Expression of Interest

I am a final-year PhD candidate in Computer Science at the University of Massachusetts Amherst. My research focuses on responsible and human-centered AI, with particular attention to how marginalized groups perceive bias in AI systems and how that affects their trust. More recently, I have also begun working in accessibility, exploring what it means for disabled people to move beyond “disability-centered” design toward having the tools and power to shape their own worlds.

To examine how bias in machine learning models shapes user trust, I conducted a behavioral study with more than 1,500 participants, adapting a trust game from behavioral economics where people chose between “robot investors” with different historical returns [1]. One model was fair, while the other displayed gender bias (favoring men or women). We found that women were more likely than men to choose the fair model, regardless of whether the bias advantaged them; and that participants were more sensitive to bias when it was presented in text rather than bar charts. However, this study necessarily treated gender in binary terms, since we had too few non-binary or gender-diverse participants to analyze separately. This limitation motivated my subsequent work on large language models, where I explicitly centered the perspectives of gender-diverse communities. In this follow-up study, I examined how gendered and neutral prompts influence the responses generated by ChatGPT and how users from different gender backgrounds perceive those responses [2]. From our interviews, women, non-binary, and transgender participants noted that outputs often reinforced stereotypes, defaulted to binary categories, or failed to represent their identities accurately. For example, women described responses that reproduced traditional gender roles, while non-binary and trans participants highlighted misgendering or the erasure of non-binary identities. Taken together, this work highlights how gender influences perceptions of bias and trust in AI, and why evaluations must account for experiences beyond binary categories.

In parallel, I am conducting a study with blind individuals that adopts a blind-centered self-design approach, positioning participants as innovators and designers [3]. By mailing physical craft materials to aid with prototyping and supporting community-led creation of assistive concepts, this work surfaces design needs often absent from mainstream AI and accessibility research, such as multisensory interaction and user agency. It hopes to demonstrate how disabled communities can actively shape technology when they have resources and decision-making power, pointing toward more inclusive forms of participation in AI development.

I see this workshop as an opportunity to extend these insights into the domain of large-scale dataset practices. I am especially interested in how communities might play a more meaningful role in shaping the data that underpins foundation models, whether through participatory approaches to curation, more transparent processes of consent, or new models of governance. The workshop's interdisciplinary focus is also appealing, as questions of responsibility in dataset curation cannot be resolved by technical fixes alone. By joining, I hope to contribute perspectives from my empirical work with gender-diverse and disabled participants while learning from others working on data governance, law, and large-scale infrastructures. My goal is to leave with new connections and conceptual tools that will help shape future research for more equitable, and inclusive dataset curation practice.

References:

1. **Gaba et al.**, "My Model is Unfair, Do People Even Care? Visual Design Affects Trust and Perceived Bias in Machine Learning," in IEEE Transactions on Visualization and Computer Graphics, vol. 30, no. 1, pp. 327-337, Jan. 2024, doi: 10.1109/TVCG.2023.3327192.
2. **Gaba et al.**, "Accuracy, Bias, and Trust: Gender Diverse Perspective on Large Language Models", (Under Review, CSCW 2026)
3. **Gaba et al.**, "Beyond "Disability-centered:" Empowering Blind-led Innovation of Assistive Technology", (Under Preparation, CHI 2026)