

STS Principles for the temporary sociotechnical system co-created by a Zoom-AI-Human virtual ‘team’

during the ‘Teaming with Robots and AI’ panel discussion, hosted by the STS Roundtable & Tavistock Institute, 25 June 2026

Twelve STS principles for Human-Robot Teams (Ang, Sankaran & Liu, 2025), adapted for our discussion of Human-AI teams · summary handout

*Marked **new for HRTs** = introduced by Ang, Sankaran & Liu (2025); **established** = long-standing STS principles, adapted for HRTs.*

STS Principle	Purpose	Definition (in brief)
1 Adaptive Autonomy <i>new for HRTs</i>	Dynamic bidirectional adjustment of autonomy, task-based flexibility, responsive system.	Autonomy levels adjust dynamically and bidirectionally between humans and robots in real time, according to task complexity, criticality and context.
2 Agile & Responsive (Future Thinking) <i>new for HRTs</i>	Iterative design, state-of-the-art knowledge, proactive approach, future thinking.	Design iteratively and stay current with state-of-the-art knowledge, anticipating fast-moving technological change rather than merely reacting to it.
3 Unpredictability Management <i>new for HRTs</i>	Address inscrutable AI-based decision-making in HRTs, human oversight, handle unexpected situations.	Strategies and safeguards — human oversight, intervention and feedback loops — for handling unexpected or inscrutable AI behaviour in dynamic settings.
4 Cognitive Workload <i>new for HRTs</i>	Manage cognitive challenges, reduce burden, complement human abilities.	Design systems that reduce rather than add to the cognitive burden on humans, complementing human cognitive abilities instead of overwhelming them.
5 Ethical Considerations <i>new for HRTs</i>	Incorporate ethical guidelines, address AI-related concerns in HRTs, build trust.	Embed ethical guidelines, transparent decision-making, accountability and informed consent to guide responsible development and build trust.
6 Joint optimisation <i>established</i>	Balance social and technical elements, synergistic, mutual adaptation.	Optimise the social and technical together, not in isolation, so combined human–robot performance exceeds what either could achieve alone.
7 Participation/co-design <i>established</i>	Stakeholder engagement, collaborative creation, ongoing involvement.	Actively engage all stakeholders — workers, end-users and developers — in design and ongoing improvement, valuing even the input of robot capabilities.
8 Multidisciplinarity <i>established</i>	Diverse expertise, cross-disciplinary collaboration.	Bring diverse expertise together so systems are technically sound, usable, and well integrated into their operational context.
9 Collaborative sensemaking <i>new for HRTs</i>	Mutual understanding, shared mental models, anticipatory strategies.	Build shared mental models and mutual understanding between humans and robots, clarifying roles and reducing ambiguity.
10 Boundary Controls <i>established</i>	Define interfaces, manage knowledge and information flow.	Identify and design the interfaces between humans and robots to manage knowledge sharing, monitoring and information flow.
11 Transparency & Explainability <i>new for HRTs</i>	Address the “black box” problem, promote understanding, inform decision-making; understandable AI processes, enable validation.	Make AI decision-making understandable so humans can interpret, question, validate and trust robot actions.
12 Trustworthiness <i>new for HRTs</i>	Design for mutual trust, multi-perspective trust, reliable HRT systems.	Design for mutual, multi-perspective trust (human–human, human–robot, robot–human) to enable safe, effective collaboration and adoption.

Acknowledgment: The research that contributed to this work was funded (partially or fully) by the Australian Government through the Australian Research Council and supported by the University of Technology who is one of the partners of the HARTCon research hub. HaRTCon is funded by Australian Research Council (ARC) through the Industrial Transformation Research Hubs (ITRH) scheme IH 240100016. **AI transparency statement:** Claude Opus 4.8 was used for minor edits; framing edits and prompts by Annja Neumann.