Trusted Artificial Intelligence Challenge for Armaments Systems Engineering

Objective. This research addresses the challenge of designing and operating systems that contain AI and autonomy with uncertain performance to provide overall systems behaviors that are responsible and trustworthy. Rather than focusing on changes to the AI model to increase trust, this research focuses on developing and improving systems engineering methods to provide this increased level of trust. This effort focuses on:

- Assured design of AI and autonomy into systems, and
- Risk-based monitoring and management for the operational use of AI-based capabilities.

Research Method. Student teams will explore the performance of the AI models over a variety of operational conditions, design the human-machine team and decision support system, and finally participate in an operational simulation of a mission scenario over the course of three semesters. Each semester, the teams will be judged on their systems engineering approach, quantitative results, and plans for the future stages.

Mission. The mission involves clearing a safe passage through a minefield using autonomous ground and aerial vehicles, remote sensing, and an AI detection model whose accuracy can vary with type of terrain, topography, season and lighting conditions. Human subject matter experts can also review imagery to detect mines, with better accuracy in some cases, but cannot assess the imagery as quickly as the AI model. For the first semester, the unmanned ground vehicle (UGV) is 100 percent effective at detecting and clearing mines, though that condition may be relaxed in future stages. The student teams have been provided with SysML models of the concept of operations, use cases, and block definition diagrams of the system elements and environmental conditions. The overall time needed to chart and clear a safe passage through the terrain is the primary measure of effectiveness.

Expected Outcomes and Relevance. This scenario provides the opportunity to design a framework for making high-consequence decisions involving trust and autonomy. Teams will be rewarded for identifying systems engineering best practices and novel approaches that provide for safe and trusted operations. Much as we can design reliable systems from less reliable components, the researchers hope to identify systems engineering methods to design and operate trusted systems from less trusted components.

Schedule. The challenge will be conducted in three stages – Summer 2024, Fall 2024 and Spring 2025. Summer submissions are due on August 9th. Each team has an opportunity to provide a 20-minute presentation to the judges on the afternoon of August 20th. The meeting starts at 12:30pm – participants are welcome to attend each other's sessions.

Deliverables. Student teams will turn in a white paper and presentation at the end of each stage describing their systems engineering approach, quantitative results, and plans for the next stage of the competition. The teams will also submit any software, MBSE models or other systems engineering artifacts developed for the competition. Each team's white paper and presentation will be shared with the other teams.

Judging Criteria. The judges will evaluate each team's deliverables against three criteria – best practices (40 percent), novel approaches (40 percent), and plans for the next stage (20 percent).

Inputs. The ability of the AI model and the human subject matter expert to detect mines was captured at four test events, which were conducted at two different locations. There were a variety of terrain conditions across a 10x10 grid, with data collected at different times of the day. Details can be found at the GitHub site below. In addition, a SysML model of the system, along with use cases and the performance characteristics of the different system elements, is also available on GitHub at https://github.com/tsherburne/aic.