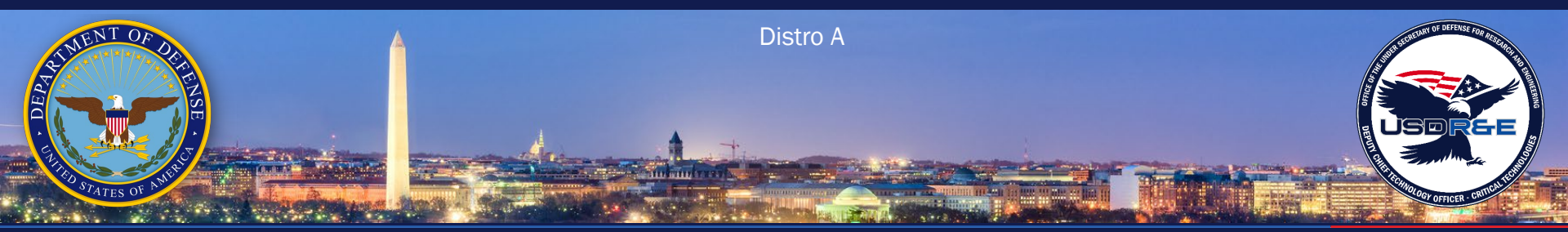


Cognitive Autonomy

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: NDIA Emerging Technologies Conference

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

Role of AI in Military Systems

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 at speed of battle*
- *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 Recent advancements in AI is not in the DOD…How can we leverage to address DOD problems rapidly and responsibly?

Challenges we're tackling…

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

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

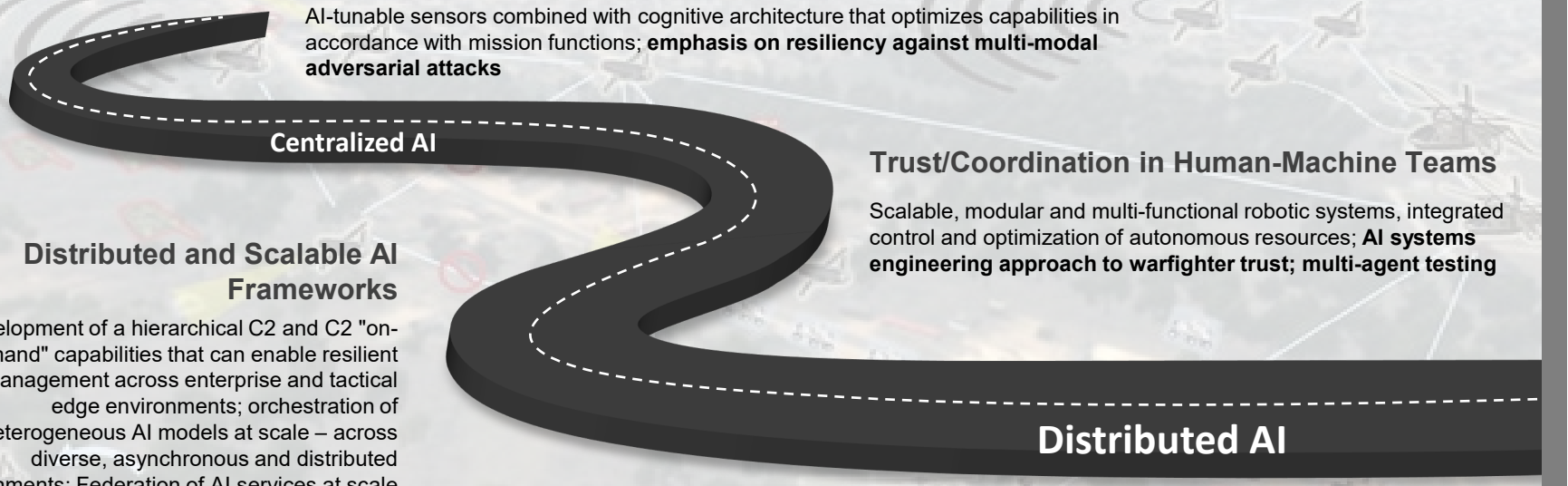
Roadmap at a Glance

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

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

Reconfigurable, Multi-modal, Trusted Perception (moving from passive to interactive)

AI-tunable sensors combined with cognitive architecture that optimizes capabilities in accordance with mission functions; **emphasis on resiliency against multi-modal adversarial attacks**



Trust/Coordination in Human-Machine Teams

Scalable, modular and multi-functional robotic systems, integrated control and optimization of autonomous resources; **AI systems engineering approach to warfighter trust; multi-agent testing**

Distributed and Scalable AI Frameworks

Development of a hierarchical C2 and C2 "on-demand" capabilities that can enable resilient battle management across enterprise and tactical edge environments; orchestration of heterogeneous AI models at scale – across diverse, asynchronous and distributed environments; Federation of AI services at scale

Cross-Echelon, Resilient Autonomous (Viral) 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

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

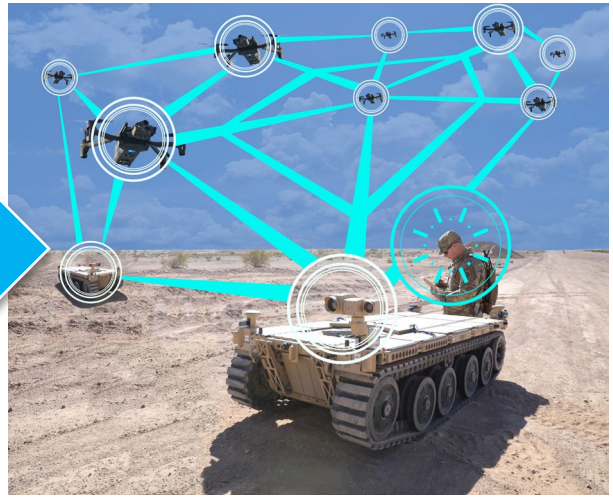
Towards Autonomous Networks of Autonomous Systems: The Power of Human-AI Synergy

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

Human ethical and policy oversight

Many controlling many UxS

One controlling many UxS



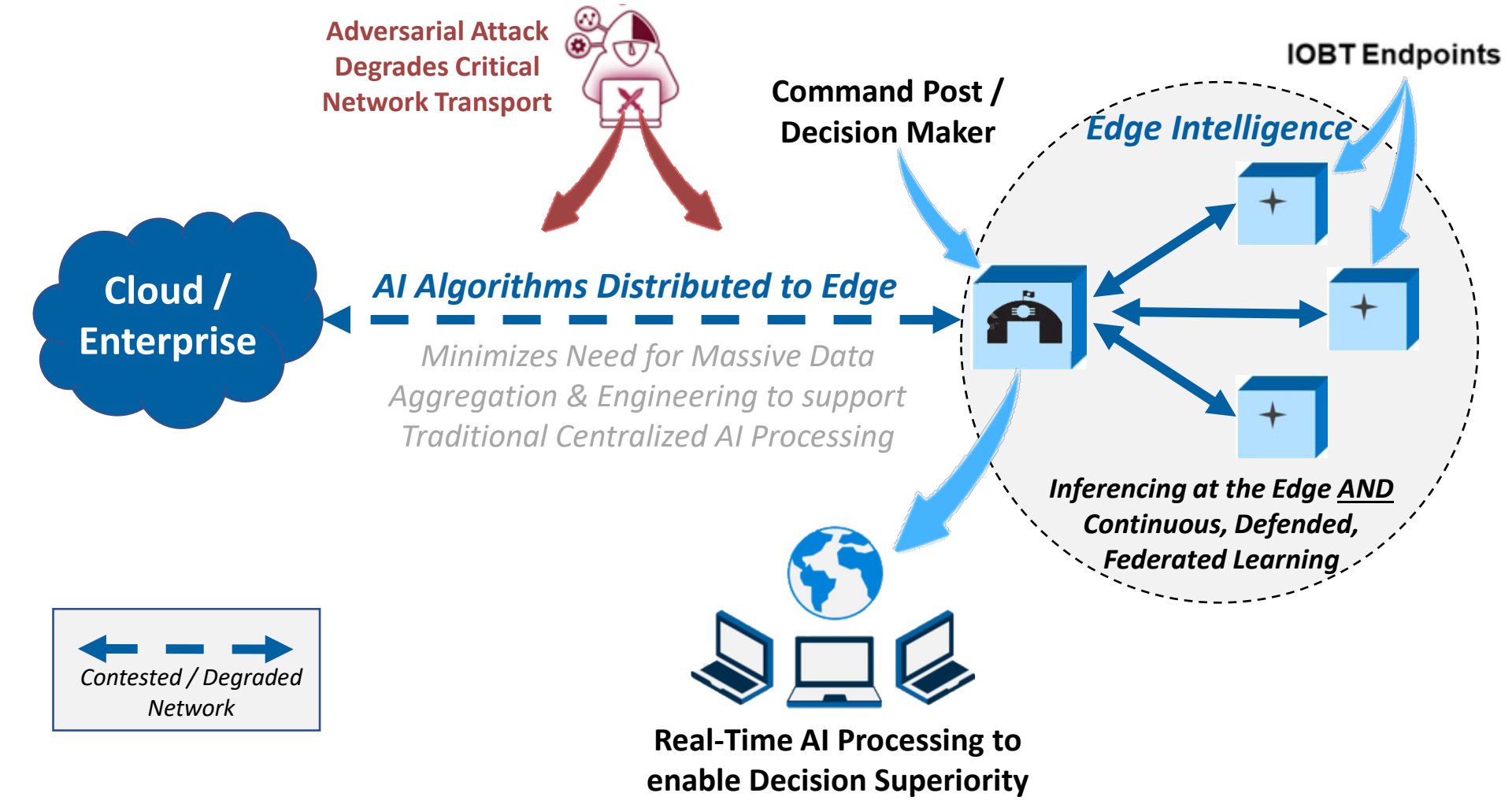
Autonomous viral networks of autonomous systems equipped with Autonomous C2 systems, common operational database (COD), AI Orchestrator to support integrated operations across domains

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

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

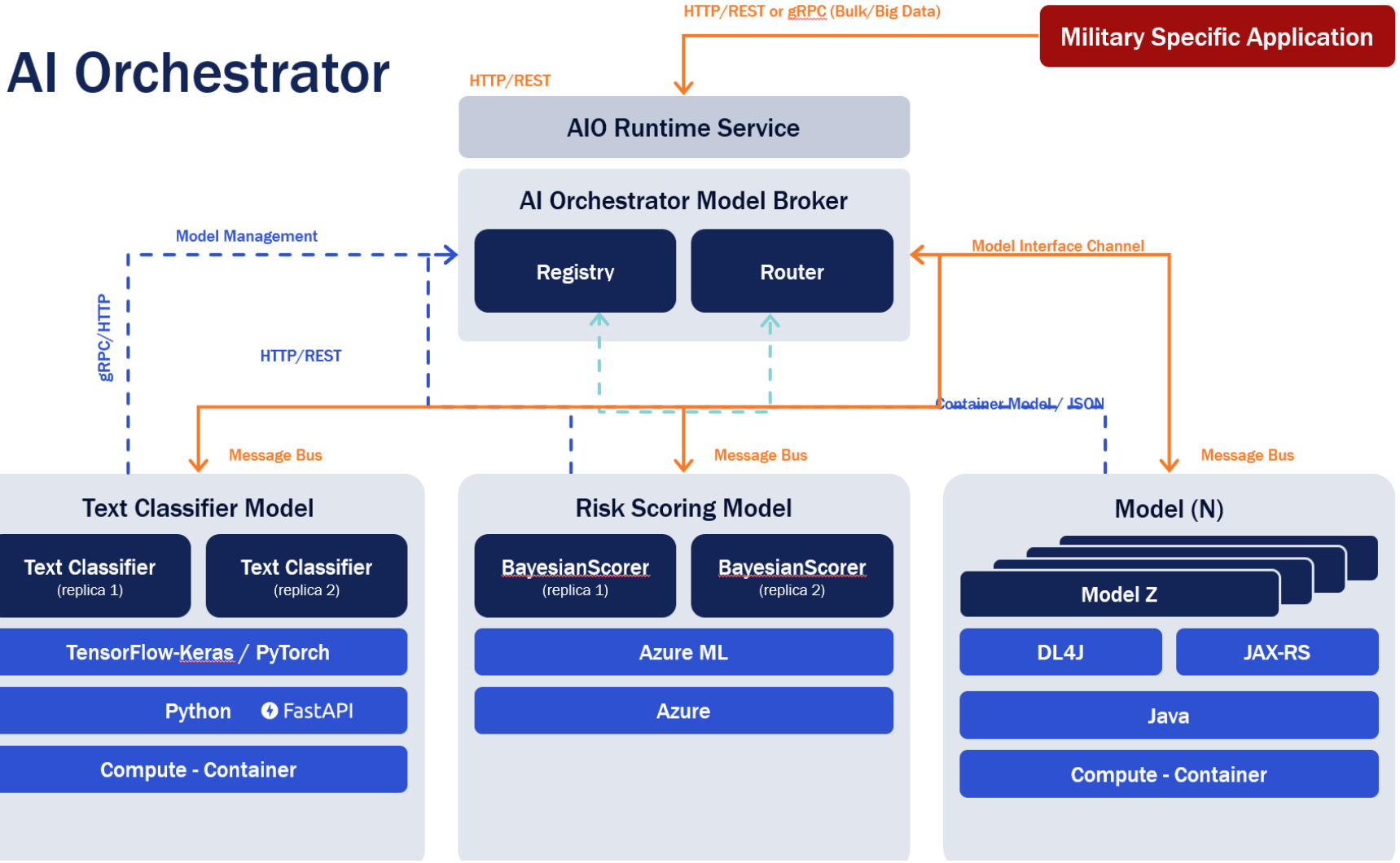
THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

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



Harmonizing Deployed AI Models at Scale

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES



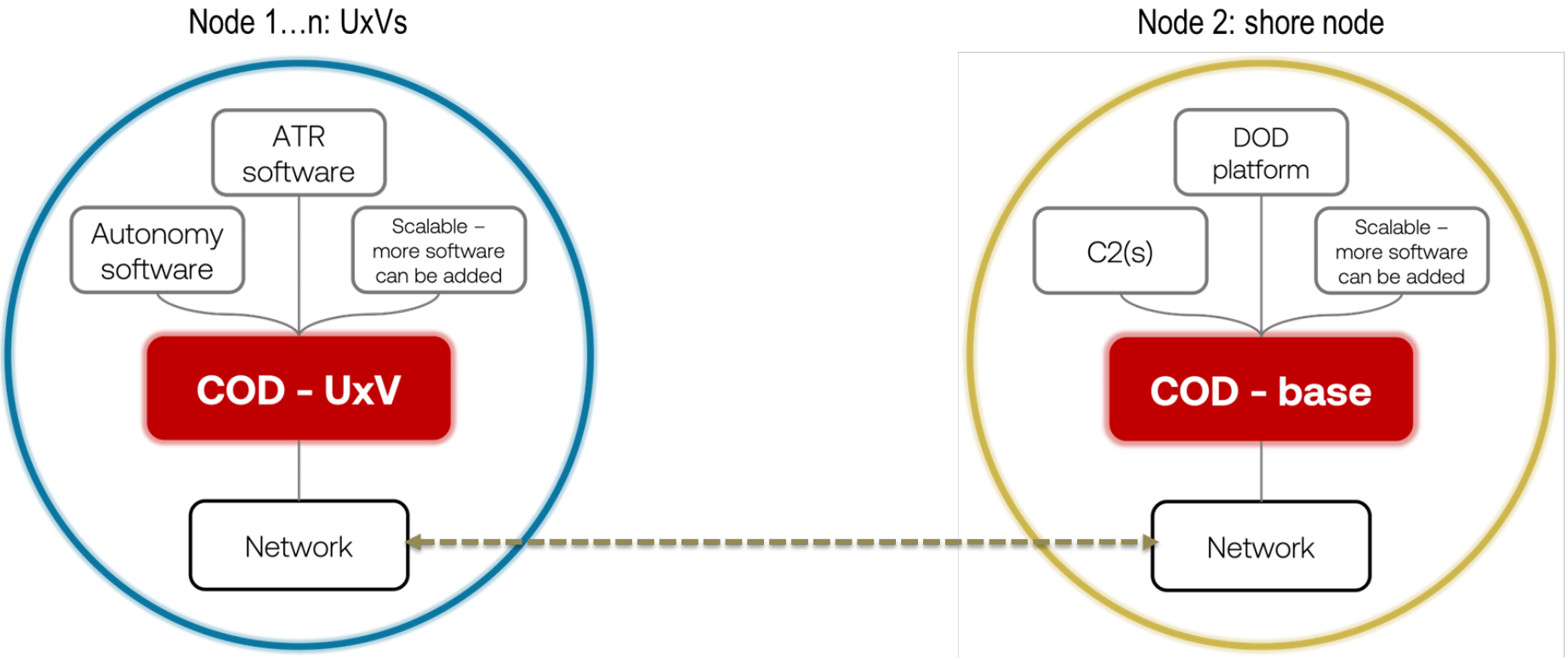
Distributing Data across Nodes at Scale

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

COD is an edge sync service that logically connects and distributes data across different nodes: UxVs, shore nodes, mobile devices, etc.

The software integrated into nodes is scalable and swappable to meet changing mission requirements

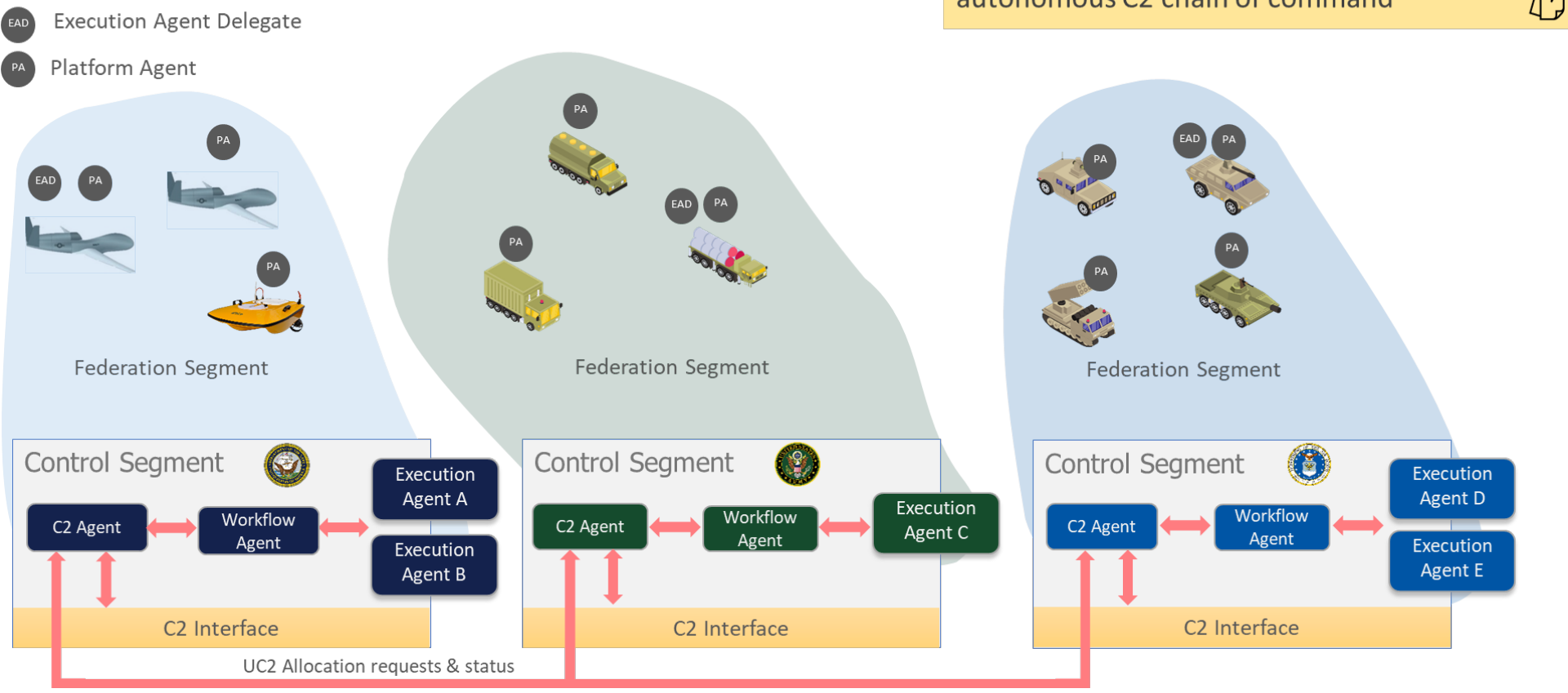
Nodes can be provided by different vendors/partners for multi-domain operations



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

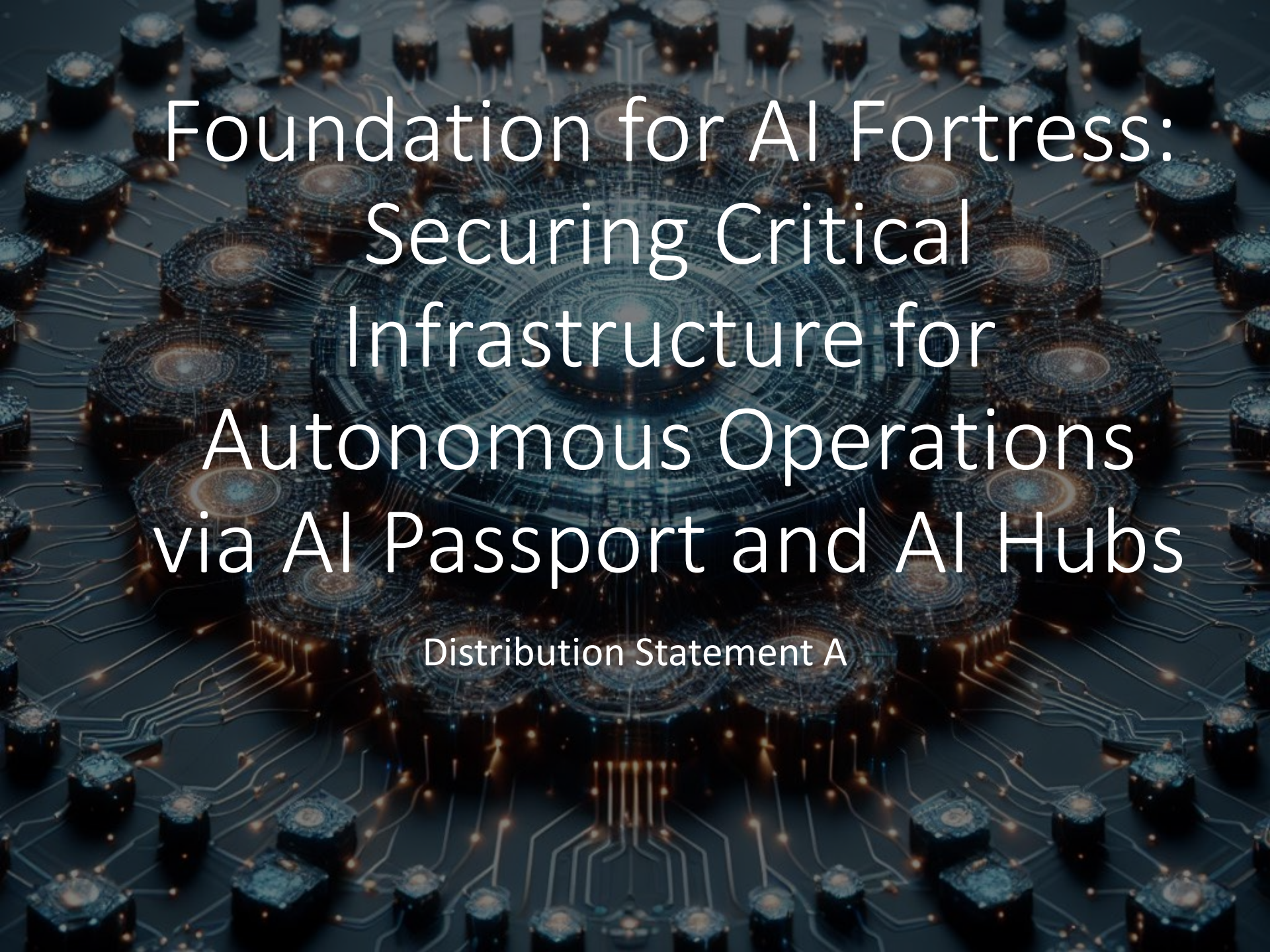


Cross-domain (air, ground, sea, underwater, etc.) uncrewed systems (UxS) are becoming critical enablers of modern warfare

As we expand implementation of AI, risks and complexities grow…

We need to have critical infrastructure for AI

- Taking a whole-of-nation approach to federating AI
- Developing the testing components - assessments of trustworthiness/confidence and ensure robustness of models



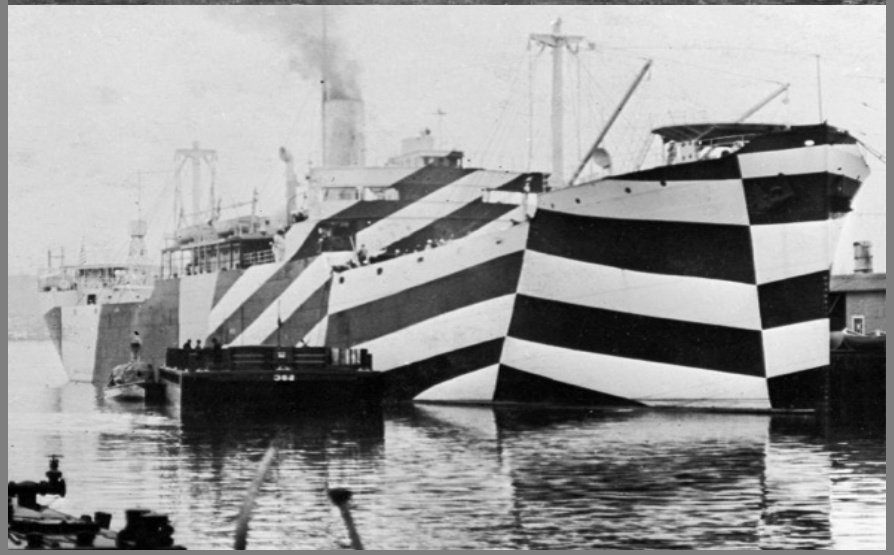
Foundation for AI Fortress: Securing Critical Infrastructure for Autonomous Operations via AI Passport and AI Hubs

Distribution Statement A

Unveiling the Hidden Risks

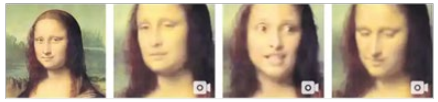
THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

Historical Deception



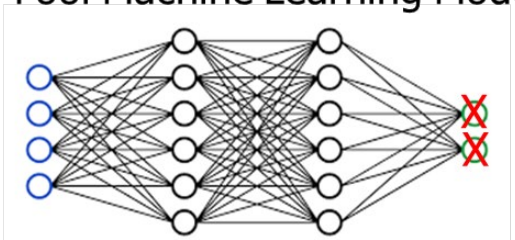
AI-Assisted Deception

Fool Human Perception



Zakharov et al. 2019

Fool Machine Learning Models



'Speed Limit 45'

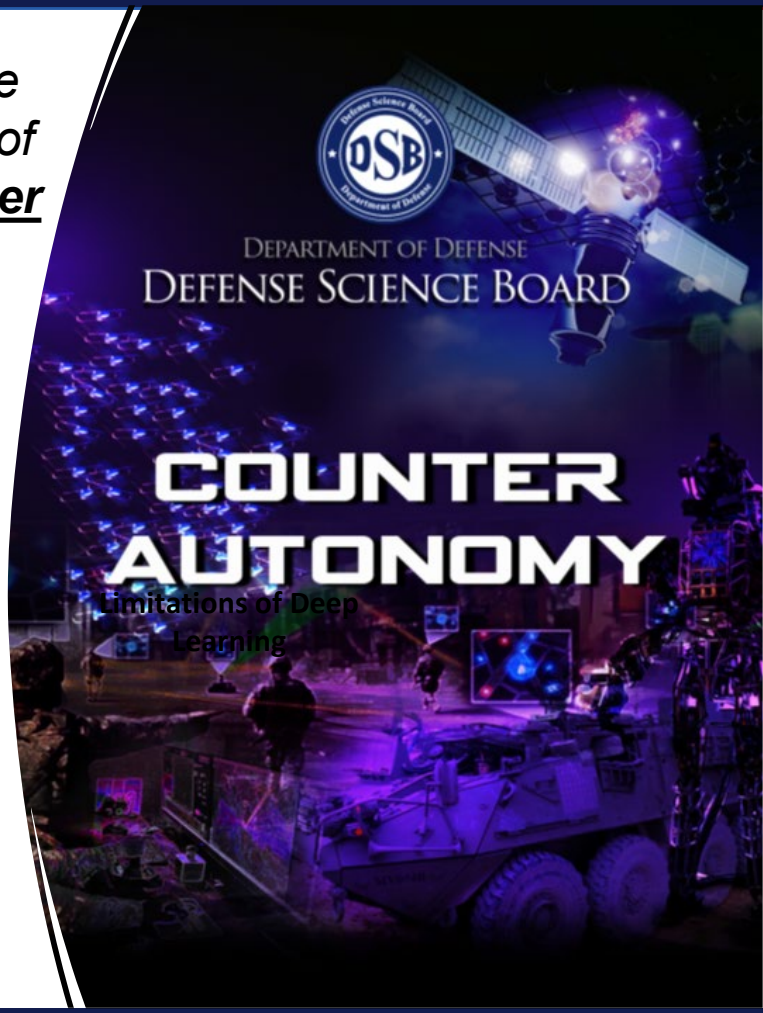
Eykholt et al. 2018

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

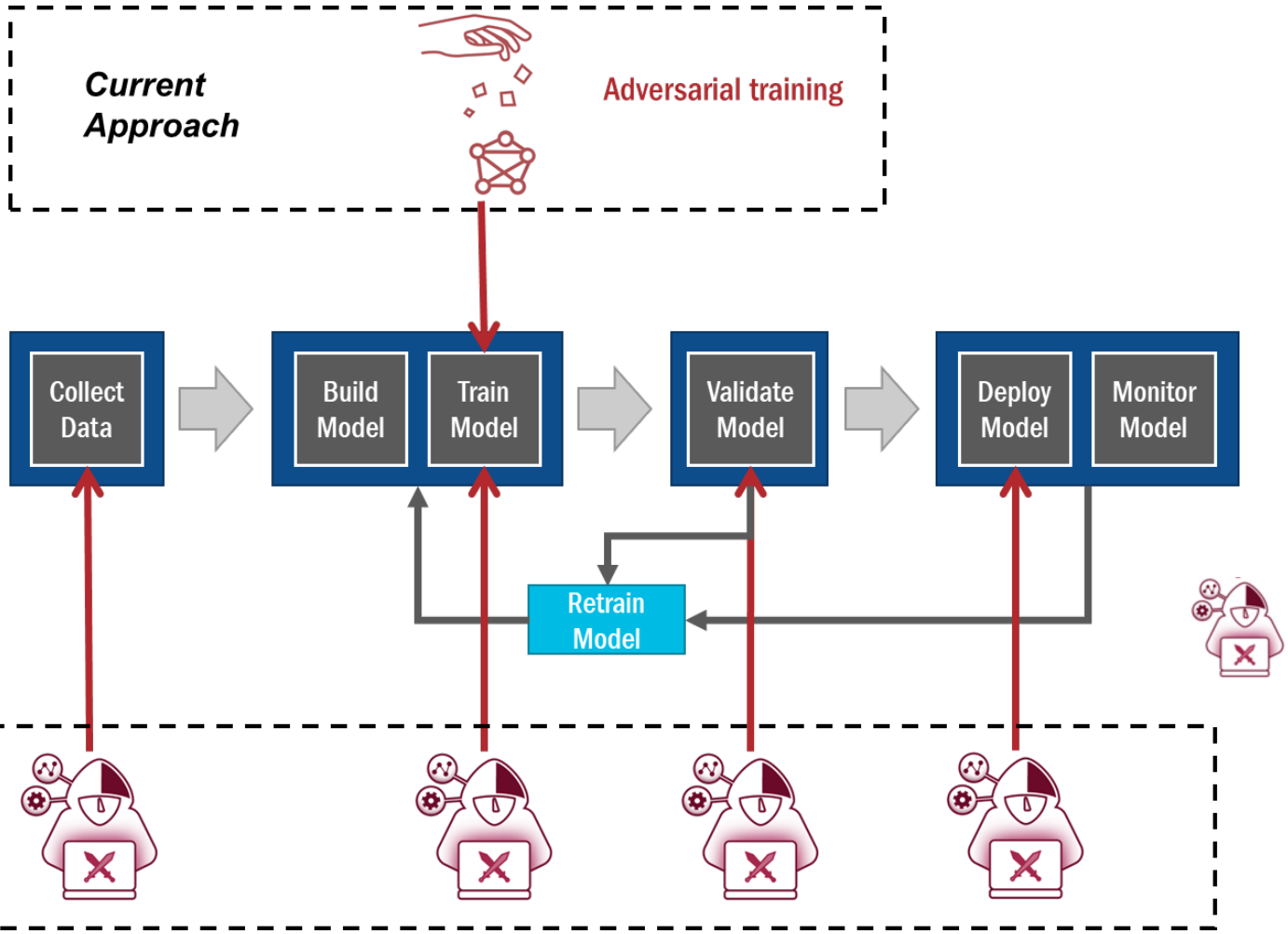
AI systems can be brittle and prone to breaking. We have to understand the vulnerabilities and fragilities of AI-enabled systems to use them ***responsibly/counter***



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

Fortifying AI Resilience via Continuous Adversarial Testing and Red-Teaming

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES



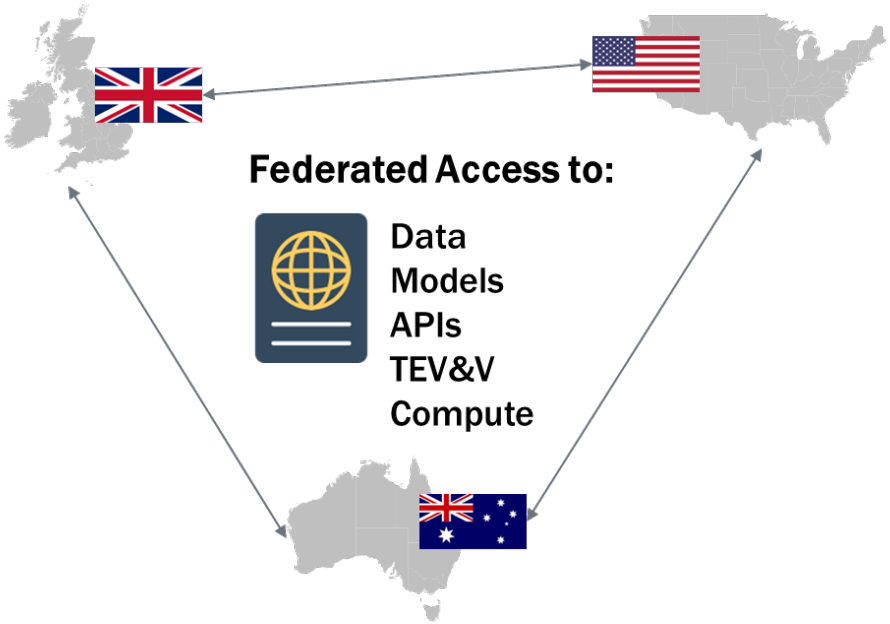
A Testbed for Continuous adversarial testing and red-teaming approaches must be applied throughout the system lifecycle – from development to deployment

Federated Access to Shared AI Components

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

AUKUS AI Passport

aim to maintain compliance with existing standards (e.g. NATO FMN Spec)



Interoperability Standards

- Data Labelling Ontology
- Open Model Standards
- Containerisation
- Data & Model Cards
- Runtime Environment

Design Principles

- Expect (& Embrace) Heterogeneity
- Modular Components by Design
- "Opt-in" – Explicit controls over sharing component AI tools/capabilities

AI Passport: New Distributed AI federation framework which enables multi-party software co-development enabling collaboration between diverse toolchains by sharing metadata that comply with AI Passport Metadata Standards

AI Passport Key Components

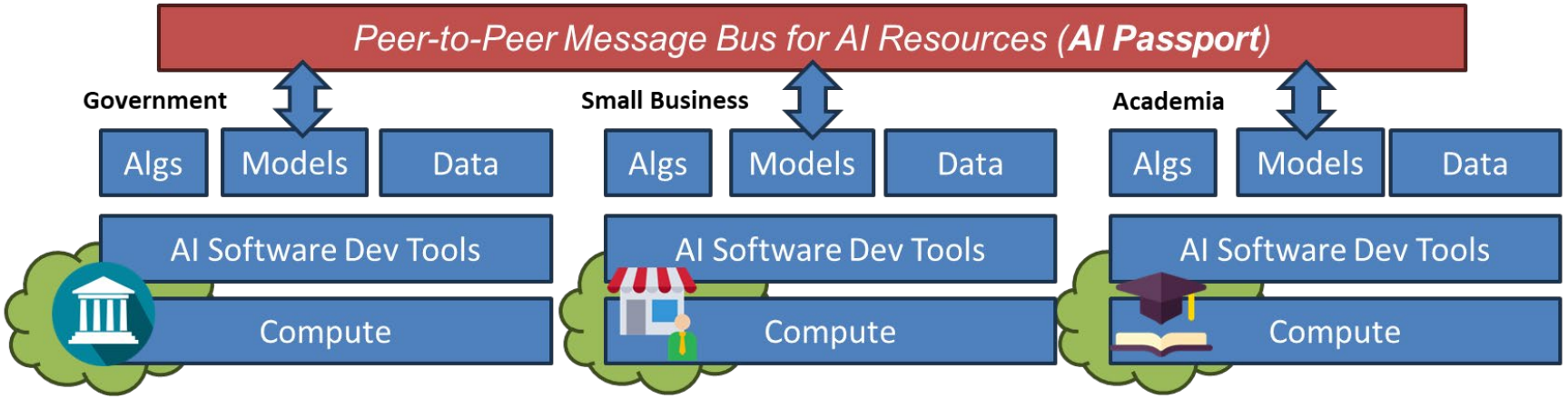
THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

Key Project Components

- AI Passport – Federated control & exchange of government & industry AI capabilities
- Data, model repository –Sharing of datasets and AI models to drive co-development
- Data & Model Cards – Virtual Networked Repository of data and model capabilities
- Evaluation framework – Enable test & evaluation of government & industry AI capabilities

Design Principles

- Expect (& Embrace) Heterogeneity
- Modular Components by Design
- "Opt-in" – Explicit controls over sharing component AI tools/capabilities

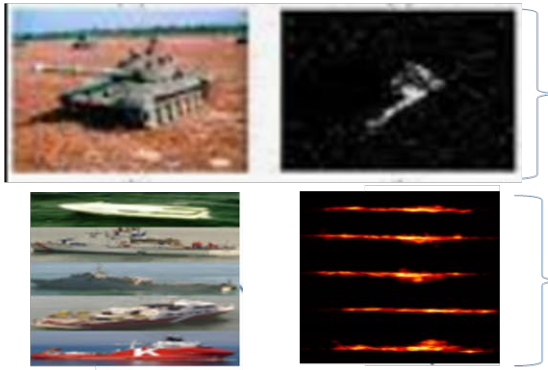


A focus on interoperability of AI capabilities across joint and multi-national boundaries – An AI partner network to leverage capabilities across coalition AI partner/industrial base

Towards Multi-Phenomenology 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***.

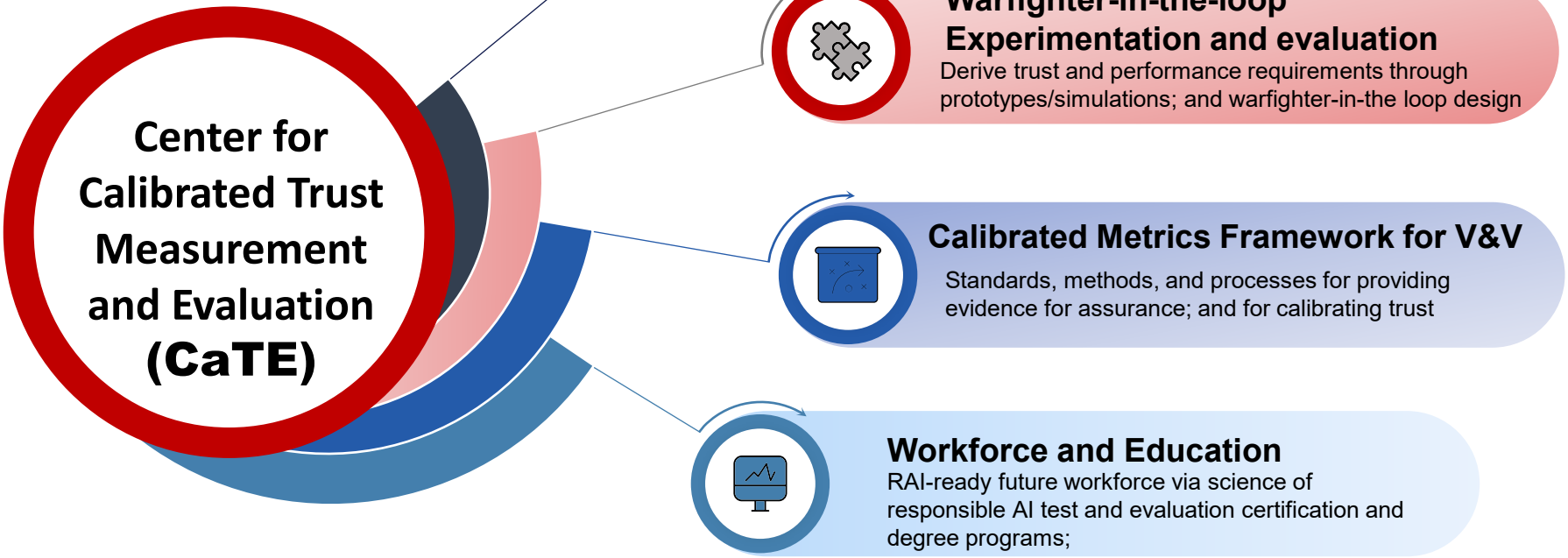
AFRL USSF DEVCOM ARMY RESEARCH LABORATORY NAVAL RESEARCH LABORATORY WASHINGTON, DC

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

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)



Connect to AI Passport – Call to Industry

THE OFFICE OF THE ASSISTANT SECRETARY FOR CRITICAL TECHNOLOGIES

Ways to connect:

Integrate your tools to comply with AI Passport API and Metadata Specification
Connect to already established AI Passport node
(Optional) Standup your own AI Passport node
(Optional) Federate your own AI Passport nodes with other nodes

<u>Experimental</u> AI Passport nodes:

HPC System – Shared with DOD and international partners
DOD AFRL DREN systems – Shared with DOD government / contractors
Future deployments – TBD

Access:

Available to Download at: **https://gitlab.dle.afrl.af.mil/aaic/ai-passport/ai-passport-rabbitmq (DOD/Contractors)**
For information and non-dod access email please email : **afrl.ris.air@us.af.mil**