First International MASST Initiative Workshop

Jeffrey M. Bradshaw, Institute for Human and Machine Cognition (IHMC), Pensacola, FL, USA
Mufti Mahmud, King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia

In Conjunction with the IEEE/WIC International Conference
on Web Intelligence and Intelligent Agent Technology
14–15 November 2025, London, United Kingdom

CALL FOR PARTICIPATION

Though "agentic AI" is understood in many ways (sometimes including single agent systems), this workshop is specifically focused on multi-agent systems (MAS), a longstanding area of research and development that has been re-invigorated by recent advances in AI. Multi-agent systems are designed to unleash Artificial Intelligence from the solitary confinement of chatbot applications so they can freely—and safely—act in the wider world as components of a network of intelligent AIs. *This workshop will welcome contributions relating to safety and teamwork challenges that either apply specifically to MAS or else to the variety of AI and other technologies that could be used as components in the creation of multi-agent systems.*

Agents functioning as part of a MAS will typically share common features and standard protocols, though their range of capabilities will vary according to who they are representing (e.g., people, robots, organizations, roles) and the kind of task they are performing (e.g., AI modeling, statistical analytics, workflow coordination, service discovery). By this approach, MAS facilitates the use of heterogeneous collections of computational methods—including neuro-symbolic reasoning and other "composite AI" methodologies. Rather than relying on a single vendor to supply AI capabilities, open standards undergirding the MAS approach are intended to foster a burgeoning marketplace of ideas and implementations.

Because MAS frameworks will be increasingly capable of understanding and considering both the personal preferences of actors and the regulatory policies of authoritative bodies, well-designed multi-agent systems will be ideally suited to cut through the inefficiencies of bureaucratic baggage endemic to modern life with increasing reliability and assurance. By loosely knitting people and services across social and technical boundaries, MAS can be seen as creating "virtual organizations" that streamline action while firmly respecting security, privacy, and resource constraints. Thus, when thoughtfully implemented in the context of broader socio-technical initiatives, multi-agent systems will also be able to function as a catalyst to organizational redesign and process improvement.

The international MASST initiative is focused on discovering and promoting scientific and engineering principles and technologies to enable safer and more effective collaboration within multi-agent systems serving hybrid human-agent collectives of all sizes and forms. Examples include, but are not limited to, the following:

- Designing for safety from the ground up rather than principally through reinforcement learning from human feedback (RLHF)
- Sophisticated, context-aware behavioral guard rails
- Representation and reasoning about ethical theories and principles
- Design-time and run-time risk-mitigation techniques
- Principles and technologies for safe agent mobility
- Learning approaches that foster an appropriate balance of consistency and responsiveness to externally defined constraints and changing contexts
- Methods to help assure mutual observability, predictability, and directability that allow humans and agents to develop appropriately calibrated trust relationships with each trusted party
- Principles and practices to assure effective Human-Agent Robot Teamwork (HART)
- Case studies of safety and teamwork (e.g., AI in medicine, self-driving vehicles, military and space applications, drones and internet of things in public safety settings)

Note that this workshop will not address issues of confidentiality, privacy, authentication, or agent credentialling. These are important topics, but solutions to this problem in MAS resemble those addressed in traditional distributed software services and likely can be addressed by general software advances already in progress. And while we think that developing principles and recommendations for government and industry initiatives on AI governance is important, many official bodies doing such work are already in existence.

Submission Information:

Full paper submission to the IEEE Conference Proceedings are *not* being requested. Instead, we invite the submission of *presentation abstracts* on topics of relevance to the topic of multi-agent system safety and teamwork. Because a variety of specific AI technologies and techniques can be used as elements of multi-agent systems, abstracts on safety and teamwork challenges of component technologies are also welcome. Each presentation abstract should be no longer than two pages and must contain the following information:

- Title
- Submitter names, affiliations, contact information, and online links to biographical information
- Descriptions of significant risks and forecasts related to the topic of the presentation
- Descriptions of relevant technical approaches and technology contributions proposed to improve safety and teamwork generally or with reference to specific application domains
- Online references to related work, when possible
- Whether you will be able to present in person (generally preferred) or will present online.

Selected abstracts will be featured as presentations at the workshop. Submitters whose abstracts are not selected are also welcome to attend the workshop in person as space allows. Anyone is welcome to attend the workshop online.

Publication:

Full papers based on the workshop and other invited contributions will be published as a book as part of a series of printed and digital MASST volumes. Selected contributions may also be featured in a special journal issue.

Important Dates:

- Presentation proposal abstracts: 15 October 2025. Send directly to Jeff at jbradshaw@ihmc.org
- Presentation acceptance feedback: 31 October 2025
- Final presentations in PowerPoint or Keynote format: 7 November 2026