# Designing Human-AI Interaction in Vehicles during Safety-Critical Events

BIOGRAPHY

Cherin Lim is a Ph.D. student in the department of Industrial and Systems Engineering at the University of Washington, advised by Prof. Prashanth Rajivan. Her research focuses on designing human-AI interaction in autonomous vehicles (AVs), specifically during safety-critical events, to support passengers through appropriate response. She applies methods such as behavioral experiments, statistical analyses, and machine learning techniques to understand how trust and risk perception towards AVs affect passenger behavior and address current interaction challenges. Cherin received the College of Engineering Dean's Fellowship and served as the vice president of the UW HFES student chapter during 2024-2025.
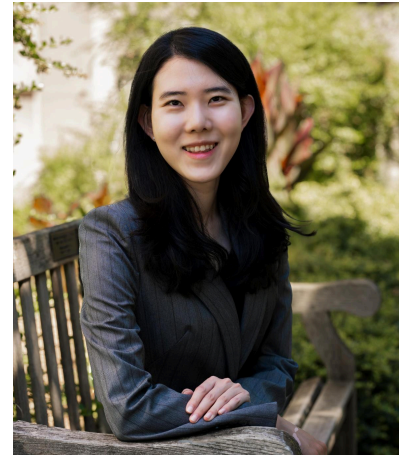
ABSTRACT

While AVs promise to facilitate the driving task and improve human safety, people often find it difficult to fully understand or trust AVs. Unlike traditional vehicles with human drivers, AVs are absent of this human presence, introducing a key interaction gap between the vehicle and passengers. This lack of interaction becomes particularly profound in safety-critical events, with passengers being left without clear explanations or reassurance, which can create uncertainty and further degrade trust. Communication and adaptive responses are thus significant in shaping how passengers perceive risk and calibrate their trust towards AVs during unexpected events. With the development of generative AI, potential to design human interaction with AVs that are context-sensitive, risk-adaptive, and emotion-aware exists. However, it remains unclear what information passengers need in safety-critical events, when it should be provided, and through what interaction mechanisms it must be provided to calibrate appropriate levels of trust towards AVs and enable the appropriate use of autonomous capabilities. This talk features four interconnected studies that address these challenges. Through an interview and driving simulator experiment, we find that trust should be understood as a dynamic construct and that the driving experience itself is significant in trust repair after an AV failure, beyond system transparency. A subsequent questionnaire study reveals that AV systems should be designed to be adaptive to risk. Lastly, social media data analysis finds that AVs should be capable of providing emotion-aware communication to passengers in unsafe situations. Overall, this line of research aims to inform the design of AV systems that can actively assume the role of a human driver, beyond technically functioning as a means of transportation.