Artificial Intelligence Test, Evaluation, Validation & Verification (AI TEV/V) for DoD: Introduction Sheet

Artificial Intelligence (AI) has the potential to enhance capabilities for multiple applications by efficiently processing massive quantities of data and elevating problem-solving beyond human capacity. These benefits are particularly relevant to the Department of Defense (DoD), where data-driven situational awareness and speed of decision-making define advantage on and off the battlefield. DoD is not alone in pursuing a competitive edge in AI: peer and near-peer competitors are heavily investing in and seek to modernize their forces with AI-enabled technologies.[1] DoD has committed to accelerating adoption of AI to meet this threat and maximize its own capability, but DoD must first test these capabilities to ensure they adhere to safety and performance standards, and this testing can prove challenging for AI-enabled systems that exhibit unpredictable and adaptable behaviors.

In the DIB's "AI Principles" report published in 2019, AI is defined as "a variety of information processing techniques and technologies used to perform a goal-oriented task and the means to reason in the pursuit of that task."[2] This definition encompasses a broad array of capabilities and behaviors that ranges from more predictable and less adaptable systems ("deterministic") and less predictable and more adaptable systems ("nondeterministic"). Nondeterministic systems are expected to provide the step-change in capability that is frequently associated with AI.

Nondeterministic AI-enabled systems require nontraditional Test, Evaluation, Validation and Verification (TEV/V) processes, and DoD will require a robust pipeline of AI-enabled systems and programs to shape its own AI TEV/V capability. The AI community writ-large (private sector, academia, and research) is wrestling with TEV/V for nondeterministic systems, and DoD cannot wait for external solutions if it hopes to adopt and deploy those types of systems in the near term.

There are currently no proven effective methods for testing nondeterministic systems. Without a strong push for education and training on this topic and a diverse range of testing programs at the developmental and operational levels, DoD will have difficulty assessing its current TEV/V processes and determining next steps to improve its AI TEV/V capability.

In the coming months, the Defense Innovation Board will construct a set of recommendations to support the development of DoD's AI TEV/V capability.

---

[1] https://tass.ru/armiya-i-opk/5485116; https://www.brookings.edu/wp-content/uploads/2020/04/FP_20200427_ai_weapons_kania_v2.pdf.
[2] https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT .PDF.