Matthew Lease (The University of Texas at Austin) - October 2024

Overview. My work spans AI modeling and human-computer interaction (HCI) design. I create annotated datasets, build AI models, and evaluate both model performance and end-user impacts. When automated AI falls short, I design human-in-the-loop approaches supported by interpretable AI models and creative user interfaces. In this area, I seek effective human-AI partnerships that capitalize on the respective strengths of each party, blending front-end HCI with back-end AI modeling of people and their tasks. Much of my annotation and modeling work has been situated in crowdsourcing and human computation (HCOMP), including the study of data workers and the design of worker-centered approaches. In HCOMP, front-end HCI design is necessary for workers to understand and complete tasks, while back-end AI modeling of workers and tasks can optimize resulting efficiency and quality. I work primarily with human language, e.g., natural language processing (NLP), though I have also worked with imagery.

I conduct fundamental research driven by use-inspired, real-world problems. Two examples:

Good Systems. My responsible AI research in explainability and fairness has been grounded in a 6-year, 7-faculty project I lead at UT Austin that seeks to develop better AI tooling to support professional fact-checkers. This project in turn is embedded in Good Systems, UT Austin's 8-year Grand Challenge to develop responsible AI technologies, with over 120 researchers engaged across 30 academic units. We believe that responsible AI should be driven as use-inspired research and grounded in addressing societal challenges. In turn, we expect these societal challenges to drive important, fundamental advances in general AI capabilities.

CosmicAI. I am Co-Director of the new NSF-Simons AI Institute for Cosmic Origins (CosmicAI). NSF AI Institutes represent a cornerstone federal government commitment to supporting long-term, fundamental AI research. By focusing AI research on key scientific domains — in our case, Astronomy — AI Institutes harness domain challenges to drive new, fundamental advances in AI capabilities, which in turn enable new advances in each scientific domain. I am particularly excited by CosmicAI's focus on developing new, generative AI capabilities to accelerate scientific discovery. I further discuss related future work below/

Past work: crowdsourcing. Data annotation is central to training and testing supervised learning systems. Fifteen years ago, the advent of crowdsourcing created an opportunity to massively scale-up annotation but also posed new risks to quality control. On-demand internet workers also created new opportunities for human-in-the-loop system design, engaging workers at run-time to perform tasks where AI fell short. Finally, new uses of human labor (often invisible) to train and supplement AI systems raised new ethical questions about treatment of the workers powering our AI advances. In these areas, I developed new techniques for task design, task allocation, annotator agreement, label aggregation, and human-AI teaming. I may be best known for my work on aggregating labels across multiple annotators to induce consensus annotations (e.g., I will receive an HCOMP Test of Time award this month for such work). My lab's more recent aggregation and annotator agreement work has developed the first general model we know of for "complex" annotation tasks (having large output spaces). While many aggregation models exist for categorical labeling tasks (with a few answers to choose from), complex tasks have vast or unbounded response spaces (e.g., translating or

summarizing text, or segmenting objects in images). Prior voting-based aggregation methods do not work when every annotator provides a different answer, requiring a custom probabilistic model or human reconciliation workflow for each different annotation task. In contrast, we developed the first task-agnostic approach that supports both label aggregation as well as measuring annotator agreement over such unbounded response spaces.

Current work: humans & generative AI. With the advent of large language models (LLMs), the frontier of AI capabilities has significantly advanced. Because we often pursue human-in-the-loop approaches when AI capabilities fall short, these greater AI capabilities have shifted the former balance and nature of work distribution between human and AI. On one hand, we can automate more today than we could before, including a variety of data labeling, human computation, and evaluation tasks that previously required manual work by people (e.g., crowdworkers or citizen scientists). From the worker perspective, when they are assigned tasks, they may also utilize LLM tools in performing the work, automating some or all of their work. As with the original rise of crowdsourcing, this creates new opportunities for efficiency gains and intriguing new possibility for human-AI configurations, but also introduces new risks, e.g., perpetuating AI errors and biases into training data, circularity of AI (rather than people) training AI, and different concerns around potential fraud and abuse by workers. Early products of my work in this area can be seen in publications at CIST 2024, EMNLP 2024, and IUI 2024.

Future work: **accelerating scientific discovery**. As mentioned above, a key goal of our new CosmicAI Institute will be to accelerate scientific discovery. This includes a new LLM co-pilot for astronomy to enhance and accelerate their work. One aspect of this is simply to automate as much rote work as possible in science just as AI co-pilots are automating such work in industry. Research questions here include: What scientific work can be delegated to an AI co-pilot vs. what must the scientist continue to do manually? Where should AI and scientist work together, and what human-AI configurations are most suitable? Given these new AI-scientist workflows, where do amateur human contributors (e.g., crowd workers or citizen scientists) fit into today's modern scientific workflows harnessing generative AI?

We did not propose citizen science for CosmicAI, focusing only on automated AI. However, this neglects my interest and experience in crowdsourcing, as well as important work by others in Galaxy Zoo (and then Zooniverse) in exploring how citizen scientists can advance science (and astronomy in particular). For this reason, the Convergence of Computational and Citizen Science Research Workshop presents an ideal opportunity to explore such ideas for CosmicAI with other researchers similarly interested in crowdsourcing and citizen science. Given LLMs, the theme of *Convergence of Computational & Citizen Science Research* is particularly apt.