Title : Ethical AI in Citizen Science: Exploring Human-Agent Collaboration for Fact-checking

- **Sanjana Gautam, University of Texas at Austin (School of Information)**

My research into social media analytics, which involves the use of AI for real-time crisis response, can inform approaches to leveraging AI assistants in computational citizen science. Tools like PIVOT, which process large-scale social media data for crisis managers, offer insights into how AI systems can be used to assist volunteers and participants in data collection and interpretation in citizen science projects. PIVOT is a sophisticated application designed to address the challenge of effectively using the vast amount of social media information available to crisis managers and individuals in real-time scenarios. We used this platform for generating real-time alerts for dispatch centers. The usability study I conducted demonstrated the current design gaps in designing for crisis scenarios. Furthermore, to better understand the workings of 911 dispatch centers during the pandemic, I conducted an interview study that was published in the Journal of Emergency Management 2021, as well as ISCRAM 2022. A statistical study conducted as a follow-up towards understanding trust in social media platforms like Twitter was undertaken next. Our study showed the daily increase in cases with the recorded number of fake news and misinformation spread.

Misinformation poses significant threats to society, including undermining trust in science, distorting public health narratives, deepening social polarization, and disrupting democratic processes. In response, professional fact-checkers and journalists have developed high-quality investigations to verify purported facts, but these manual efforts are overwhelmed by the sheer scale of misinformation. To address this challenge, Natural Language Processing (NLP) technologies have been proposed to automate and scale fact-checking efforts. Despite rapid advancements, practical adoption of NLP tools by fact-checkers remains limited.

Another study (that I conducted) explored TikTok's role in shaping political awareness, particularly among younger users. By deploying a browser extension that collected videos and user annotations, the research found that an average of 16.36% of content in users' feeds is political, with users who have stronger political inclinations encountering even more (up to 21%). Common political topics include police, crime, racism, homelessness, and income disparities. The findings provide insights into TikTok's impact on political discourse and engagement from a user-centered perspective.

This further highlights the importance of identifying misinformation on social media platforms. My research not only contributes to advancing responsible AI for citizen science but also highlights the importance of fostering trust, collaboration, and ethical practices in human-AI teaming. The need for early involvement of fact-checkers in the design process, ensuring that AI tools are tailored to meet their needs and can be seamlessly integrated into real-world workflows. This aligns with the *Human-computer Teaming* and *Citizen Science Trust, Equity, Ethics, and Responsible AI* focus areas by emphasizing transparency, trust, and ethical AI deployment. Additionally, by addressing biases in language models that may inadvertently propagate misinformation, my research seeks to mitigate harmful consequences associated with AI-generated content. I propose that these adaptive and socially-conscious AI systems can help broaden participation in computational citizen science by reducing barriers and enabling more inclusive engagement. Through this work, I aim to foster a collaborative environment where AI systems

enhance human capabilities, particularly in combating misinformation in high-stakes contexts, such as public health and political discourse. Ultimately, I advocate for developing benchmarks that prioritize the practical utility of AI tools in fact-checking and misinformation mitigation, ensuring they contribute to the creation of a more informed and resilient society.