

Thirdly, the various elements of our event, including the expert presentations, informative posters, and notably the winners of our two leader-board, are set to offer a substantive yet succinct overview of the latest and significant trends, research directions, and innovative concepts in the realm of multimodal agents. These presentations will encapsulate pivotal findings and developments, shining a light on new systems, ideas, and technologies in the field of multimodal agent AI. This assortment of knowledge is not only beneficial for the attendees of our forum, who are looking to deepen their understanding and expertise in this domain, but it also serves as a dynamic and rich resource board. Those visiting our forum’s website can tap into this reservoir of information to discover and understand the cutting-edge advancements and creative ideas steering the future of multimodal agent AI. We strive to serve as a useful knowledge base for both newcomers and veterans in the field. By engaging with these resources, we hope participants and online visitors alike can remain informed of the transformative changes and novel approaches that are shaping the exciting landscape surrounding multimodal agent AI.

11 Ethical Considerations

Multimodal Agent AI systems have many applications. In addition to interactive AI, grounded multimodal models could help drive content generation for bots and AI agents, and assist in productivity applications, helping to re-play, paraphrase, action prediction or synthesize 3D or 2D scenario. Fundamental advances in agent AI help contribute towards these goals and many would benefit from a greater understanding of how to model embodied and empathetic in a simulate reality or a real world. Arguably many of these applications could have positive benefits.

However, this technology could also be used by bad actors. Agent AI systems that generate content can be used to manipulate or deceive people. Therefore, it is very important that this technology is developed in accordance with responsible AI guidelines. For example, explicitly communicating to users that content is generated by an AI system and providing the user with controls in order to customize such a system. It is possible the Agent AI could be used to develop new methods to detect manipulative content - partly because it is rich with hallucination performance of large foundation model - and thus help address another real world problem.

For examples, 1) in health topic, ethical deployment of LLM and VLM agents, especially in sensitive domains like healthcare, is paramount. AI agents trained on biased data could potentially worsen health disparities by providing inaccurate diagnoses for underrepresented groups. Moreover, the handling of sensitive patient data by AI agents raises significant privacy and confidentiality concerns. 2) In the gaming industry, AI agents could transform the role of developers, shifting their focus from scripting non-player characters to refining agent learning processes. Similarly, adaptive robotic systems could redefine manufacturing roles, necessitating new skill sets rather than replacing human workers. Navigating these transitions responsibly is vital to minimize potential socio-economic disruptions.

Furthermore, the agent AI focuses on learning collaboration policy in simulation and there is some risk if directly applying the policy to the real world due to the distribution shift. Robust testing and continual safety monitoring mechanisms should be put in place to minimize risks of unpredictable behaviors in real-world scenarios. Our “VideoAnalytica” dataset is collected from the Internet and considering which is not a fully representative source, so we already go through-ed the ethical review and legal process from both Microsoft and University Washington. Be that as it may, we also need to understand biases that might exist in this corpus. Data distributions can be characterized in many ways. In this workshop, we have captured how the agent level distribution in our dataset is different from other existing datasets. However, there is much more than could be included in a single dataset or workshop. We would argue that there is a need for more approaches or discussion linked to real tasks or topics and that by making these data or system available.

We will dedicate a segment of our project to discussing these ethical issues, exploring potential mitigation strategies, and deploying a responsible multi-modal AI agent. We hope to help more researchers answer these questions together via this paper.

12 Diversity Statement

By examining the adaptability of AI agent models in various domains, we inherently embrace a diversity of leader-board, perspectives, and solutions. In this vein, our project aims to build a diverse community by exploring the wide array of subjects in multimodal and agentic AI.