Trusted AI Challenge Series

Air Force Research Laboratory (AFRL), State University of New York (SUNY), IBM, NYSTEC, National Security Innovation Network (NSIN)

Presented by Innovare Advancement Center

Request for White Papers
Deadline: June 4, 2021, 5:00pm (EST)

I. Overview

Innovare Advancement Center is a globally connected, world-class facility acting as a lightning rod for top scientific, engineering, and entrepreneurial talent to leverage highly specialized resources and accelerate both expertise and innovation in critical research areas, including artificial intelligence/machine learning, cybersecurity, and quantum information science. As part of Innovare Advancement Center's outreach, it is announcing the Trusted Artificial Intelligence (TAI) Challenge Series. Interested participants will have an opportunity to submit a two page white paper after the competition is announced virtually on April 29, 2021.

The TAI Challenge Series will cover one of four distinct topic areas:
Topic #1 - Verification of Autonomous Systems
Topic #2 - Human-Artificial Intelligence Performance Optimization: Trust and Joint Action for Digital Data Analysis
Topic #3 - Dynamic Bi-Directional Trust in Human-AI Collaborative Systems
Topic #4 - Trustworthy AI Certification

Each topic represents critical areas for AFRL and its partners, and the goal of this competition is to help advance the mission to build a magnetic ecosystem in which the world's leading scientific and entrepreneurial talent tackle the greatest challenges to national security and economic competitiveness for the TAI realm. Please see Section IV for topic details and eligibility for academic, small business, and international R&D communities.

II. Background

This TAI Challenge Series event follows Event 1 of the series "Building the Vision," held Oct 14, 2020 that covered a set of thought-provoking talks and included an interactive panel offering industry, research, and government perspectives, and insights into the critical path requirements for building reliable, robust AI and autonomous systems that can be widely adopted. While current machine learning and AI technologies are focusing on many issues for static data and systems, dynamic systems such as autonomous vehicles, drones, and unmanned aerial vehicles are increasingly being deployed in both civilian and military contexts. Of special interest to this forum are formal methods, protocols, and standard certifications for testing, validation, and certification of trustworthy systems along with the supporting infrastructure and tools. Further, the next generation of technologies will involve evolutionary computing that focuses on the system's ability to learn, prioritize and discount knowledge as it evolves through interaction with people, the environment, and other systems. Through these challenge problems, we hope to uncover novel solutions that move the community closer to addressing these needs, in the context of today's concrete problems.

A recording of Event 1: Building the Vision" can be found using the following links:

https://youtu.be/Uuk0k59I7Y4
https://www.innovare.org/events/trusted-ai-challenge-series

III. Competition Details

Funding: Approximately $500,000 will be available to fund up to one (1) year grants to successful proposers, of approximately $50,000 - $100,000 per grant.

IV. Topic Descriptions and Request for White Paper Submission Details:

INNOVARE
ADVANCEMENT CENTER

Topic #2: Human-Artificial Intelligence Performance Optimization: Trust and Joint Action for Digital Data Analysis

Sponsor: AFRL and AFOSR

Eligibility: US and International Academic Institutions may apply

Target: Up to $100,000 per effort for 1 year. Expecting to fund 1-3 proposed efforts**.**

Objective: This topic seeks to optimize human-AI performance and efficiency through the lens of shared joint action for digital data analysis and reporting. Achieving this objective requires new methods and greater precision for human-AI analytic joint action that consider trust calibration and require system learning and adaptation, with preference for human-AI collaboration that have real-world applications.

Description: Current research on trust in AI is nascent; however, the state of the literature suggests that the trust construct for AI must be evaluated from the side of the human trustee and the AI trustor, and should consider shared context or situation awareness to achieve joint action. In addition, these results also suggest that the traditional trust model for automation continues to be viable for human-AI collaborations, with ability, integrity, and benevolence as the primary bases of trust[1], with the caveat that these indicators generally require intent to cooperate versus coordinate.

To achieve this objective, trust must be appropriately calibrated between the human operator and its AI partner to effectively and efficiently deliver analytic products. Such products will vary across timelines, when the human-AI partnership have disparate goals beyond the required collaboration (e.g. scheduling, push-pull demands) to produce required scheduled and on-demand analyses. In these cases, human analysts must trust the data pushed and pulled from the AI, and the AI must learn to adapt to the preferences, topics of interest, expertise, and scheduling demands for one or more human analysts to calibrate trust. The topic sponsor may provide a subject-matter expertise and evaluation criteria by which to evaluate the effectiveness of the approach. Experts may suggest different capabilities, whether in-house or external, for potential testing scenarios and provide input and scope for shared analytic joint action tasks of interest to current and future lab programs.

Guidance: Success for this topic would entail demonstrable improvement of human-AI performance and efficiency in joint analytic tasks that involve analysis and reporting. A desired solution would include measures or mechanisms of trust calibration and agent learning and/or adaptation to one or more human analysts. Selected efforts will be for a short-term effort, such as a seedling effort, and the anticipation is that the final solution could be further developed. All solutions would include suggested measures of performance and efficiency that consider trust, learning, and adaptation. Solutions may also be selected for extension and additional funding. Proposers should provide a research extension plan at the effort's conclusion to be considered for additional research funding and extension.

Summary: This topic seeks ideas that consider optimizing performance and efficiency in analytic joint action

tasks. Potential solutions should address operator-trustee trust calibration during said action, with a preference for research, development, and testing using real-world or simulated real-world joint action analytic contexts, which could range from human resource or training data and reporting requirements to tasks where analyses using big data, or large amounts of structured and/or unstructured data in one or more media formats, could benefit from human-AI collaboration.

White Paper Submission: Submit a 2-page white paper of proposed research project in the format described below. The deadline for 2-page white papers is **5:00 pm (EST), June 4, 2021**. Proposals should be submitted in PDF form via e-mail as instructed below.

1. Proposers are limited to one submission. Multiple submissions or a single proposal addressing multiple problem areas will not be accepted or further evaluated. Proposers are eligible to submit additional proposals under Topics 1, 3 or 4 subject to eligibility criteria identified therein.

2. Email Topic 2 submissions to afrl.ri.taichallenge@us.af.mil.

3. For questions please email to afrl.ri.taichallenge@us.af.mil.

4. All white papers should be 11-point Times or Arial font, single spaced and be a maximum of two (2) pages (not including references).

5. Applicants submitting white papers must follow the white paper template at Attachment A.

6. White papers must clearly address the challenge problem identified in each submission.

Evaluation and Award Process: The Government will employ a two-step process to select proposals for grant funding:

1. White papers will be reviewed by members of the AFRL selection committee using the following evaluation and selection criteria:

A. Primary Evaluation Criteria
   i. The technical merits and innovative aspects of the proposed research and development; and,
   ii. Relationship of the proposed research and development to United States Department of Defense missions.

B. Other Evaluation Criteria
   i. The applicant's capabilities integral to achieving U.S. Air Force objectives. This includes principal investigator's, team leader's, or key personnel's qualifications, related experience, facilities, or techniques or a combination of these factors integral to achieving U.S. Air Force objectives, and the potential risk of this effort to the U.S. Air Force.

2. Submitters will be notified on **June 18, 2021** if they have been selected for award. White papers selected for award will be invited to submit a formal proposal within 30 days to an AFOSR Broad Agency Announcement (BAA). Proposals selected under the BAA will be awarded a one year grant with awards up to $100,000 each.

3. Proposers from US and international universities are eligible to submit to this challenge competition. Proposers who have not previously received a grant from the USAF are also strongly encouraged to apply.

4. Awardees will be invited to present their challenge solution as part of the third and final TAI event in the series, i.e. the "Trusted AI at Scale" event. Trusted AI at Scale is scheduled to take place virtually and will feature top researchers, as well as leaders from governmental, academic, and industrial organizations, from Jul 27-28, 2021.

[1] Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20, 709–734.