complexity underscores the need for a nuanced, holistic, pluralistic, and context-sensitive understanding of AI model quality.

Motivated by this observation, this talk will delve into the concept of "positive normative choices" in the context of AI, a term coined by Wachter et al. (2021). 'Positive normative choices' refer to explicit, considered, and justified decisions integral to AI model development and deployment. These decisions are multifaceted, encompassing various dimensions of quality measures. A 'positive normative choice' can be a decision regarding a specific design choice within a quality measure, or a decision to prioritize one quality measure over another. These choices are not merely technical but also ethical, reflecting a commitment to certain values and principles. Thus, 'positive normative choices' represent a proactive stance towards shaping the ethical landscape of AI systems, ensuring alignment with societal values and norms.

Given that the justification for these decisions is arguably context-dependent, responsible and trustworthy AI design becomes a highly nuanced and context-specific task. This underscores the need for a comprehensive understanding of the broader social and decision-making contexts in which AI models operate.

In this talk, we will explore how these choices can be made ethically, considering the broader social and decision-making contexts in which AI models are embedded. We will advocate for a truly interdisciplinary and integrative approach to ensure that AI models do not unduly contribute to harmful outcomes or exacerbate existing inequalities. We will advocate for a holistic assessment of AI models, considering not only their technical performance but also their societal impact within larger societal and decision contexts.

By exploring these topics, this talk aims to deepen our understanding of the ethical dimensions and societal implications of AI. It seeks to foster the development of best practices for creating responsible, trustworthy, and beneficial AI technologies. Through this lens, we will address the challenges posed by AI in real-world applications, focusing on ethical considerations but acknowledging the interconnected legal, psychological, economic, and societal dimensions. This talk is a call to action for AI practitioners and stakeholders to take the "positive normative choices" they have to make seriously, thereby ensuring the responsible evolution of AI technologies.