

Final Session Title

Capability Gaps in Social Humanoids: Methods for Robust Interaction Across Diverse Users

Organizers and Contact

Name: Kazuhiro Nakadai (Main Organizer)

Affiliation: Institute of Science Tokyo, Department of Systems and Control Engineering,
Institute of Science Tokyo, Tokyo, Japan

Phone: +81-3-5734-2550

Email Address: nakadai@ra.sc.eng.isct.ac.jp

Name: Bowen Wu

Affiliation: RIKEN

Phone: (+81) 774-95-1360

Email Address: bowen.wu@riken.jp

Name: Kotaro Funakoshi

Affiliation: Institute of Science Tokyo

Phone: +81-45-924-5294

Email Address: funakoshi@first.iir.isct.ac.jp

Name: Koji Inoue

Affiliation: Kyoto University

Phone: +81-75-753-5396

Email Address: inoue.koji.3x@kyoto-u.ac.jp

Aims and Scopes:

Social humanoid robots are often developed and evaluated around a “default user” assumption, typically healthy adults in controlled environments. As humanoids move into homes, schools, hospitals, public services, and care contexts, this assumption becomes a central technical bottleneck: core interaction technologies (speech/audio, perception, dialogue, nonverbal behavior, and embodied control) can show steep performance drops across user groups and real-world conditions.

This Special Session aims to identify and address these **capability gaps** as a **technology and system-design** challenge, not only a dataset issue. We invite peer-reviewed contributions that,

- (i) Diagnose failure modes and performance cliffs across diverse users and contexts
- (ii) Propose robust and adaptive methods across the interaction stack, and
- (iii) Establish evaluation protocols that go beyond aggregate scores through stratified reporting and interaction-level outcomes.

Topics include robust multimodal and crossmodal perception; adaptive language and dialogue (grounding, turn-taking, clarification, repair); generation and evaluation of legible nonverbal social behavior (gaze, gesture, timing, proxemics); and embodied feasibility and safety under physical constraints.

The session emphasizes cross-stack integration and interface failures that commonly emerge only after deployment, and promotes methodologies that enable humanoids to interact reliably with diverse populations (e.g., children, older adults, deaf and hard-of-hearing users, users with speech/motor differences, neurodivergent users) and in realistic acoustic/visual settings.