



Prof. Jeffrey M. Bradshaw

Senior Research Scientist at IHMC

Institute for Human and Machine Cognition (IHMC)

Jeffrey M. Bradshaw (PhD, Cognitive Science, University of Washington) is a Senior Research Scientist at the Institute for Human and Machine Cognition (IHMC). He helped pioneer the research areas of multi-agent systems and human-agent-robot teamwork, and his first book on the topic, *Software Agents*, became a classic in the field and a best-seller for The MIT Press. He is an emeritus member of the board for the International Foundation for Autonomous Agents and Multi-Agent Systems. In 2025, he inaugurated the international Multi-Agent System Safety and Teamwork initiative (MASST). Jeff was appointed an Honorary Visiting Researcher at the University of Edinburgh and a Visiting Professor at the Institute de Cognitique at the University of Bordeaux, following earlier research in France as a senior Fulbright scholar. Jeff chaired the Scientific Advisory Council for the Nissan Research Center—Silicon Valley autonomous vehicle program and for many years was a co-editor of the Human-Centered Computing Department for IEEE Intelligent Systems. Jeff was a member of the US Defense Science Board 2015 Study on Autonomy, the Board on Global Science and Technology for the National Academies of Science, and the National Research Council Committee on Emerging Cognitive Neuroscience Research. He is a former chair of ACM SIGART (now SIGAI) and the RIACS Science Council for NASA Ames Research Center. He was a scientific advisor to the Cognitive Science Program at Sandia National Laboratories, the HCI and Visualization program at the German Research Center for Artificial Intelligence (DFKI), and to the Japanese NEC Technology Paradigm Shifts initiative. He received the Web Intelligence Consortium Outstanding Contributions Award.

Trust, Policy, and Teamwork: Missing Layers in Standards for Multi-Agent System Interoperability

Multi-Agent Systems: Trust, Policy & Teamwork

Multi-Agent Systems (MAS) enable AI agents to collaborate across people, tools, cloud platforms, and devices to perform complex tasks. Recent advances, including Google's Agent Development Kit (ADK), are accelerating MAS development, but three critical interoperability challenges remain:

- **Trust:** Continuously verifying the trustworthiness of agents, collectives, and their principals.
- **Policy:** Dynamically governing agent behavior at runtime through external policy controls.
- **Teamwork:** Moving beyond simple task delegation toward resilient and adaptive human-agent-robot teamwork (HART).

Proposed Approach:

Integrate blockchain-based trust monitoring, the KAoS policy management system, and the Luna teamwork framework to support secure and effective MAS collaboration.

MASST Initiative:

Develop and evaluate open-source safety and teamwork layers for interoperable MAS, including applications in cyber situation awareness.



27/09/2026



03:00 P.M



Building 24
Room 121