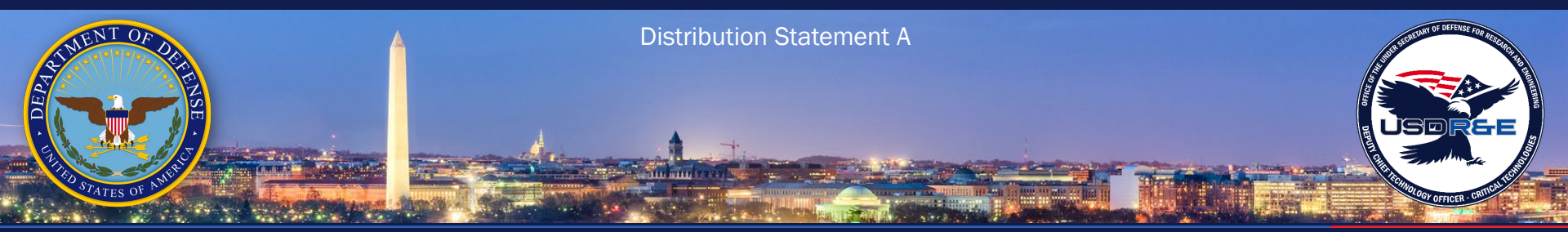


Secure, Robust and Scalable AI/Autonomy - A Holistic, system of systems approach to development of trustworthy AI-enabled systems

Name: Dr. Kimberly Sablon
Title: Principal Director TAI&A
Organization: OUSD (R&E)/Critical Technologies
Event: George Washington University

Controlled by: OUSD (R&E)
Controlled by: Critical Technologies
Category: Critical Technology
Distribution: A
POC: Dr. Kimberly Sablon, 703-692-6930

OV-1: Future Operational Environment…

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES

Dynamic, Contested, Congested, Resource Constrained with Novel Adversary Technologies

- The battlefield will be highly lethal, involving all-domains, with adversaries increasingly on technical par with the U.S.
- Warfighter tasks will be increasingly augmented or performed by machines
- Warfighters will be disaggregated for extended periods of time, often subjected to intermittent comms
- All combatants will demonstrate an increased reliance on CCD and decoys
- Persistent enemy surveillance, GPS denial, spoofing and deception will be the norm
- Behaviors of massive urban populations may significantly influence the execution of military operations
- ***The volume of information and required speed of decision-making will accelerate dramatically***

AI technologies should provide multi-domain real-time situational awareness, improved targeting and multi-target fire control, advanced maneuverability and protection, low-detectable communications, logistics optimization, and faster, improved command and control.

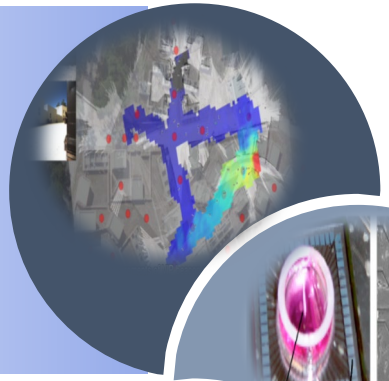
AI-enabled Behaviors in a Military Context

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES

AI Perception, Reasoning, Understanding, Planning → Data Driven Decisions at Speed and Scale

Assess

Distributed reconfigurable sensing;
Sensor AI;
Interactive data fusion;
Context-aware data



Plan/Decide

Hierarchical distributive and collaborative intelligence;
AI-accelerated planning and decision making;
Complex uncertain data environments
Distributed lethality

Act

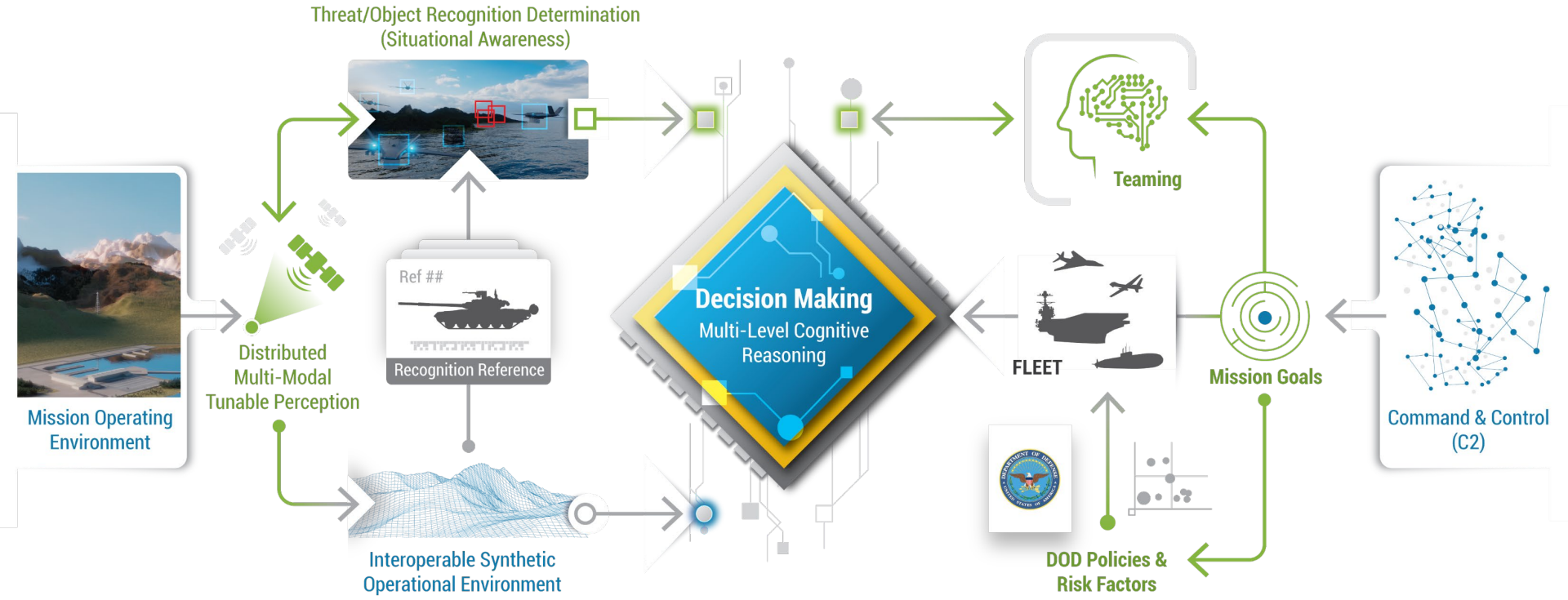
Heterogeneous robotic cooperation;
High degree of freedom actuator;
All-terrain mobility;
Long duration autonomy in complex environments
Collaborative Engagement

Team

Rapidly adaptable human-AI teams; Hybrid cognitive architectures; Soldier training/ performance
Trusted human-machine teams
Autonomous operations

Cognitive Autonomy: A System of Systems

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES



We must take into consideration all complexities of the real-world environment and we need to integrate value judgement systems

Focus Areas

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES

Robust AI

Performance across Multi-environments/Domains

- Systems need to achieve mission success across novel conditions and environments
- Interactive perception systems to enable robust autonomy behavior in multi-modal environments

Resilient against Inadvertent Bias and Adversarial Manipulation

- Resist adversarial tampering from manipulation of input or training data
- Detect and recognize deception

Scaling AI

Distributed Hierarchical AI Architectures to enable Edge Intelligence

- Within contested environments, systems must make inferences at the edge
- Decision making tools must use appropriate abstractions, accounting for uncertainty

Autonomous Multi-level Orchestration and C2 of AI/A Components and Systems

- The architecture must know which model (or ensemble) to push for the mission and operational environment

Human-Centered AI and Autonomy (Warfighter Trust)

Center for Calibrated Trust Measurement and Evaluation (CaTE)

- A system engineering approach to developing trust with the operator during design and testing
- Must include warfighter training

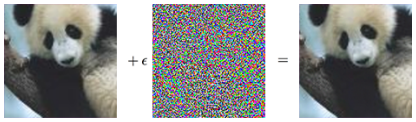
Merging Calibrated Trust Frameworks with Mandatory Model-based System Engineering Workflows

- Standards, methods, and processes for providing evidence for assurance; and for calibrating trust

AI Infrastructure and Community of Action

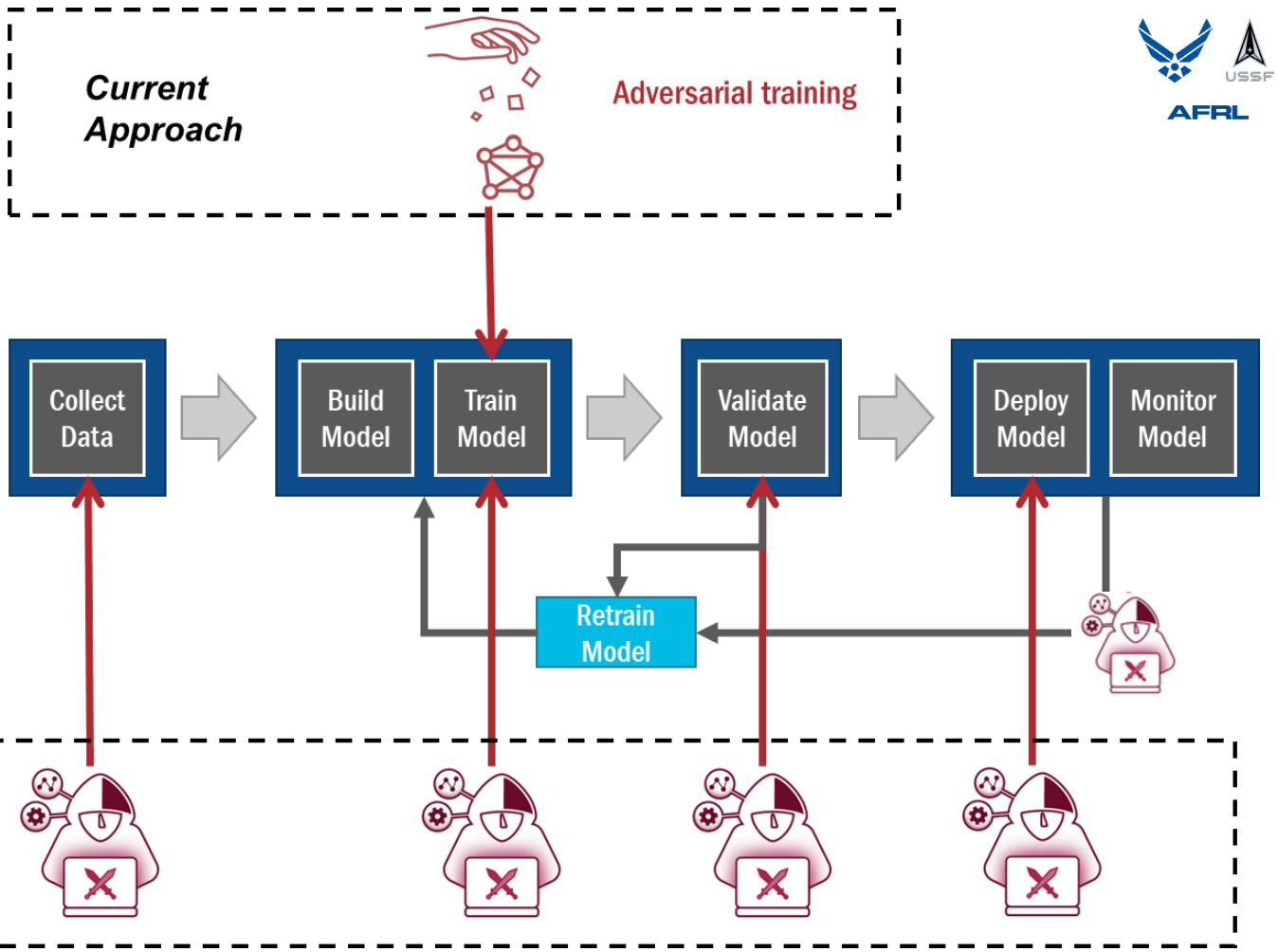
Robustness to Adversarial Attacks

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES



"SpeedLimit65"

"rifle"

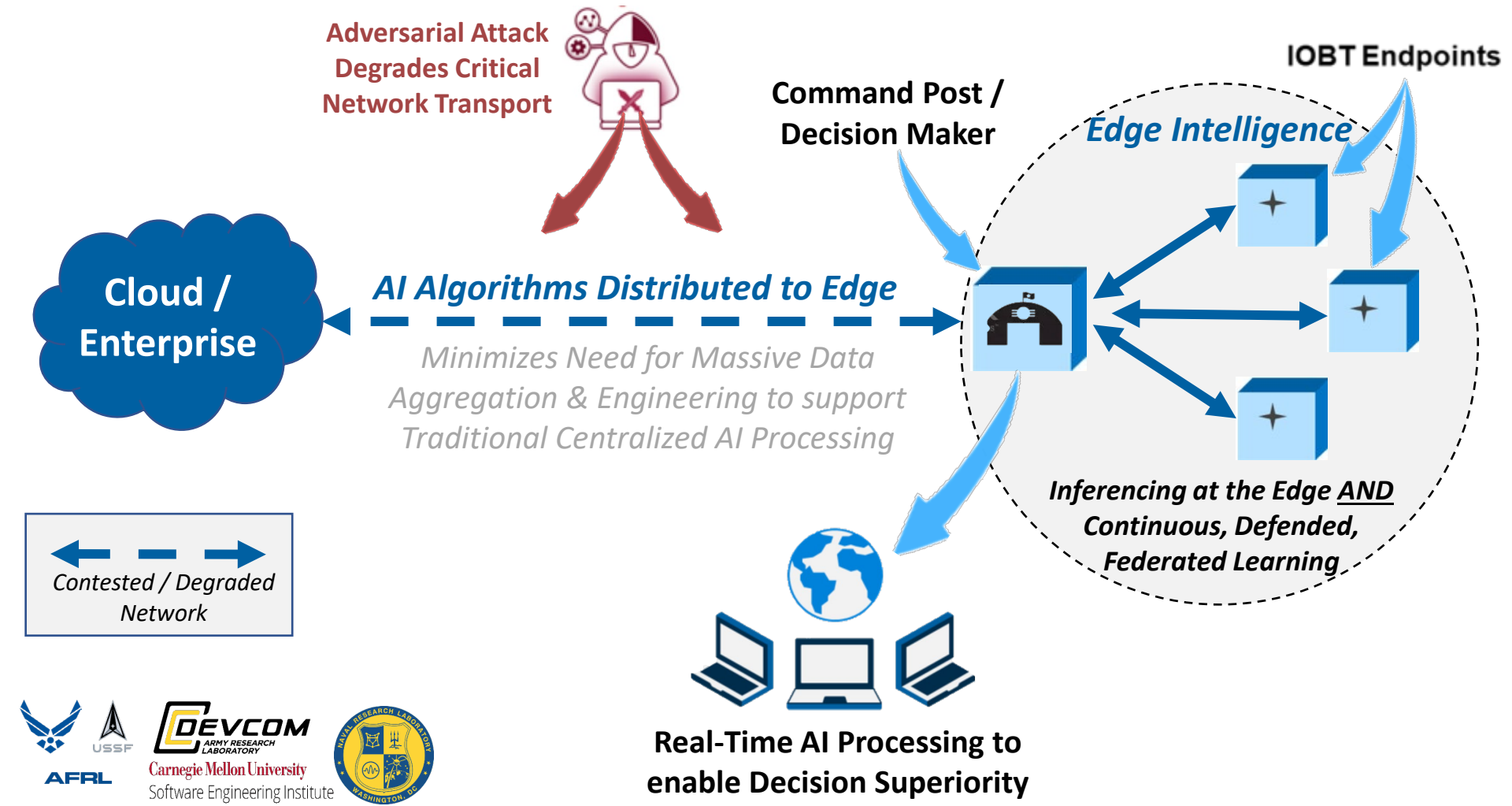


Continuous adversarial testing and red-teaming approaches must be applied throughout the system lifecycle – from development to deployment

DDIL-Resilient Decision and Control Distributed AI Architectures to Enable AI at Scale

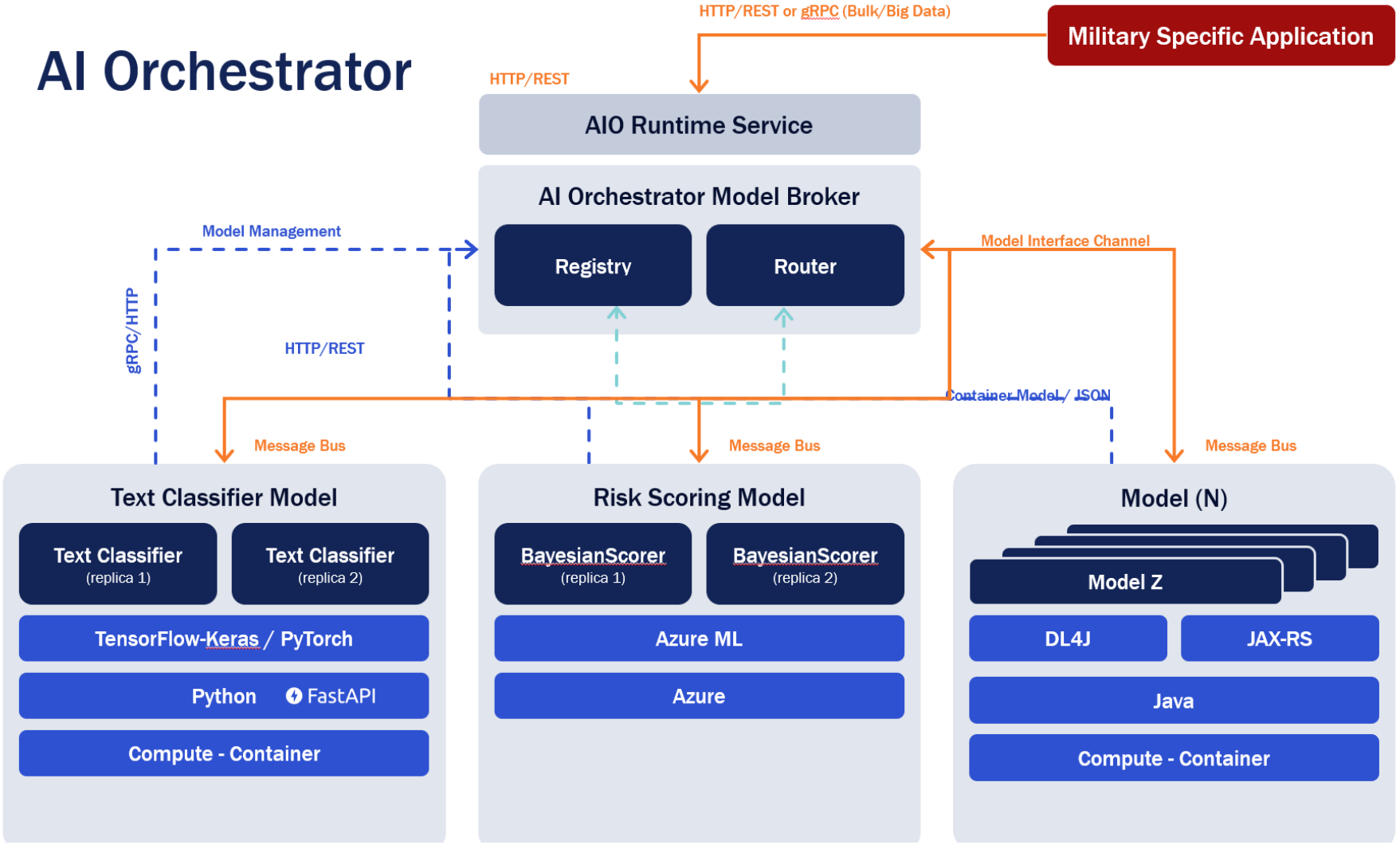
THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES

We must leverage advances in Distributed ML and Edge computing to enable Distributed AI architectures that minimize the dependency on massive data aggregation for AI-empowered Decision Superiority



AFRL USSF DEVCOM ARMY RESEARCH LABORATORY Carnegie Mellon University Software Engineering Institute

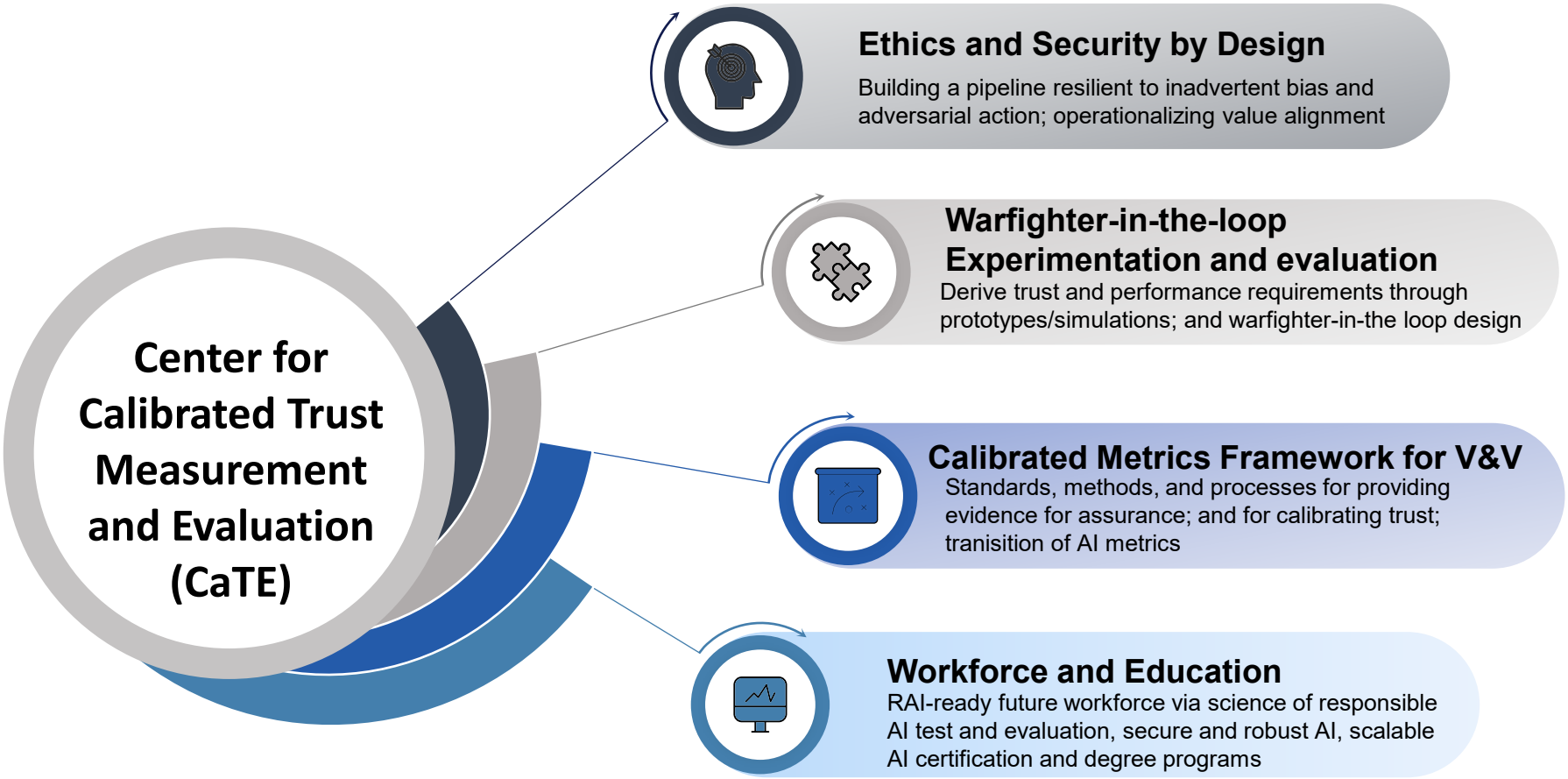
Multi-level Orchestration of AI/A Components



Pushing towards Warfighter Trust

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES

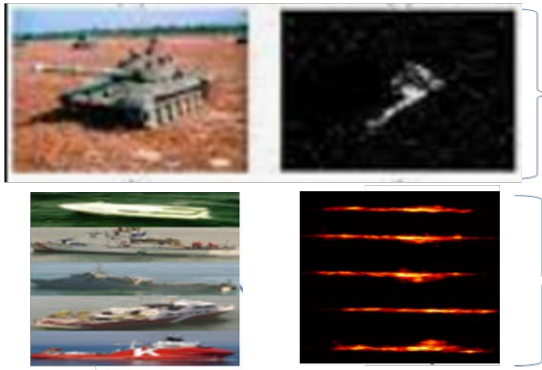
Bringing TEV&V, R&D and Acquisition under the same umbrella to develop common frameworks for providing evidence for assurance and to develop calibrated levels of trust in human-machine teams (HiL/HoL)



Getting the AI Infrastructure right – Driving towards an Integrated, Multi-phenomenology Capability

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES

AI Research Hubs provide common infrastructure (networks, data storage, and computing) for researchers across the DoD S&T enterprise to share previously siloed data, establish common standards and development tools (labeling, synthetic data generation, M&S environments, and test harnesses)



Fusing EO/IR and SAR imagery for Ground targets

Fusing EO/IR and SAR imagery for Maritime targets

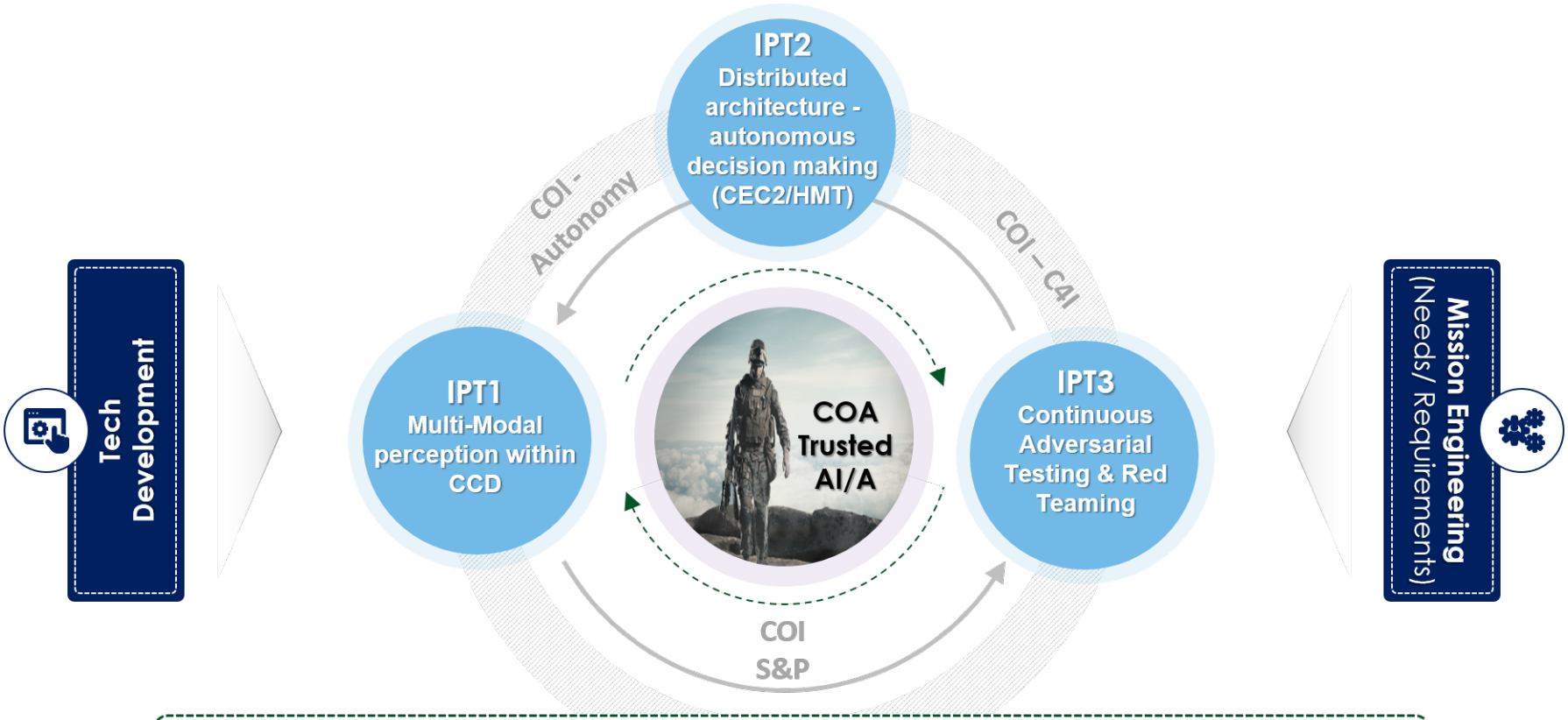
1. Accelerate the development and sharing of Automatic Target Recognition (ATR) / Machine Learning (ML) tools, across the services, for multi-domain applications.
2. Develop ATR tools that can render significant advantages in delivering accurate target ID and shortening kill chain in contested environments.
3. Amplify existing investments made by the services and foster productive collaboration resulting in cross-service dataset sharing and tool development.
   i. Develop tools for rendering single or limited aspect views of EO/IR imagery of targets into 3D models to facilitate association with corresponding SAR data.
   ii. Develop co-registration tools for ***EO/IR and SAR imagery***.

Common, security-appropriate infrastructure across DoD S&T to share previously siloed data, establish common standards and development tools, addressing core data-specific and multi-sensor fusion problems

Building a Community of Action - The Warfighter is the Center of Gravity

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES

Driving Joint Operational Capabilities with Mission Partners



TAI/A CoA will be organized under the JRASE to focus efforts of researchers and engineers on a system of systems approach to research, development and integration of AI-enabled components across warfighting functions, echelons and domains with emphasis on rapid reaction experimentation

What does this mean for autonomous systems?

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES

Where are we across the Department? Where do we want to be?

Cross-Echelon, Resilient Autonomous Networks of Autonomous Systems

Collaborative and federated learning, reasoning in complex and adversarial environments, collaborative real-time AI-generated courses of action with systems that understand functions and limitations, decentralized coordination across Warfighting functions

Far-Term

Intelligent and Collaborative Robotics and Human-Machine Teams

Scalable, modular and multi-functional robotic systems, integrated control and human-machine interfaces, control and optimization of autonomous resources

Mid Term

Navigation, Hyper-active Maneuvers and Mobility

Unmanned air, ground, and maritime platforms; multi-environment navigation and maneuvers in complex, adversarial environments; contested logistics.

Near-Term

Multi-Object Recognition, Targeting and Fires

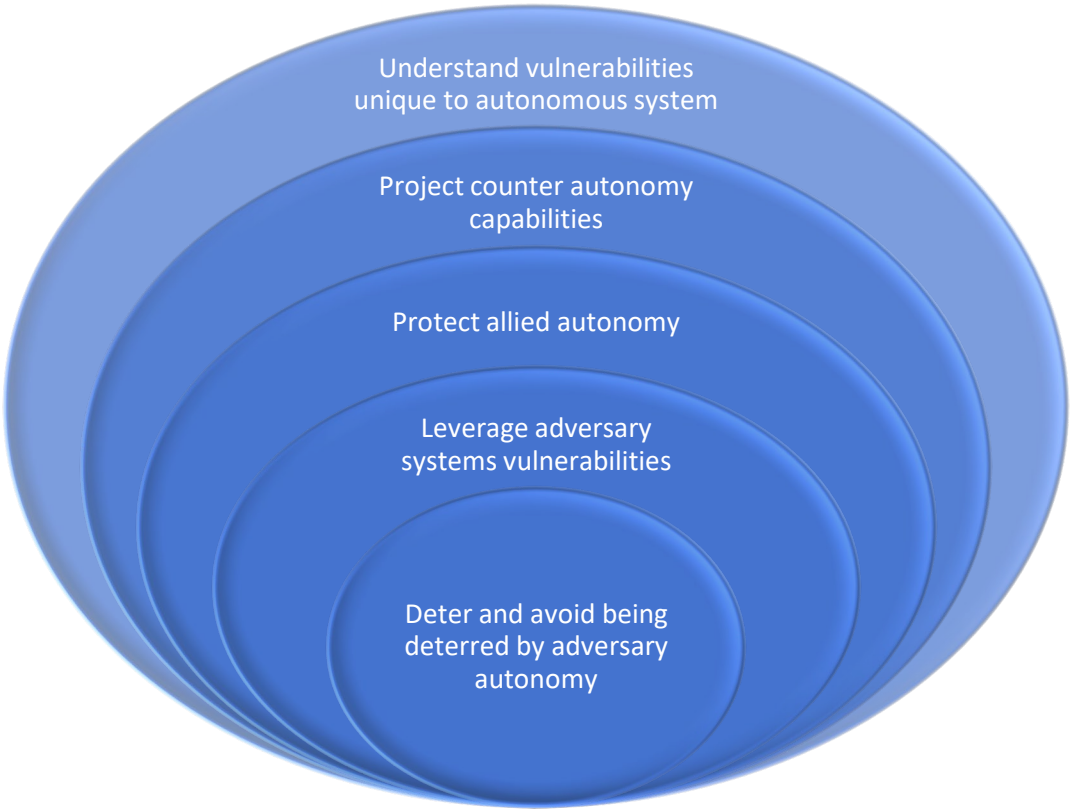
Collaborative targeting and fire control, Intelligent, on-board multi-modal threat recognition, synchronization of complex tasks, AI-enhanced weapon-target paring, AI-agent based decision making

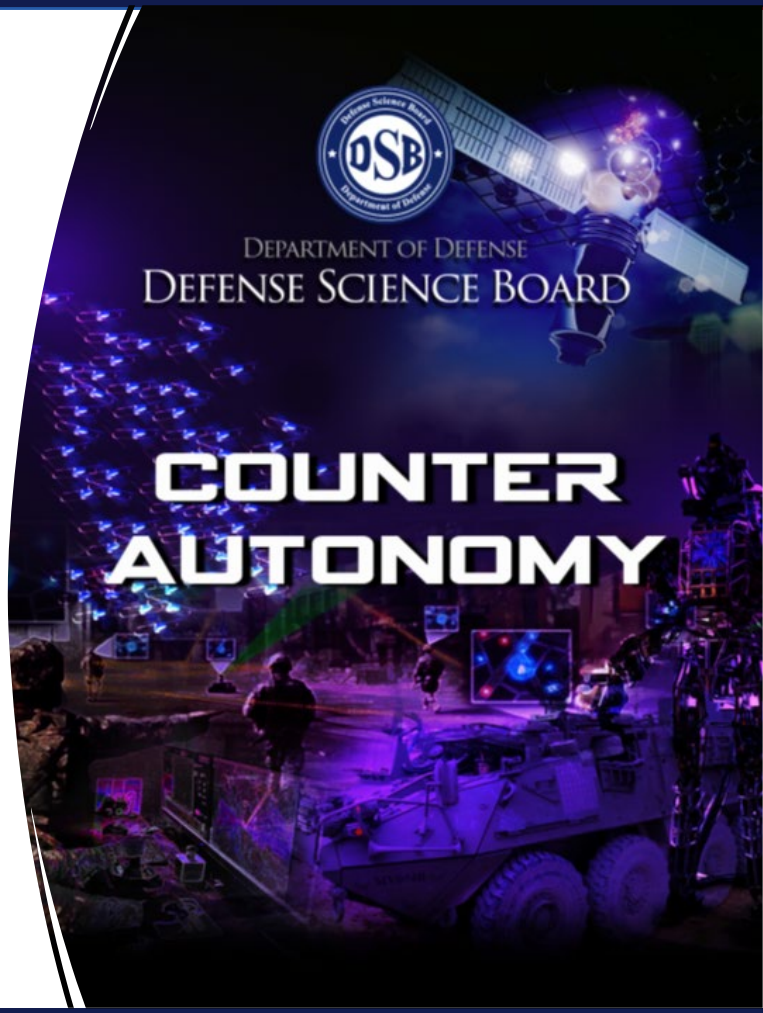
Distributed, Safe, Secure Collaborative Behaviors with emphasis on Human-Machine Teaming – A framework to integrate AI/ML solutions at multiple layers and across warfighting functions is required!

Counter-Autonomy

THE OFFICE OF THE DEPUTY TECHNOLOGY OFFICER FOR CRITICAL TECHNOLOGIES

How do we protect our systems? How do we counter opposing autonomous systems?





Counter autonomy begins with resiliency of US autonomy systems to deter adversaries and create trust in their use

Summary

- **The identified critical technology goals for Trusted AI and Autonomy include:**
  - Inferencing at the Edge AND Continuous, Defended, Federated Learning are critical for operating in the DDIL environment
  - Trust and Resiliency
  - AI Federated Infrastructure
  - Collaboration and Workforce
  - Considerations for the Defence Industrial Base
- **To advance these goals, Trusted AI and Autonomy is standing up the following, supporting and expanding as funding allows:**
  - AI Hub S&T Capability Incubators – expanding from the 5 pilot Hubs covering EO/IR, sonar, SIGINT, Modelling and Reasoning, and Maneuver to new data modalities including SAR, radar, LIDAR, and HSI
  - A Community of Action to focus on a system of systems approach to AI-enable capability. Initial IPTs:
    - Multi-modal perception within CCD
    - Distributed architectures-autonomous decision making
    - Continuous Adversarial Testing & Red Teaming
  - Center of Calibrated Trust Measurement and Evaluation (CaTE)
    - FFRDC Calibrated Trust Center @ Software Engineering Institute
    - Academic Autonomous Systems Test & Evaluation Center
  - International Initiatives – including AUKUS RAAIT, CENTCOM experimentation support, and US/UK Agile Defence Proposal Process