R4 Trustworthy AI: Responsible, Reliable, Robust, and Resilient Human-AI Symbiosis

By Ming Hou, Brad Cain, and Shaina Raza

The 1st Collaborative Human-AI Symbiosis Workshop was held on 6th, October 2025 at the SMC 2025 Conference in Vienna.

Context:

When the society delves into the realm of Human-AI collaboration, we encounter a complex web of reliability, accountability, trust, and safety considerations. Crafting workflows and processes that up-hold stringent AI system requirements is not just a technical challenge, but a design imperative given the challenge with increasingly autonomous AI decision-making capabilities.

To address the challenge, IEEE Systems, Man, an Cybernetics Society (SMCS) started a new Initiative in 2025 to formulate a Collaborative Human-AI Symbiosis (CHAIS) program and prepare, enable, and demonstrate SMCS as a world leader in CHAIS. It benefits from the internationally recognized SMCS interdisciplinary knowledge, expertise, and technology in Systems, Man, Cybernetics, and particularly Human-AI Teaming. The goal is to design, develop, and demonstrate the intricacies of Human-AI teaming and unlock the potential of this transformative alliance. The CHAIS workshop is the first event of the Initiative for engaging with world leaders in various related expertise areas as well as SMCS and IEEE regular and student members working in the area of concepts and tools in human and AI decision-making processes.

Specifically, the objective of this workshop is to brainstorm, identify trustworthy AI implementation challenges and priorities to advise the IEEE SMCS Board of Governors. Ultimately, the goal is to support development of an SMCS roadmap to address the emerging expertise and capability gaps for AI applications.

Attendees:

Ten invited speakers presented their perspectives about challenges and gaps facing responsible development of advanced automation and AI. The speakers were selected from three sectors, although several have cross-domain experience during their careers

Academia:

- Dr. Mary Cummings, Professor, George Mason University, College of Engineering and Computing (US)
- Dr. Keeley Crockett, Professor, Computational Intelligence, Manchester Metropolitan University (GB)
- Dr. Shaina Raza, Scientist in Applied Machine Learning, Vector Institute for Artificial Intelligence (CA)

Industry:

- Mr. Brad Cain, PEng, Independent Human Factors Engineering Consultant (CA)
- Dr. Dale Richards, Principal Human Factors Engineer, Thales Group (GB)

- Ms. Patricia Shaw, LLM, Global AI and Data Ethics Adviser, Chair, Society for Computers and Law (GB)
- Dr. Qingsong Wen, Head of AI & Chief Scientist, Squirrel Ai Learning Inc. (U.S.)

Government:

- Dr. Ming Hou, Principal Scientist, Defence Research and Development Canada (CA)
- Mr. Edgar Reuber, Principal Advisor, EUROCONTROL (DE)
- Dr. Ashitey Trebi-Ollennu, Principal Robotics Engineer, Deputy Chief Engineer, Assistant Division Manager, Autonomous Systems Division 34, NASA Jet Propulsion Laboratory (US)

Observations:

Currently, there is a lot of hype in the economy about a future with AI, particularly with the release of Large Language Models (LLMs) and generative AI (GenAI) versions of AI; there are also a few who foresee a dystopian AI future. Somewhere in the middle, there must be a region of applicability where AI may provide utility and make life better for society with a risk that is acceptable. Defining the boundaries of this region will require a team with the interests of society at the forefront of concern, willing to encourage responsible development of AI that leads to its applications being trusted appropriately by users.

The expert panel observed that there is a need for increased guidance on process documents and standards to help guide risk management strategies within industry. For some industries, like learning and education, AI may be readily integrated into a system that provides oversight, so risks are minimized. However, for other industries, like self-driving vehicles or Air Traffic Control, development strategies such as "move fast and break things" are highly inappropriate for obvious safety reasons. But government bodies have a difficult time regulating novel technologies that are continually changing. An unbiased, or at least balanced perspective, approach of stakeholders in a responsible organization such as IEEE may be our best chance of advocating for society in responsible application of AI products and services.

There appears to be a role for educating developers of cutting-edge AI systems that IEEE SMC can fulfil, providing guidance and processes to facilitate risk management appropriately without stifling innovation within research and development (R&D) organization. Documentation that identifies what "good design" looks like, in a general sense, for industry and society, respecting legal and ethical concerns of the public. IEEE SMCS could produce an accreditation program for engineers, researchers, designers, and managers that provides knowledge necessary to accountable and responsible development and application of reliable, robust, and resilient (R4) trustworthy AI products and services, including an AI-specific focus in knowledge areas such as Human Factors Engineering, Engineering Law and Engineering Ethics. A complementary role for educating government policy development may also be advantageous so that there is less regulatory uncertainty as well as shorter time between product demonstration and regulation requirements.

Although the large technology companies are front and centre in AI application announcements, AI is a technology that will find its way into companies across the spectrum of sizes. This suggests that regulatory frameworks should be manageable by all, with low-cost or free standards

publications. Further, because of the rapidly changing domain, guidelines and best practices should be published for public consumption to help small or midsize businesses develop. These publications must be distinguished from standards so that they both help development and educated small capital developers about mitigating risk responsibly, with benchmark performance metrics.

Finally, IEEE SMCS might, in a role of enabler and educator, collaborate with other respected, non-partisan organizations such as Association of Computing Machinery (ACM) and Human Factors and Ergonomics Society (HFES) to establish a repository for pre-experiment protocol registration and resulting data such that it can be made publicly available. The intent of such a repository is to both publicize best practices for ongoing experimentation as well as provide training data to facilitate R&D for a Responsible, Reliable, Robust, and Resilient (R4) Trustworthy Human-AI Symbiotic Partnership.

Recommendations:

The CHAIS 2025 workshop concluded that realizing a Responsible, Reliable, Robust, and Resilient (R4) Trustworthy Human-AI Symbiosis requires a concerted, interdisciplinary effort. It was agreed that IEEE including SMCS and other sister societies/councils should establish a dedicated Task Force on Trustworthy Human-AI Collaboration program to advance this agenda. Pilot initiatives will be encouraged in safety-critical sectors such as healthcare, transportation, and defense to demonstrate responsible integration pathways. In parallel, IEEE including SMCS and other sister societies/councils should explore education and certification programs to equip engineers, researchers, and policymakers with the competencies necessary for responsible AI design and deployment. Finally, collaborative engagement with international standard bodies and policy organizations should ensure that CHAIS outcomes inform and align with emerging global governance frameworks for trustworthy AI.

CHAIS 2025 Workshop Guest Speakers and IEEE SMC Representatives.