# Digital Transformation in Test and Evaluation for AI/ML, Autonomous, and Continuously Evolving Systems – Base Year

## EXECUTIVE SUMMARY
## SEPTEMBER 2023

**PRINCIPAL INVESTIGATOR:**
**Dr. Laura Freeman,** *Virginia Tech National Security Institute*

**PRINCIPAL INVESTIGATOR:**
**Mr. Geoff Kerr,** *Virginia Tech National Security Institute*

**VIRGINIA TECH**

# EXECUTIVE SUMMARY

The Department of Defense (DoD) 2022 National Defense Strategy has realized the need for advanced technology and more rapid development and fielding of that technology to sustain dominance against peer/near-peer threats. In support of these objectives, the Director, Operational Test and Evaluation (DOT&E) engaged the Acquisition Innovation Research Center (AIRC) University Affiliated Research Center (UARC) in a multi-year contract to advance test and evaluation (T&E) methods within the DoD, and the research is concluding the initial years' worth of effort. The multi-university AIRC research team that has partnered with DOT&E has focused research efforts over the past year in support of the DOT&E Implementation Plan. Though established as two contracts under WRT-1070 Test and Evaluation Methods for Middle Tier Acquisition (MTA) and WRT-1071 Digital Transformation (DT) in Test and Evaluation for AI/ML, Autonomous, and Continuously Evolving System, the research team has united the efforts for efficiency and for alignment to the Implementation Plan (I-Plan). The research team focused on, under advisement from DOT&E technical leadership, supporting three of the five key pillars within the Implementation Plan.

- Pillar 1 – Test the Way We Fight
- Pillar 2 – Accelerate the Delivery of Weapons that Work
- Pillar 4 – Pioneer T&E of Weapon Systems Built to Change Over Time

As the greater research team attacked this multifaceted space with focused efforts, there were commonalities in the basic approach. First, the team researched the current best practices, challenges, and strategies within industry and the government through detailed literature reviews and subject matter expert engagements. Second, the team summarized their findings and initial conclusions in a series of publications and presentations, as seen in Figure 1 below. Finally, the team explored test cases to evaluate the initial findings and recommendations. In the upcoming year, the research team will expand on these foundational efforts.
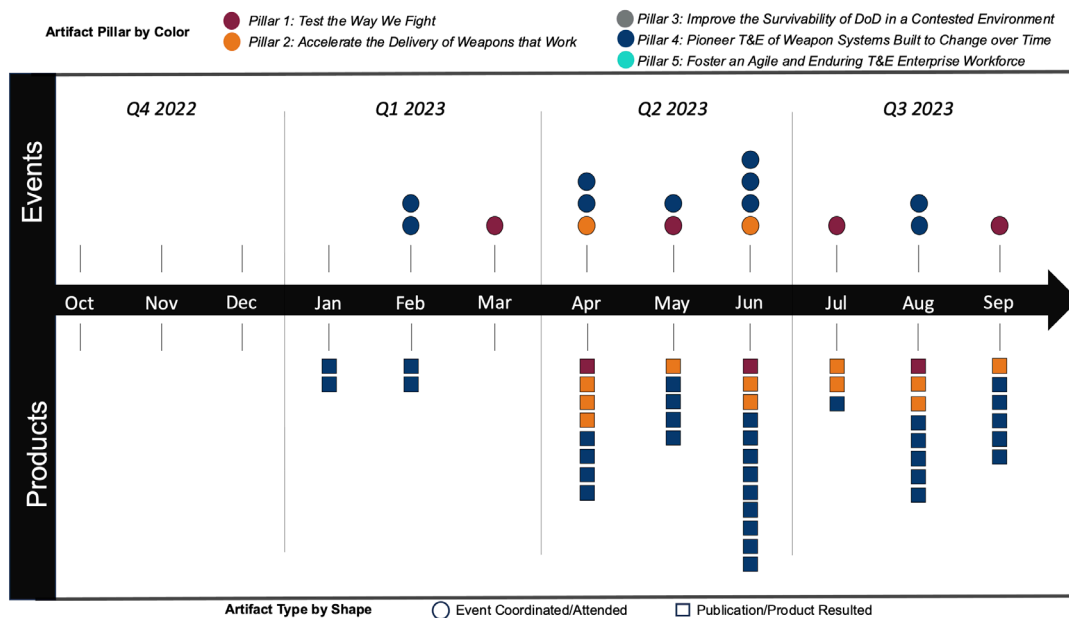


*Figure 1: AIRC Year 1 Resulting Products and Event Participation*

In support of Pillar 1, the AIRC research team focused on maturing evaluation methodologies for Joint Warfighting Concepts with the development of a fundamental Joint Test Concept (JTC). The team established a broad Community of Interest (COI) to explore numerous aspects and challenges around the test and evaluation of Joint Warfighting Concepts. The team conducted a three-phased research approach, each of which concluded in a workshop that first shaped and bound the focus of the research team, developed initial JTC priorities, and concluded with a tabletop simulation of joint operation assessment.

In support of Pillar 2, the research team explored current threats to DoD data and how these threats can be mitigated. The team responded with a Data Security paper that expands upon the findings. In addition to the data security aspects of accelerating Weapons that Work, the research team developed Bayesian-based approaches to leverage contractor testing, development testing, live fire testing, modeling and simulation results, and operational test to ensure more rapid test and evaluation of weapon systems to expedite fielding of warfighting capabilities. In the future efforts, the research team will be working on a Model Based Test and Evaluation Plan (MB TEMP) and supporting the advancement of the Integrated Decision Support Key (IDSK) beyond the initial efforts described in the Digital Engineering for T&E section of this report.

In support of Pillar 4, the research team investigated improved methods for test and evaluation of Artificial Intelligence (AI) /Machine Learning (ML) -enabled DoD systems. The team has developed recommendations on how to ensure AI/ML systems have proper training data while continuing to evaluate and inform system performance while in use. The team further explored methods to evaluate AI/ML systems when the T&E community has limited or no detailed understanding of the underlying AI/ML algorithms within a system. Also, in support of Pillar 4, the multi-university team helped further the application of digital engineering practices in support of Test and Evaluation. With a realistic look at tooling challenges and current cultural challenges, the team made practical recommendations on the use of Model Based Systems Engineering (MBSE) methods for test planning and execution, in addition to digital linkage from mission requirements through operational assessment. Lastly, Pillar 4 efforts also included maturing a framework to automate security Penetration Testing. After conducting cyber security research, the team developed early test software methods that enable cyber physical system penetration testing to be automated in concert with a continuous innovation, continuous deployment product development environment. In the next year, the team intends to test these capabilities with case study DoD products.

Finally, in support of DOT&E emerging priorities, the research team supported numerous tasks that included providing policy reviews and recommendations, offering workforce planning concepts, developing exemplar IDSK, and delivering technical support to T&E meetings, workshops, and conferences. The research impact has been wide reaching. The AIRC team hosted 3 workshops (1 under 1071), 22 presentations/sponsor briefs (12 under 1071, 3 under both), 1 webinar (under 1071), 17 reports/publications (10 under 1071), 2 apps/programming tools.

Based on the foundational work performed in this initial year, the AIRC research team looks forward to building on the year 1 recommendations and tooling by performing case study evaluations on actual DoD programs of record and on conceptual programs. The team also intends to introduce workforce development efforts in the follow-on year in support of DOT&E's fifth pillar, Foster an Agile and Enduring T&E Enterprise Workforce.

## DISCLAIMER