



Department of the Air Force  
Scientific Advisory Board

**DEPARTMENT OF THE AIR FORCE**  
**HEADQUARTERS AIR FORCE**  
**WASHINGTON DC**

**Responsible Artificial Intelligence for Supporting Combat Engagements**  
**Abstract**

While it has been identified that AI may significantly enhance combat operations, the seriousness of lethal decisions in combat situations requires that human judgement be augmented, not replaced, by AI. The Secretary of the Air Force tasked the Department of the Air Force (DAF) Scientific Advisory Board (SAB) to conduct a study on "Responsible Artificial Intelligence for Supporting Combat Engagements (RAI)." The study sought to provide the means to deploy AI at speed and scale, while remaining consistent with DoD Ethical Principles, U.S. ethical norms, Law of Armed Conflict, and cognizant of societal ethical concerns.

The study panel found that there are commercial practices and frameworks that the DAF can immediately leverage to accelerate implementation of RAI principles across its relevant programs. One of the key study results is a science and technology roadmap for near-, mid-, and far-term investments that will enhance the responsible use of AI including rapid model updates, resilient and robust AI, platform interoperability, human-in-the-loop, transfer learning, and multi-domain live-virtual-constructive test ranges to accelerate operational capabilities. Additionally, the study observed that the DAF needs a RAI-cognizant, rigorous verification and validation (V&V) methodology to support continuous DevSecOps for both effective AI systems development, and fielding/sustainment pipelines. Responsible AI systems (operator-machine) need to be developed in the context of enhancing DAF mission-level effectiveness and are particularly relevant to accelerating solutions to the meet DAF Operational Imperatives.

The study identified four specific recommendations for the DAF:

- Invest in the physical and digital infrastructure required to train AI models and incorporate RAI principles, including support for data curation and provenance. In addition, continued use and enhancement of the department's data fabric structure, to assure accessibility to developers and operators as these capabilities evolve. The study's framework facilitates continuous V&V and red teaming to increase model robustness and resilience, capable of countering potential adversarial actions targeting undesired behaviors.
- Expand human-machine teaming and integration into a DevSecOps flow that include multidisciplinary teams over the full lifecycle. As a part of this process, identify data simulation needs and limitations from rare or high-consequence, and corner-case events.
- Establish system-engineering boundaries (technical, risk, and performance) for AI agent delegation, grounded in the test and training contexts. Model processing products and derived results should be recorded for later review, with functions similar to an enhanced black box, to ensure auditability.

- Incentivize and promote the AI workforce that can implement the responsible use of AI-enabled technologies. Personnel management tools such as the use of Highly Qualified Individuals and Special Education Indicator trackers will enable rapid deployment of strong teams with multidisciplinary experts. Constructive service credits, already used in medical fields, would advance accession levels and integrate technically specialized personnel into operational units that are starting to introduce RAI practices.