Institute of Artificial Intelligence

THE IAI DISTINGUISHED SPEAKER SERIES



Prof. Mohit Bansal

AI Challenges: Trustworthy Collaboration, World Discovery, and Long-Horizon Memory

September 16, 2026 | 11am | SPRK 101 & Zoom

Abstract

In this talk, I will discuss 3 major pitfalls and challenges of current state-of-the-art AI agents, and present potential solutions for: (1) Teaching agents to be trustworthy and reliable collaborators based on: social/pragmatic multi-agent interactions via speaker-listener confidence calibration, learning to balance positive and negative persuasion, and multi-agent AI safety through the lens of compositional attacks, theory-of-mind, and belief-steering; (2) Discovering and improving skills/world models needed for efficient, robust action and collaboration based on: learning programmatic skills, weakness-driven adaptive data/environment generation for skill improvement, and structured, selective world model discovery and inference; (3) Planning of long-horizon memory for multi-step reasoning and generation over continuously evolving, conflicting, and scattered information. We will cover diverse domains (math, commonsense, coding, tool-use, computer use, etc.), modalities (text, images, videos, audio, layouts, etc.), and real-world applications (early medical diagnosis and classroom education engagement).

Biography

Dr. Mohit Bansal is the John R. & Louise S. Parker Distinguished Professor, Director of the MURGe-Lab (UNC-AI Group), and Core AI Lead of the ENGAGE NSF-AI Institute in the Computer Science department at the University of North Carolina (UNC) Chapel Hill. He received his Ph.D. from the University of California at Berkeley (where he was advised by Dan Klein) and his B.Tech. from the Indian Institute of Technology at Kanpur. His research expertise is in multimodal generative models, reasoning and planning agents, faithful language generation, and interpretable, efficient, and generalizable deep learning. He is an ACL and AAAI Fellow and recipient of the Presidential Early Career Award for Scientists and Engineers (PECASE), IIT Kanpur Young Alumnus Award, DARPA Director's Fellowship, NSF CAREER Award, Google Focused Research Award, Microsoft Investigator Fellowship, Army Young Investigator Award (YIP), DARPA Young Faculty Award (YFA), and outstanding paper awards at ACL, CVPR, EACL, COLING, CoNLL, and TMLR. He has been a keynote speaker for the IEEE/CVF WACV 2027, IEEE MLSP 2026, ECAI 2025, ACM-CODS 2025, AACL-IJCNLP 2023, CoNLL 2023, and INLG 2022 conferences. His service includes EMNLP Program Co-Chair, Associate Editor-in-Chief for TPAMI, CoNLL Program Co-Chair, ACL Executive Committee, ACM Doctoral Dissertation Award Committee, ACL Doctoral Dissertation Award Co-Organizer, ACL Mentorship Program Co-Founder, and Associate Editor for TACL, CL, IEEE/ACM TASLP, and CSL journals.

Host: Dr. Amrit Singh Bedi | **ai.ucf.edu**

UCF