

## Cover Letter

Dr. Fazelpour,

Hi! I am a PhD student (taking my candidacy exam this semester) working on the challenges that arise in AI when moving from the sandbox to the real world, primarily at the intersection of fairness, reliability, privacy, and safety, focusing on uncovering the learning dynamics of ML/AI systems. More recently, my research has focused on the socio-technical evaluation of generative AI systems and uncovering sources of variance, biases, and limitations of existing assessments and metrics. My research closely aligns with the core themes of this position - social impact evaluations, fairness, and harms - which makes me genuinely excited about the position.

My research lies between theoretical and empirical inquiry, drawing from fairness analysis, conceptual analysis, large-scale AI evaluation, and policy and governance. My research has revolved around diverse methods, including qualitative and quantitative studies, systematic reviews, and human studies. At its core, my work focuses on empirical evaluation and simulation studies aimed at operationalizing fairness in AI/ML systems, minimizing the gap between the technical feasibility of approaches and social impact. An example of my diverse research is the novel application of fairness principles to drone authorizations in U.S. airspace, where I am collaborating closely with NASA and the FAA.

Through this opportunity, I am interested in applying software engineering and safety engineering principles to develop systematic approaches to evaluating the social impact of generative AI systems. I am particularly interested in drawing from defeasible reasoning used in software assurance cases for generative AI evaluation, parallel to red teaming. Moreover, I would like to evaluate intersectional harms in generative AI systems and new safety risks/harms, drawing from my previous work on intersectional fairness. Lastly, I am also interested in analyzing the shortcomings of current evaluation metrics (including uncertainty) and assessments.

I am very excited about collaborating and would love the opportunity to discuss how my background and interests align with the visions for this summer program. Thank you!