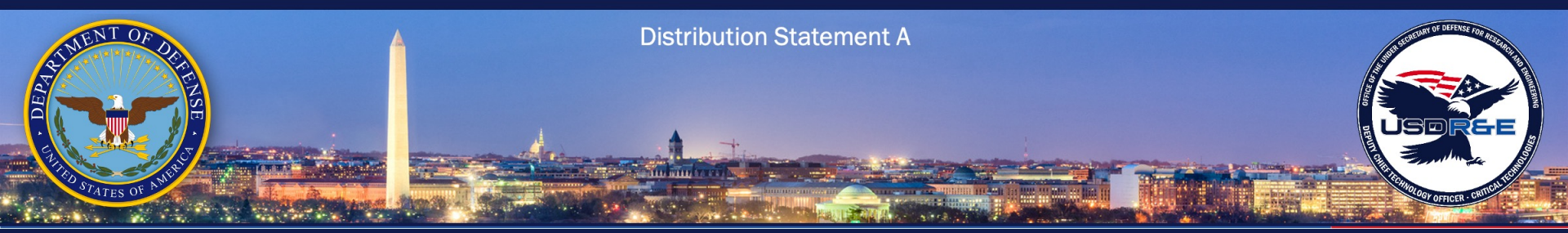


Cognitive Autonomy: The Next Frontier in Intelligent Decision Making

A Holistic, system of systems approach to development of resilient AI and autonomy

Name: Dr. Kimberly Sablon
Title: Principal Director, Trusted AI and Autonomy
Event:

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

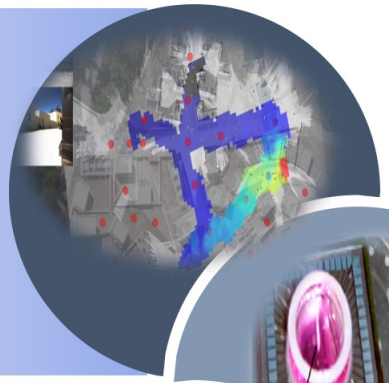
Unlocking Decision Superiority

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

The DOD has emphasized use of AI for a decision advantage
– Commanders Decision Cycle -

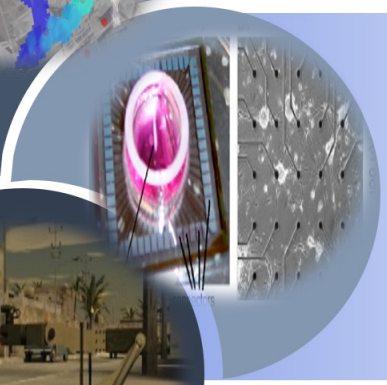
Sense

Reconfigurable AI-sensing;
Interactive data fusion in DDIL environments;
Context-aware data;
Logistics status;
Adversary Actions/Reactions



Plan/Decide

Hierarchical distributive and collaborative intelligence;
AI-accelerated planning and decision making;
Complex uncertain data environments;
Rapidly adjust plans based on operational situation



Communicate/ Act

Rapidly adaptable and trusted human-AI teams; Data interfaces to support sharing and management in autonomous operations

Role of AI in Military Systems

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

AI in the **Commercial** Sector

- *Computer vision for image and video analysis*
- *Predictive analytics for consumer service and sales forecasting*
- *Autonomous vehicles and drones*
- *Robotic process automation*
- *NLP /LLM for virtual assistants/chatbots*

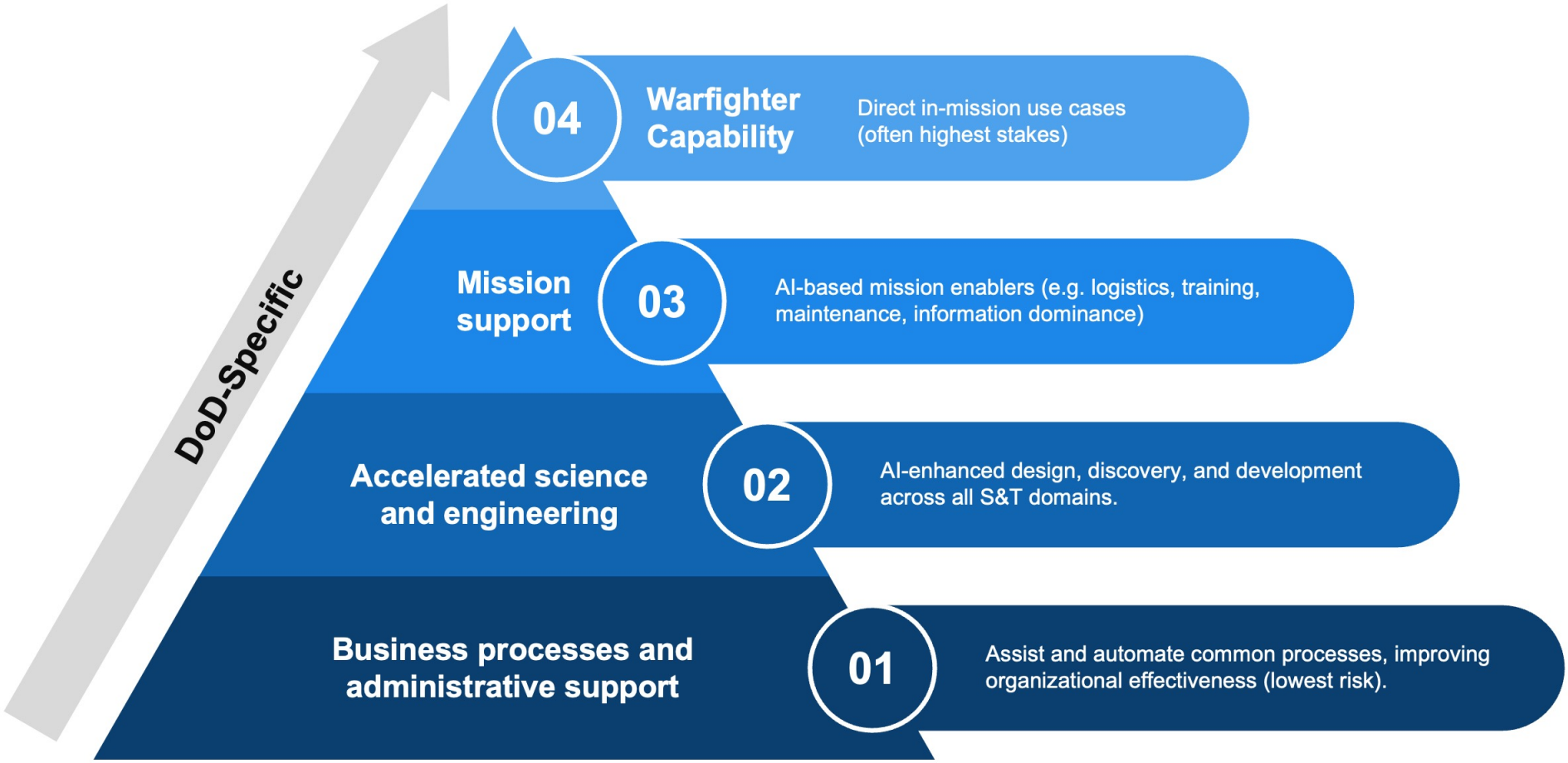
AI for the **Defense** Sector

- *Adapt in highly dynamic, uncertain, unstructured environments*
- *Distributed and scalable across diverse, asynchronous environments (edge-cloud-fwd centers)*
- *Need to operate in resource-constrained environments (energy efficient)*
- *Discriminate between targets and decoys (multi-scale)*
- *Robust and Resilient*

The Best AI Work is not in the DOD…How can we leverage to address DOD problems rapidly and responsibly?

AI Transformation Impact

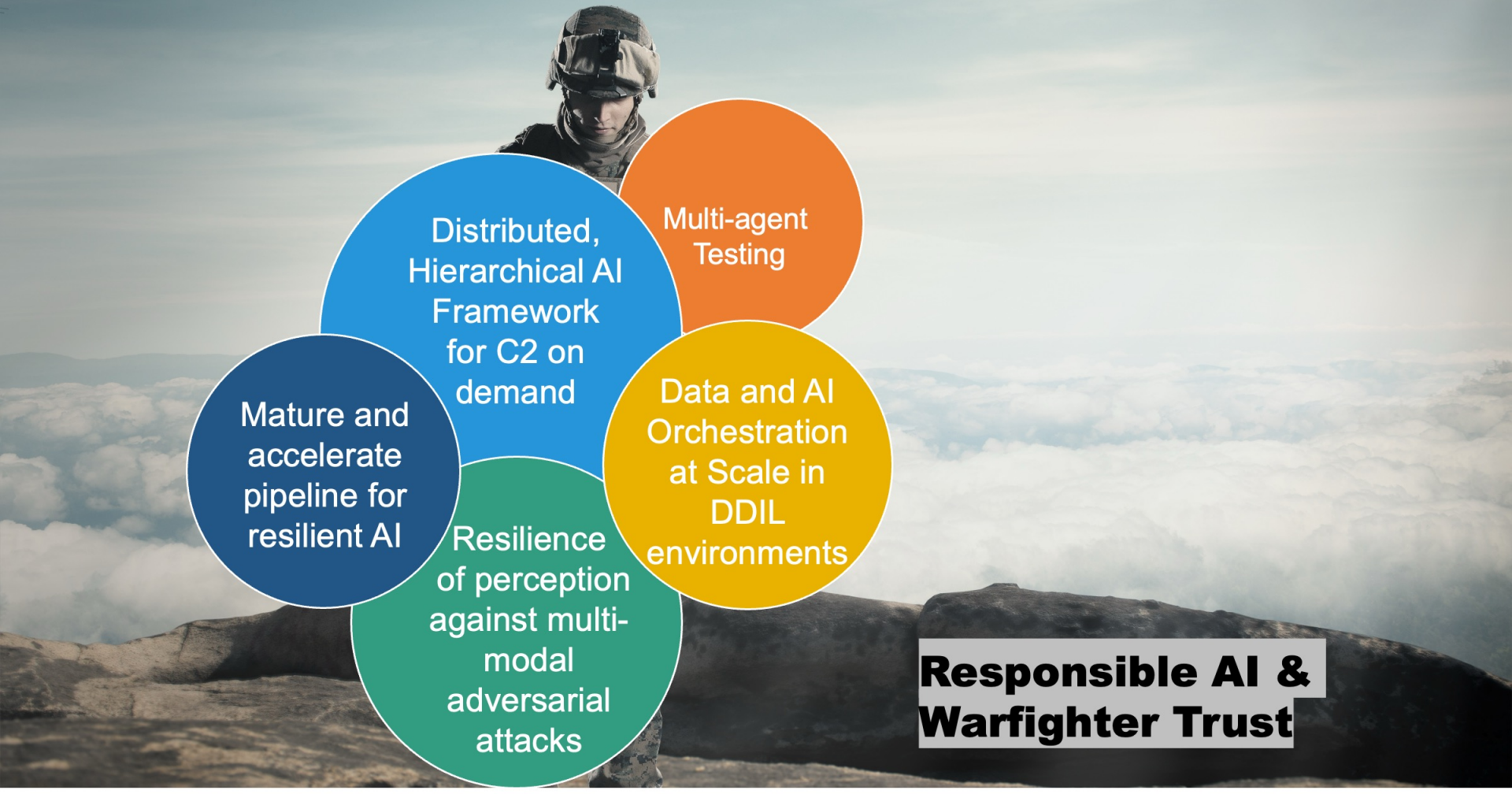
THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES



Challenges we're tackling...

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

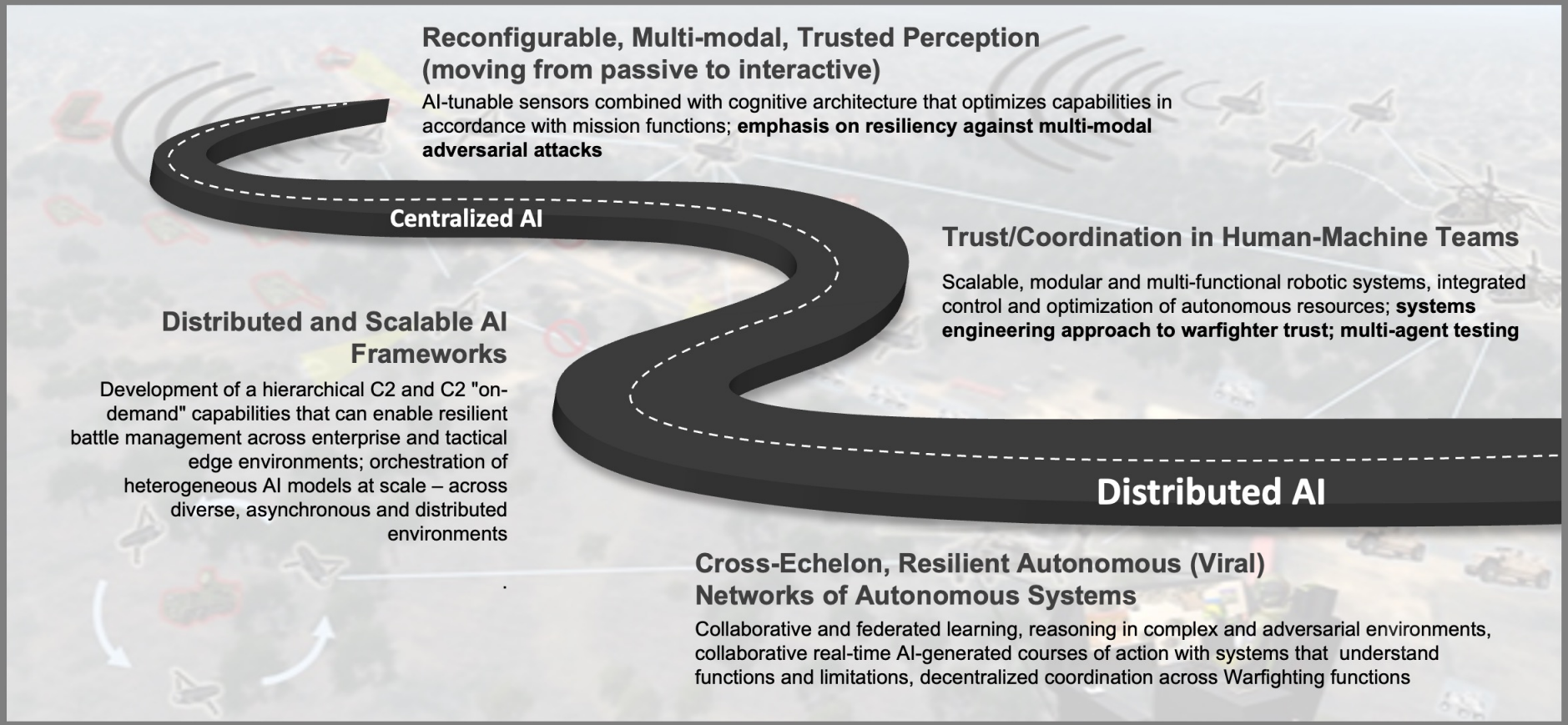
Dynamic, Contested, Congested, Resource Constrained with Novel Adversary Technologies



Roadmap at a Glance

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

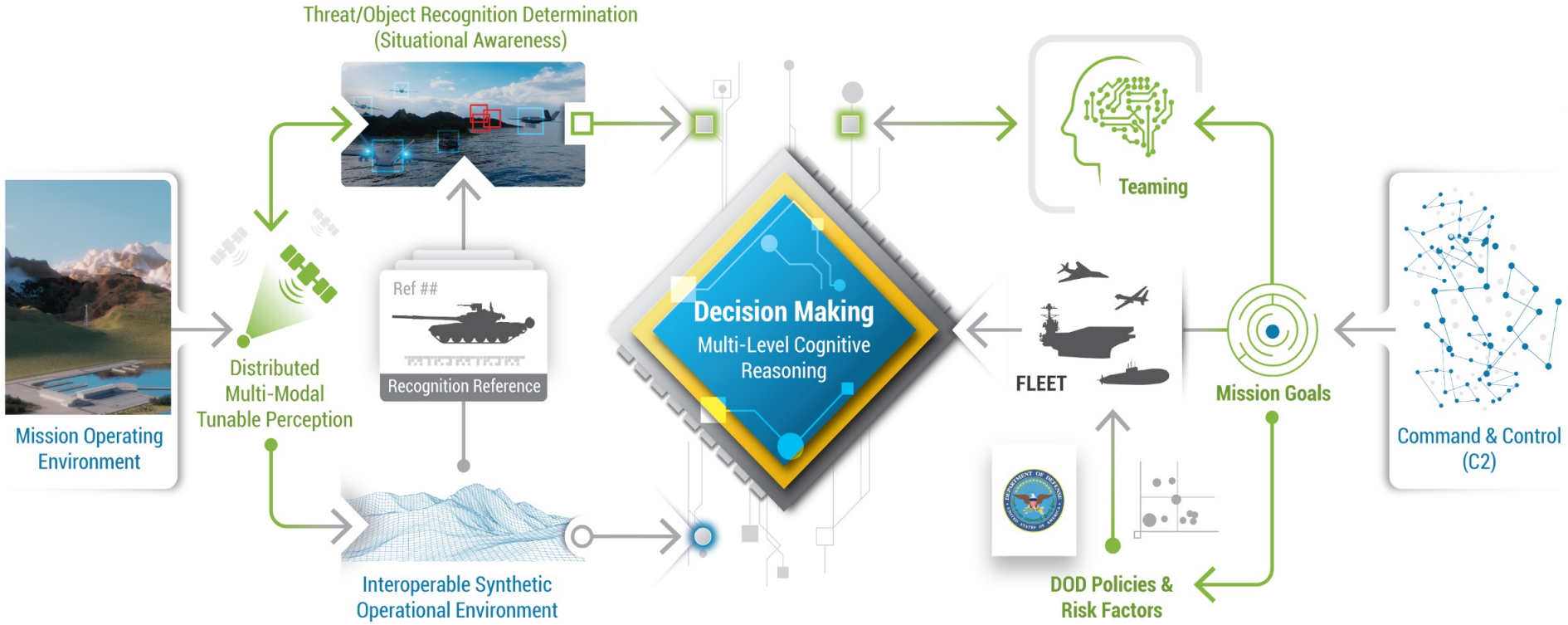
Vision: Robust, Resilient and Scalable Autonomous Operations with the Warfighter as the Center of Gravity



Developing the infrastructure to support AI development and acceleration across the R&D enterprise

Integrated Intelligence: A System of Systems Approach for Enhanced Capabilities

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

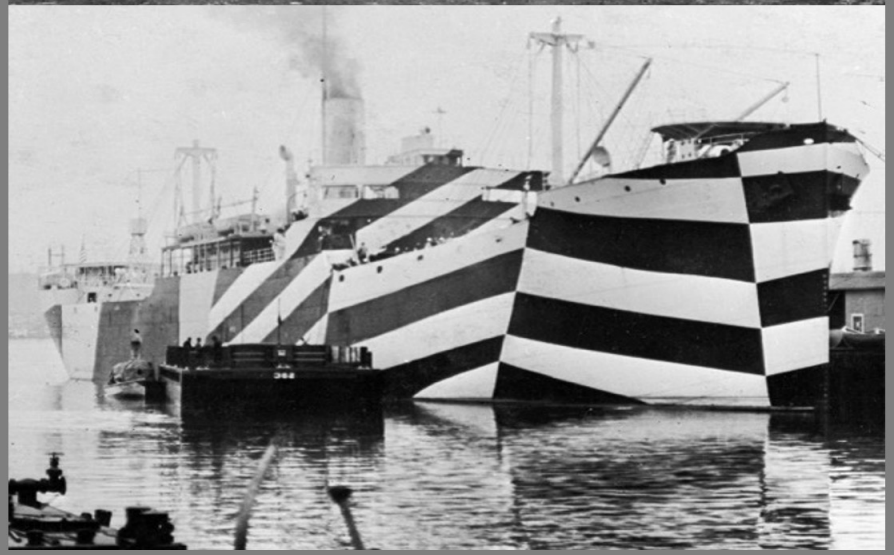


Context parsing in detecting and mitigating deception tactics is critical for survivability and should be considered up front!

Unveiling the Hidden Risks

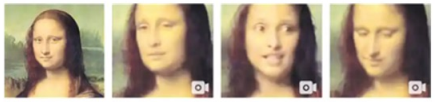
THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

Historical Deception



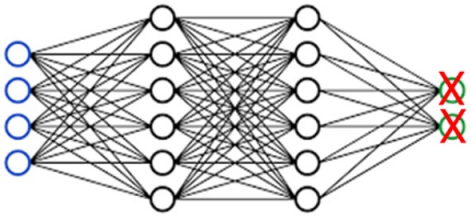
AI-Assisted Deception

Fool Human Perception



Zakharov et al. 2019

Fool Machine Learning Models



'Speed Limit 45'

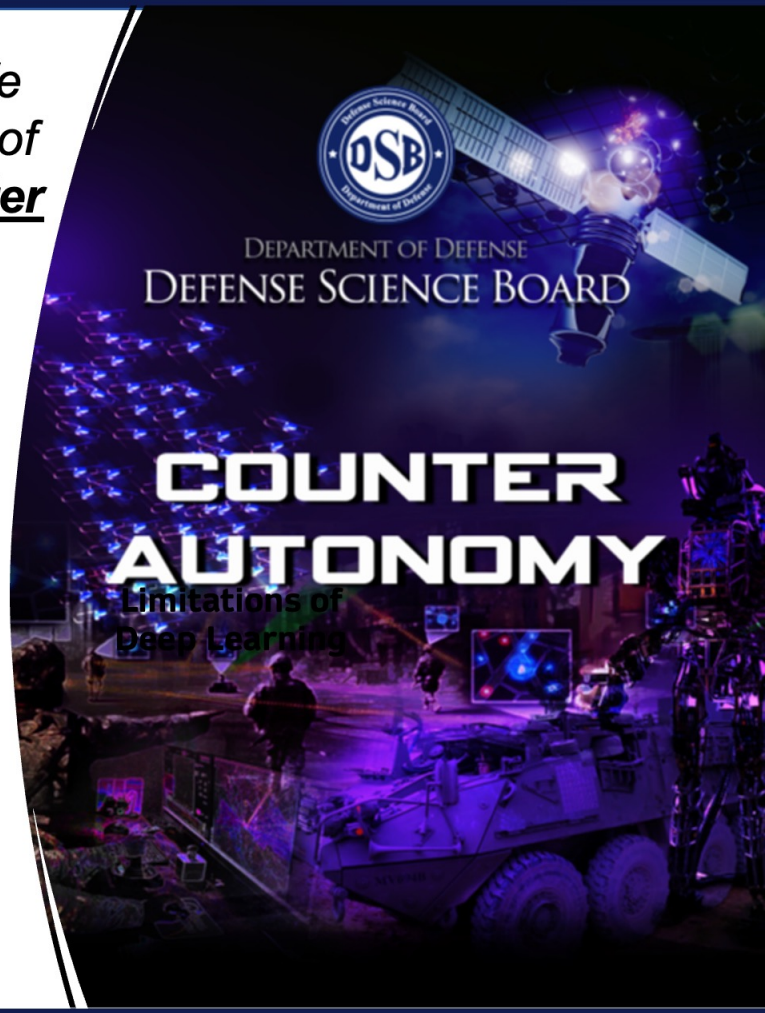
Eykholt et al. 2018

Enemy tactics assisted by AI will create a much more dynamic threat environment and will continuously change as their mission progress. This is very different from what our traditional tactics are good at…

The Brittle Truth…

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

AI systems can be brittle and prone to breaking. We have to understand the vulnerabilities and fragilities of AI-enabled systems to use them ***responsibly/counter***

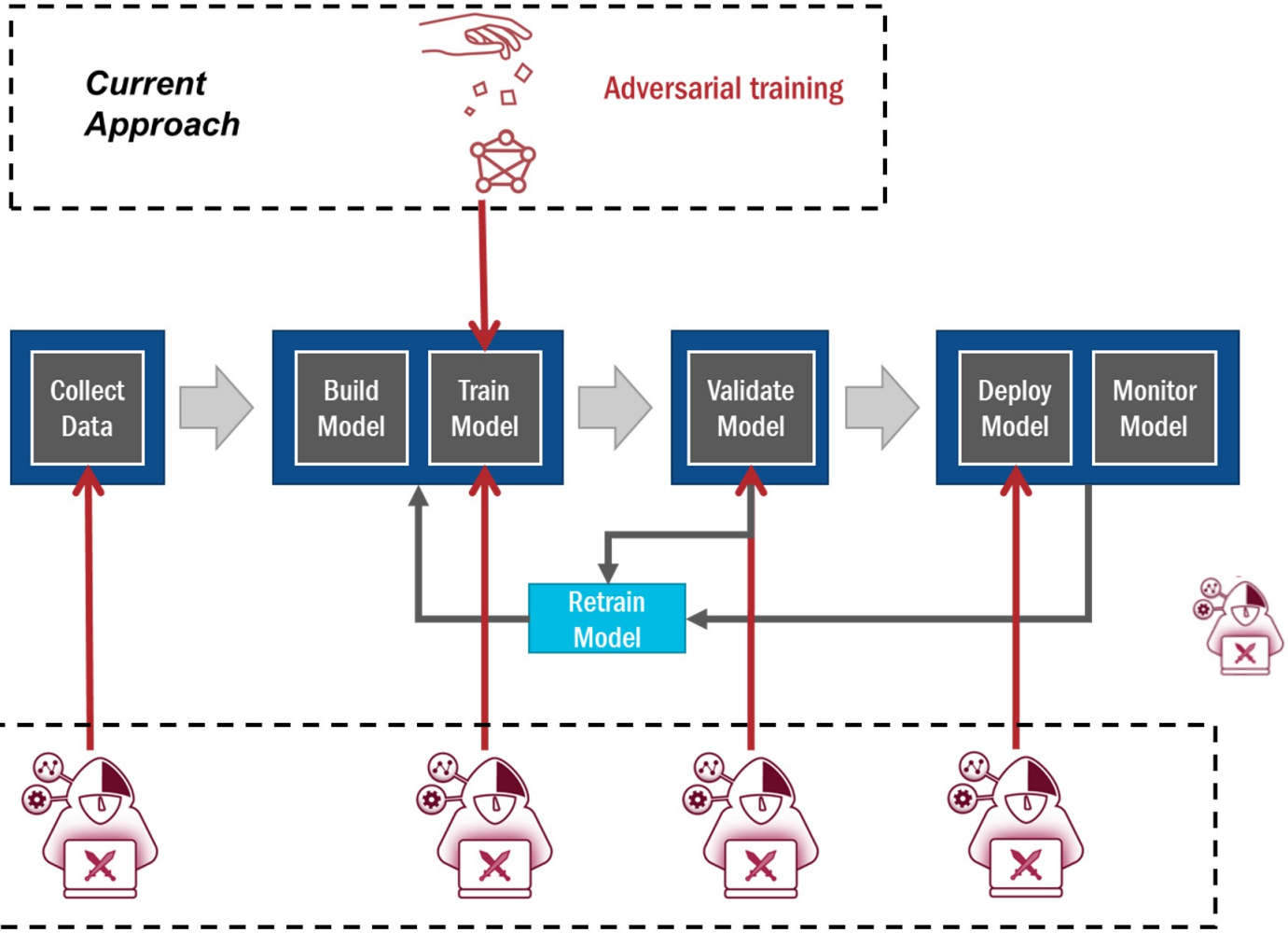


How do we protect our systems? How do we counter opposing autonomous systems?

Initiatives

Fortifying AI Resilience via Continuous Adversarial Testing and Red-Teaming

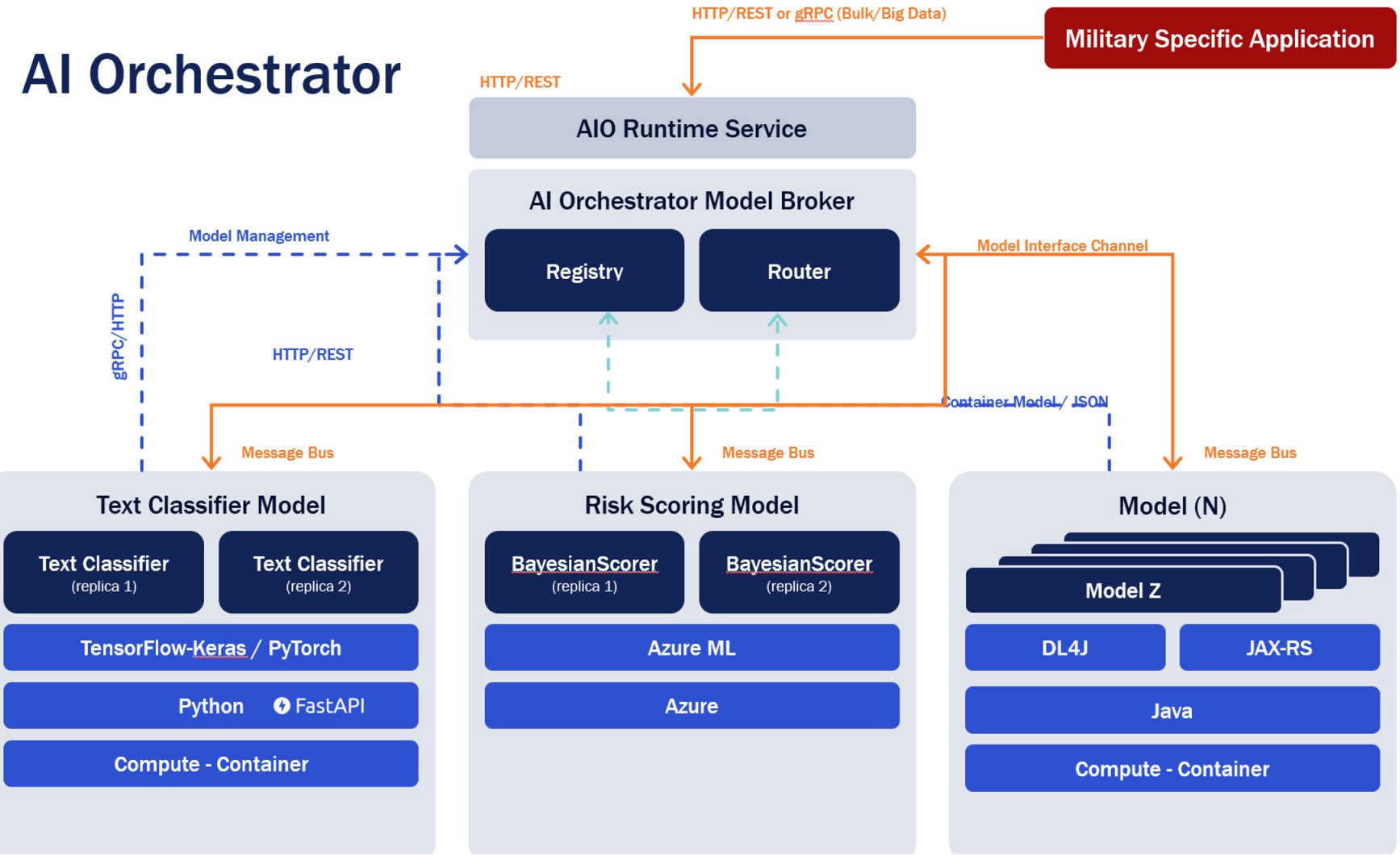
THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES



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

Harmonizing Deployed AI Models at Scale

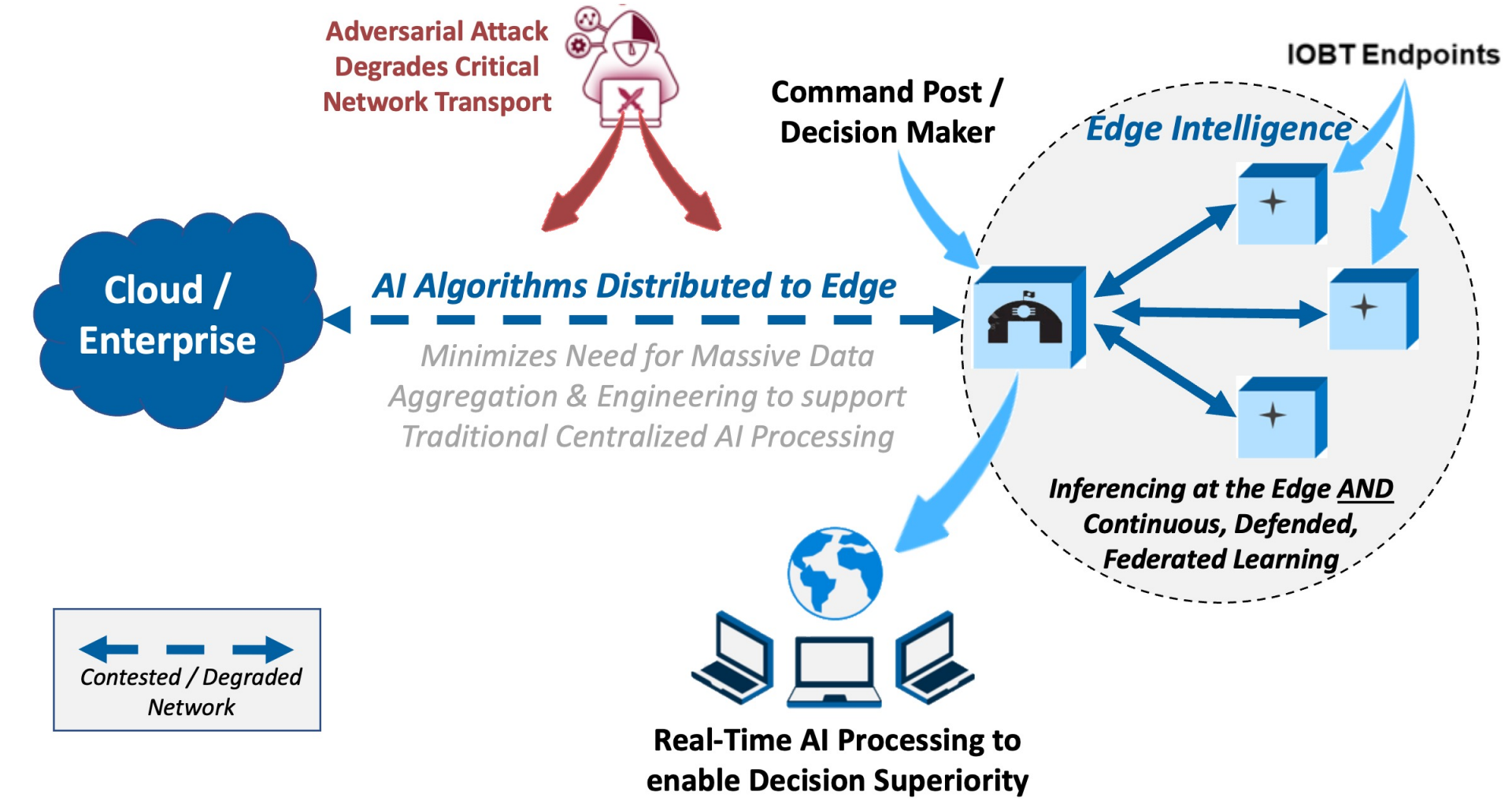
THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES



Scaling Harmony: Distributed Architecture for Model Orchestration and C2 on Demand

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

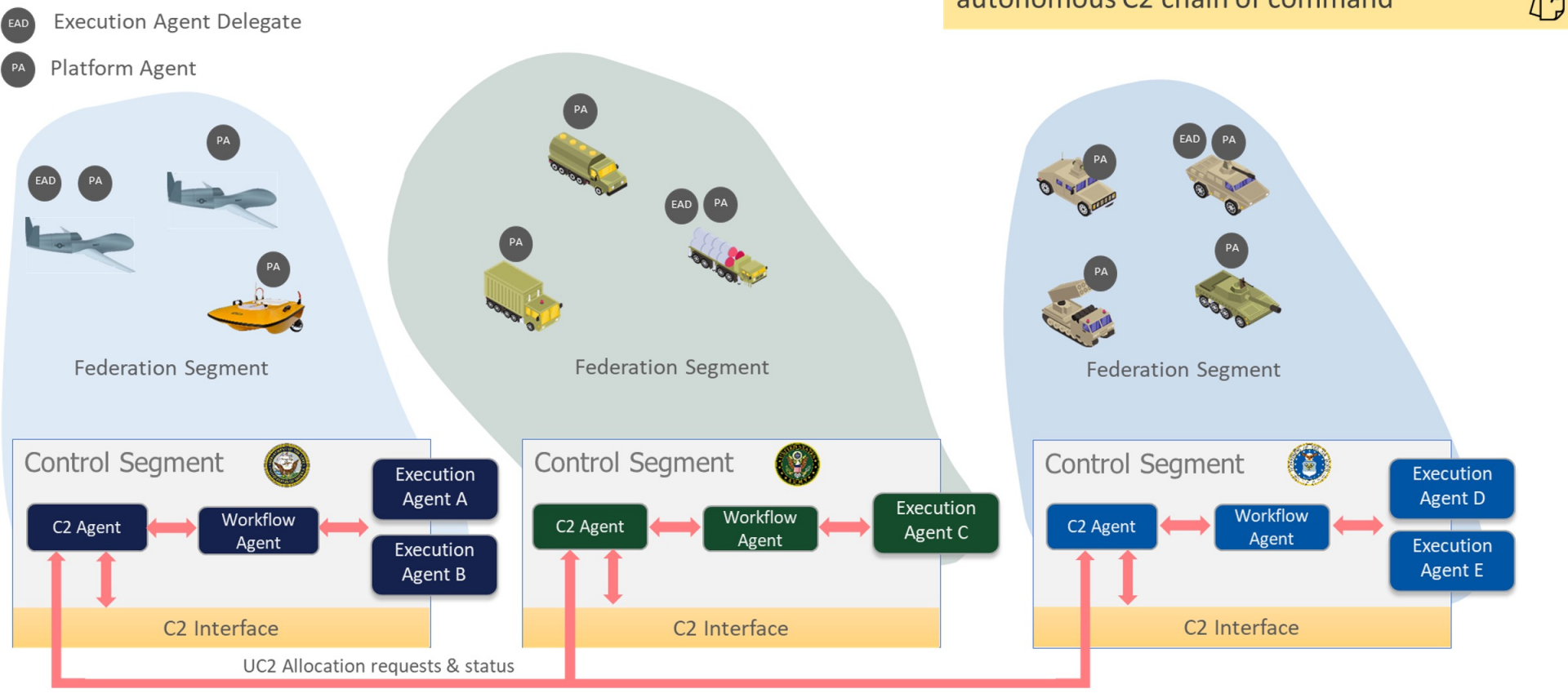
We must leverage advances in Distributed ML and Edge computing to enable Distributed AI architectures that minimize the dependency on massive data aggregation for AI-empowered Decision Superiority



Distributed, Hierarchical C2 and C2 on Demand Capabilities

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

Hierarchical agent structure implements an autonomous C2 chain of command



Cross-domain (air, ground, sea, underwater, etc.) uncrewed systems (UxS) are becoming critical enablers of modern warfare

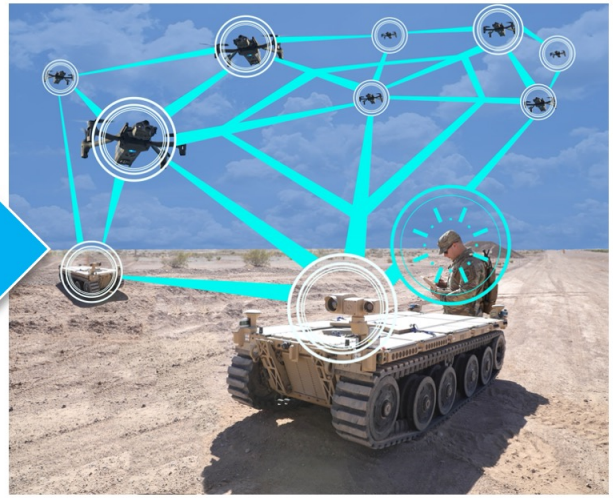
Towards Autonomous Networks of Autonomous Systems: The Power of Human-AI Synergy

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

Human oversight/policy

Many controlling many UxS

One controlling many UxS



Autonomous viral networks of autonomous systems equipped with Autonomous C2 systems, common operational database (COD), AI Orchestrator to support integrated operations across domains

Note: An autonomous viral network should be able to rebuild, reconnect and reorganize

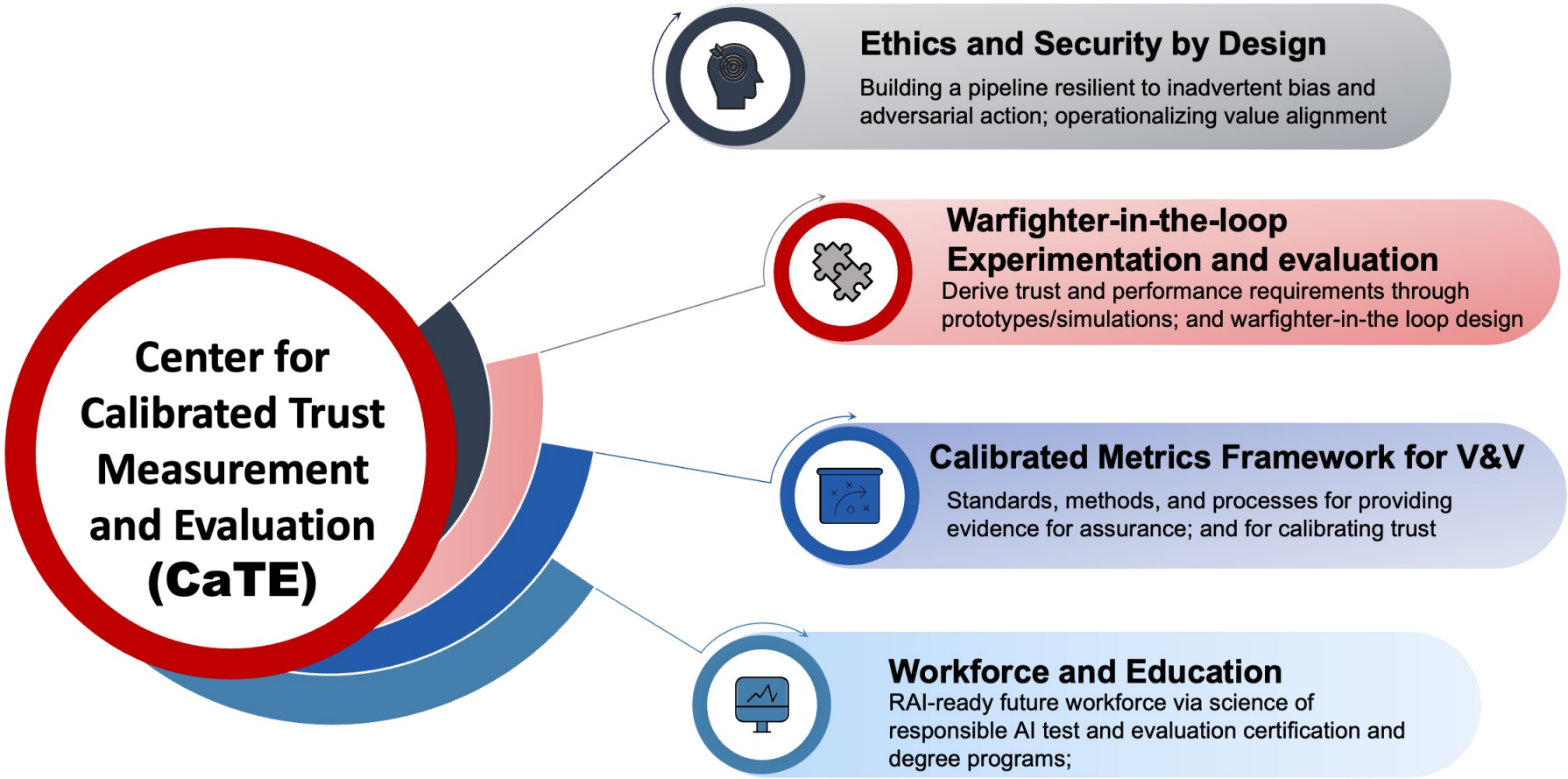
DoD Responsible AI (RAI) Foundational Tenets

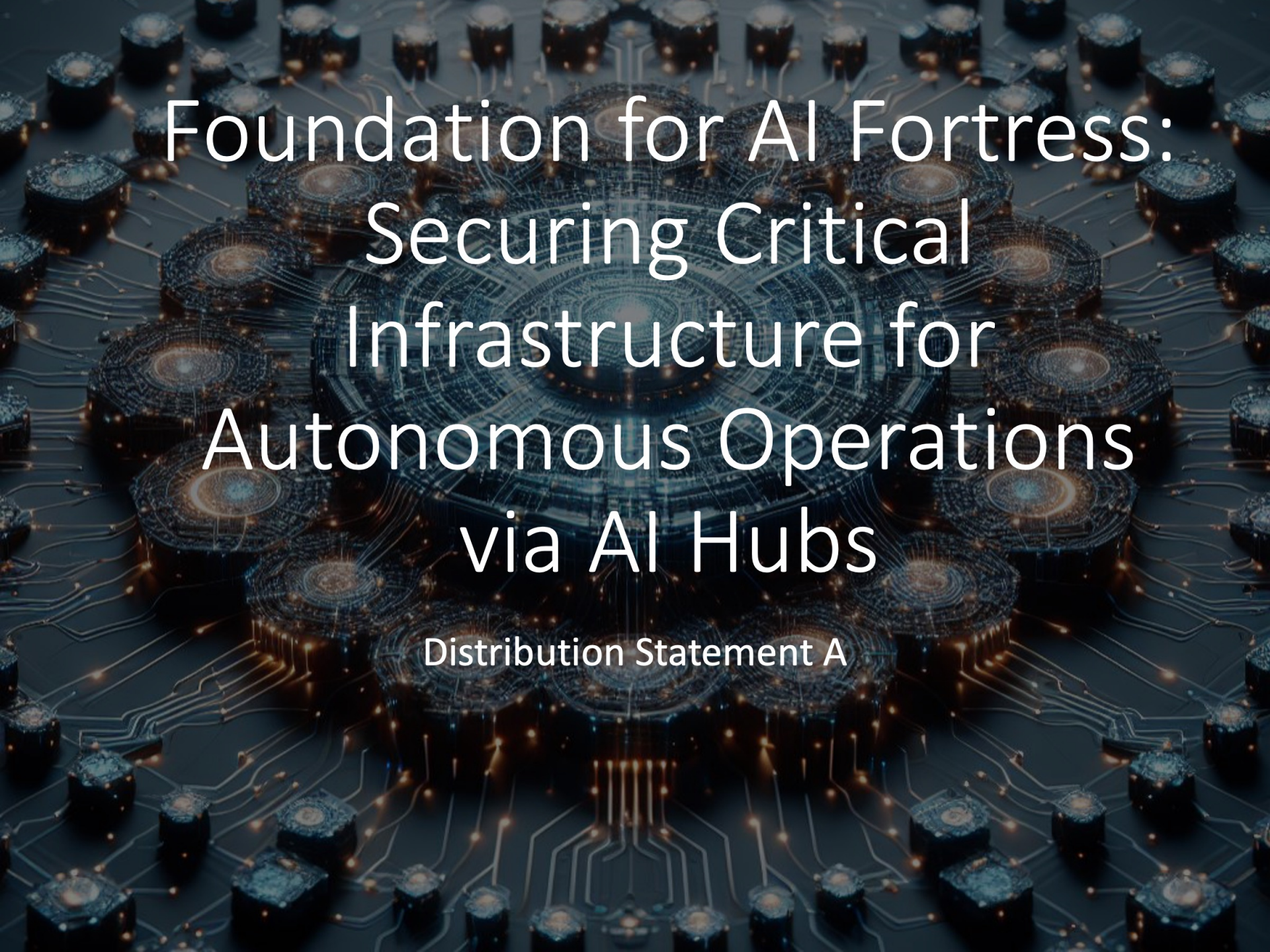
- **RAI Governance:**
  - Ensure disciplined governance structure and processes at the Component and DoD-wide levels for oversight and accountability and clearly articulate DoD guidelines and policies on RAI and associated incentives to accelerate adoption of RAI within the DoD.
- **Warfighter Trust:**
  - Ensure warfighter trust by providing education and training, establishing a test and evaluation and verification and validation (TE/VV) framework that integrates real-time monitoring, algorithm confidence metrics, and user feedback to ensure trusted and trustworthy AI capabilities.
- **AI Product and Acquisition Lifecycle:**
  - Develop tools, policies, processes, systems, and guidance to synchronize enterprise RAI implementation for the AI product throughout the acquisition lifecycle through a systems engineering and risk management approach.
- **Requirements Validation:**
  - Incorporate RAI into all applicable AI requirements, including joint performance requirements established and approved by the Joint Requirements Oversight Council, to ensure RAI inclusion in appropriate DoD AI capabilities.
- **Responsible AI Ecosystem:**
  - Build a robust national and global RAI ecosystem to improve intergovernmental, academic, industry, and stakeholder collaboration, including cooperation with allies and coalition partners, and to advance global norms grounded in shared values.
- **AI Workforce:**
  - Build, train, equip, and retain an RAI-ready workforce to ensure robust talent planning, recruitment, and capacity- building measures, including workforce education and training on RAI.

Symphony of Minds: Warfighter Trust

THE OFFICE OF THE ASSISTANT SECRETARY 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)




Foundation for AI Fortress: Securing Critical Infrastructure for Autonomous Operations via AI Hubs
Distribution Statement A

Distribution Statement A

Bridging the Gap: Uniting Research & Operations

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

R&D efforts often are supported by mission focused programs with immediate operational goals or purely with research funding that is not sustained, resulting in short-term storage for data, limiting of sustained and common data, development, and simulation platforms, and limited connection within the development pipeline.

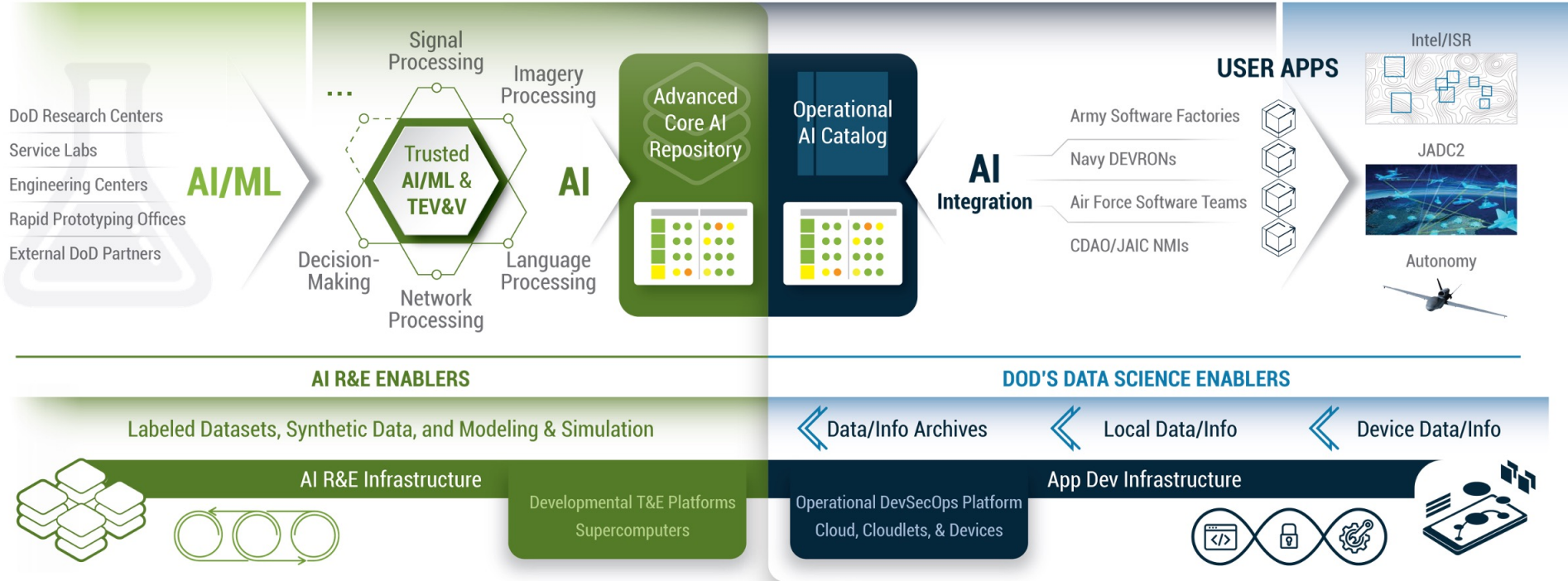
DoD's AI Workforce and Envisioned REDO Pipeline (OUSD(R&E) & (OUSD(A&S))

R AI/ML Research Programs

E Engineering Programs

D Acquisition/Development Programs

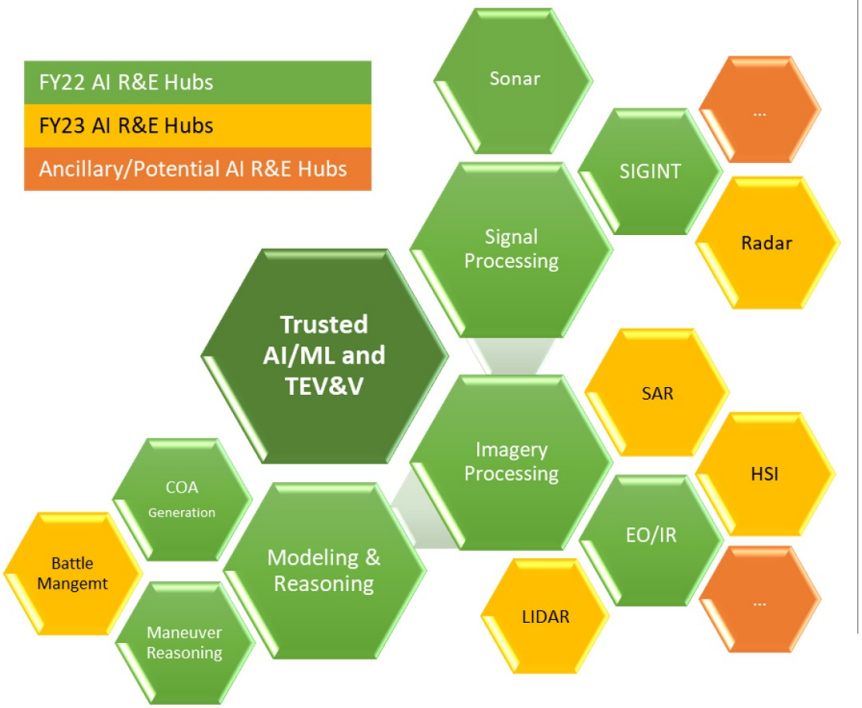
O O&M Programs

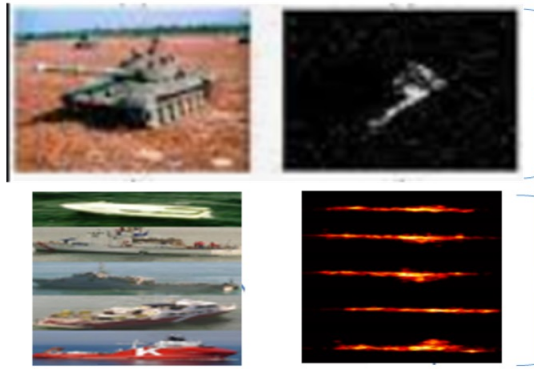


Data-Specific and Multi-Data Integration

THE OFFICE OF THE ASSISTANT SECRETARY 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***.

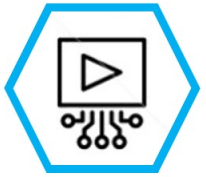
AFRL USSF DEVCOM ARMY RESEARCH LABORATORY NAVAL RESEARCH LABORATORY WASHINGTON, DC

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

Unlocking Multi-Modal Potential: Leveraging Foundation Models for New Operational Capabilities

Large Language Model

Generative Voice Model



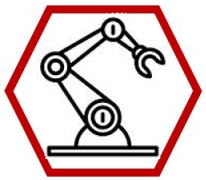
Generative Video Model

Generative Image Models

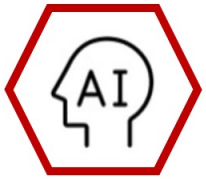
Generative Code Model

Generative Signal Models

Generative Scientific Models



Generative Autonomy Models



Generative Decision Models

Multimodal Foundation Model

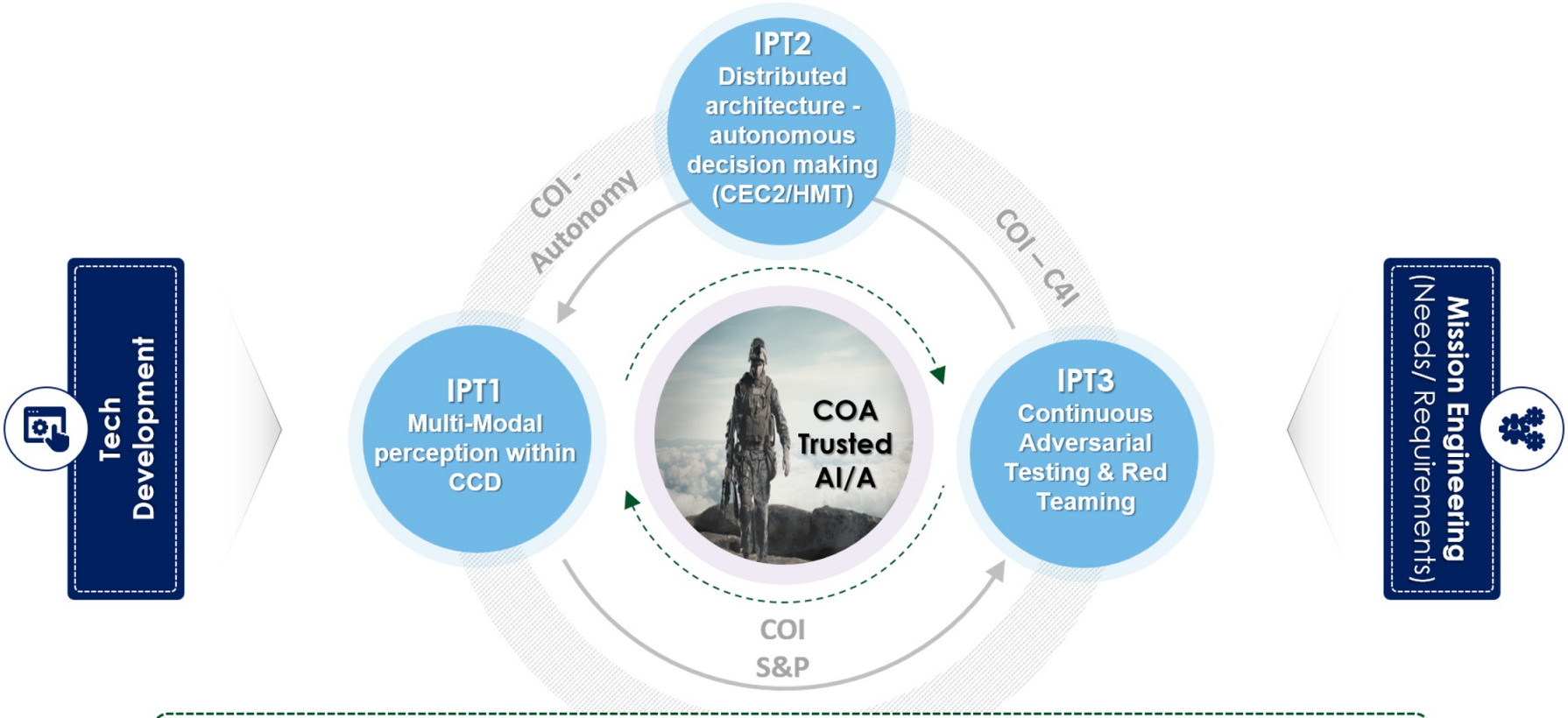
- Rapid autonomy/ robotics dev
- Revolutionary HMT
- Next gen UIs
- New wargaming/ COA approaches
- Rapid coding
- Cyber warfare
- Improved psy-ops
- Intel analysis process improvement
- Multimodal synthetic data
- Increased research efficiencies
- ….
- Move toward AGI

Initial creation of unimodal generative models developed and maintained in each hub will over time lead to multimodal systems that will create new operational capabilities for DoD personnel and increase efficiency for S&T

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

THE OFFICE OF THE ASSISTANT SECRETARY 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

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