



# Operational Test & Evaluation Considerations for Emerging Technologies

**Honorable Dr. Douglas C. Schmidt**  
Director, Operational Test & Evaluation,  
Office of the Secretary of Defense

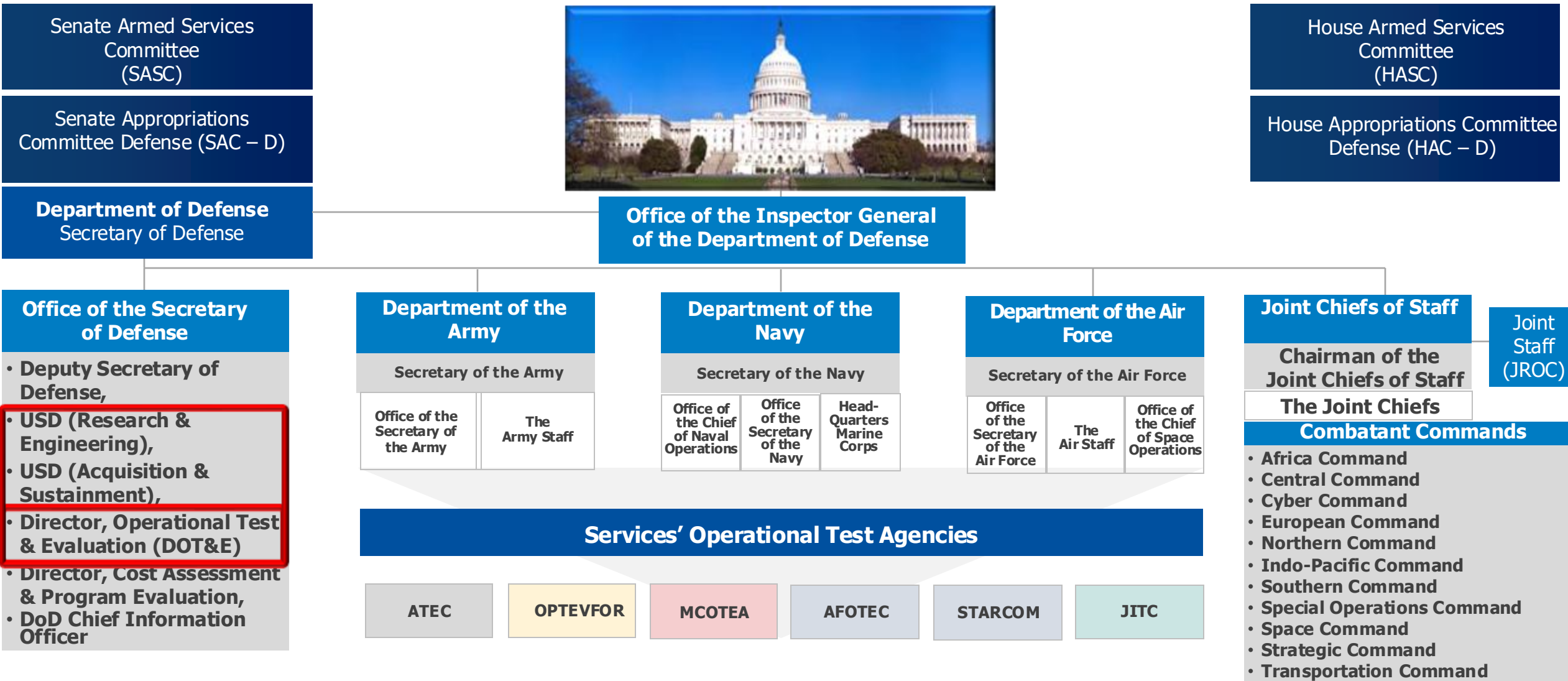


# Operational & Live Fire Test & Evaluation (OT&E/LFT&E)



**Mission: Evaluate the operational *effectiveness, suitability, survivability,* & (when necessary) *lethality* to defend our homeland & prevail in conflict**

# DOT&E in Context of the Department of Defense



**Acronyms in slide:** SASC – Senate Armed Services Committee; SAC-D – Senate Appropriations Committee Defense; HASC – House Armed Services Committee; HAC-D – House Appropriations Committee Defense; USD – Under Secretary of Defense; JROC – Joint Requirements Oversight Council; ATEC – Army Test & Evaluation Command; OPTEVFOR – Operational Test & Evaluation Force; MCOTEA – Marine Corps Operational Test & Evaluation Activity; AFOTEC – Air Force Operational Test & Evaluation Center

# What is the Director, Operational Test & Evaluation (DOT&E)?



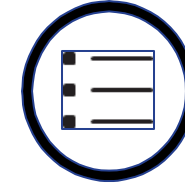
**Policy & Guidance**



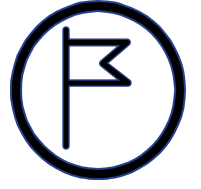
**Oversight**



**Reporting**



**Congressional Tasking**



**Strategic Initiatives**



January 2024  
This report satisfies the provisions of title 10, United States Code, section 129. The report summarizes the operational test and evaluation activities (including live fire testing activities) of the Department of Defense during the preceding fiscal year.

*Raymond D. O'Toole, Jr.*  
Dr. Raymond D. O'Toole, Jr.  
Acting Director



Test the way we fight

Architect T&E around validated mission threads & demonstrate the operational performance of the Joint Force in multi-domain operations



Accelerate the delivery of weapons that work

Embrace digital technologies to deliver high-quality systems at more dynamic rates



Improve survivability of DoD in a contested environment

Identify, assess, & act on cyber, electromagnetic spectrum, space, & other risks to DOD mission – at scale & speed



Pioneer T&E of weapon systems built to change over time

Implement fluid & iterative T&E across the entire system lifecycle to help assure continued combat credibility as the system evolves to meet warfighter needs



Foster an agile & enduring T&E enterprise workforce

Centralize & leverage efforts to assess, curate, & engage T&E talent to quicken the pace of innovation across the T&E enterprise

# What Do These Systems All Have In Common?



All lacked adequate operational testing & evaluation!

# Transformation of Today's Battlefield...



**SEAMLESS COLLABORATION ACROSS ALL DOMAINS**

**ATTRITABLE SYSTEMS AT SCALE**

**ENABLED BY AUTONOMY & AI**

# Transformation of Today's Battlefield...



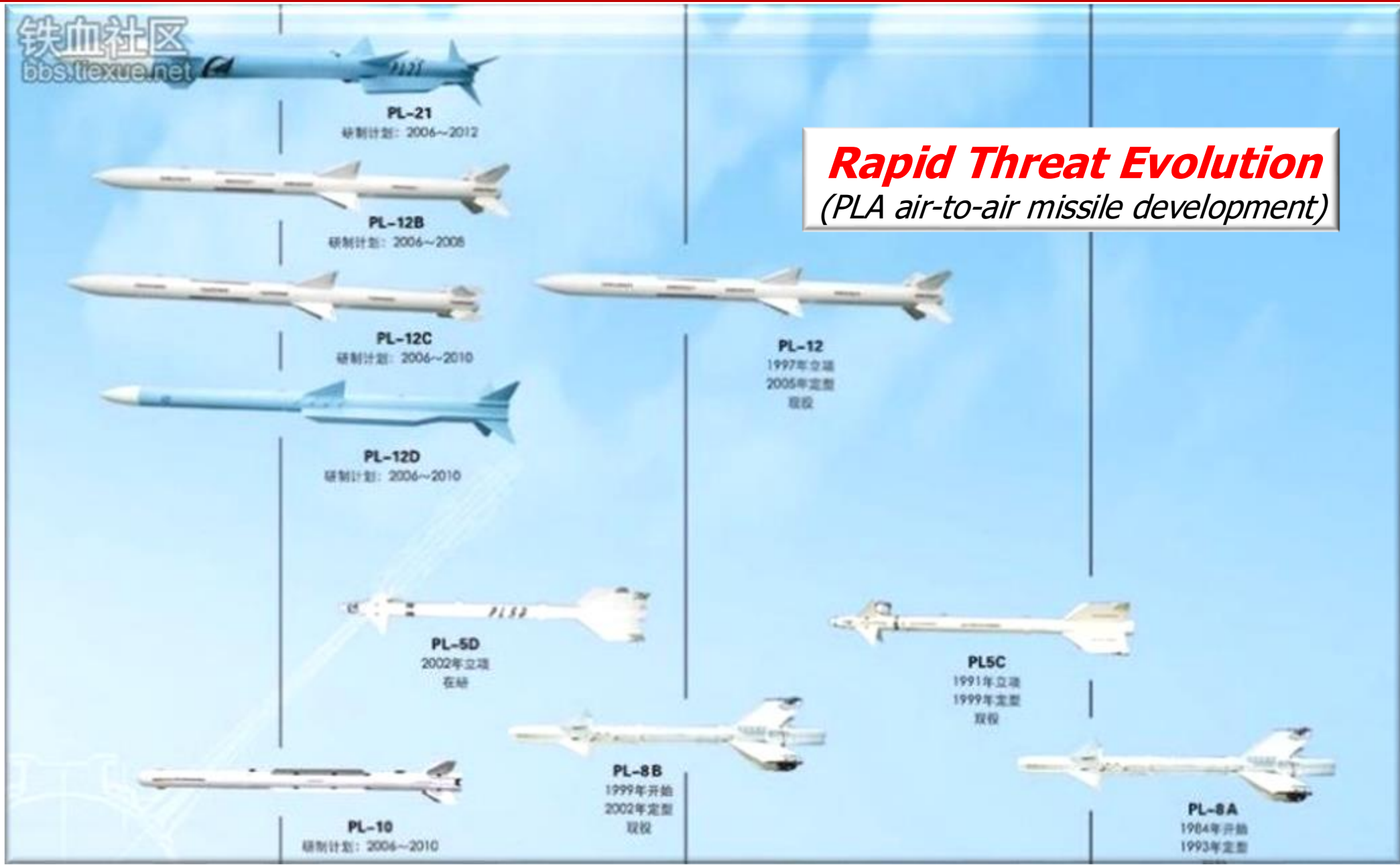
**LONG-RANGE LETHALITY**

**ATTRITABLE SYSTEMS AT SCALE**

**CONTESTED ELECTROMAGNETIC SPECTRUM**

**ANTI-ACCESS / AREA DENIAL (A2/AD)**

# Transformation of Today's Battlefield...





# DOT&E Strategic Pillars

## DOT&E Strategy Implementation Plan (I-Plan)

### 1. Test the way we fight



- Standardize the development of a scalable and adaptive representation of the multi-domain operating environment
- Implement measures, tools, and processes to efficiently evaluate kill webs and system-of-systems performance

### 2. Accelerate delivery of weapons that work



- Develop and implement an enterprise-level T&E data management solution
- Integrate T&E in model-based engineering to operationalize and optimize the Shift Left approach

### 3. Improve the survivability of DoD in a contested environment



- Standardize and automate mission-based risk assessments
- Emphasize cyber and electromagnetic spectrum survivability
- Evaluate operational performance in a contested space environment

### 4. Pioneer T&E of weapon systems built to change over time



- Increase the use of credible digital twins in T&E
- Evaluate the operational and ethical performance of AI-based systems
- Advance the evaluation of software-reliant systems' operational performance

### 5. Foster an agile and enduring T&E enterprise workforce



- Identify and track T&E workforce competencies and capabilities
- Assess and address critical T&E workforce professional development needs

# DOT&E Strategic Pillars



## **PILLAR 1**

*Test The Way We Fight*



## **PILLAR 3**

*Improve the  
Survivability of DoD in a  
Contested Environment*



## **PILLAR 4**

*T&E of Weapon Systems  
Built to Change Over Time*



## **PILLAR 2**

*Accelerate The Delivery  
of Weapons That Work*



## **PILLAR 5**

*Foster an Agile & Enduring  
T&E Enterprise Workforce*

# Test the Way We Fight



## PILLAR 1

*Test The Way We Fight*

*Provide tools, methods, & processes to evaluate complex missions in challenging all-domain operations*



*Open Air Battle Shaping (OABS)*



*Joint Simulation Environment (JSE)*



*Contested Operational Environments*



*Threat Radar Emulator*



*5<sup>th</sup> Gen Aerial Target (5GAT)*

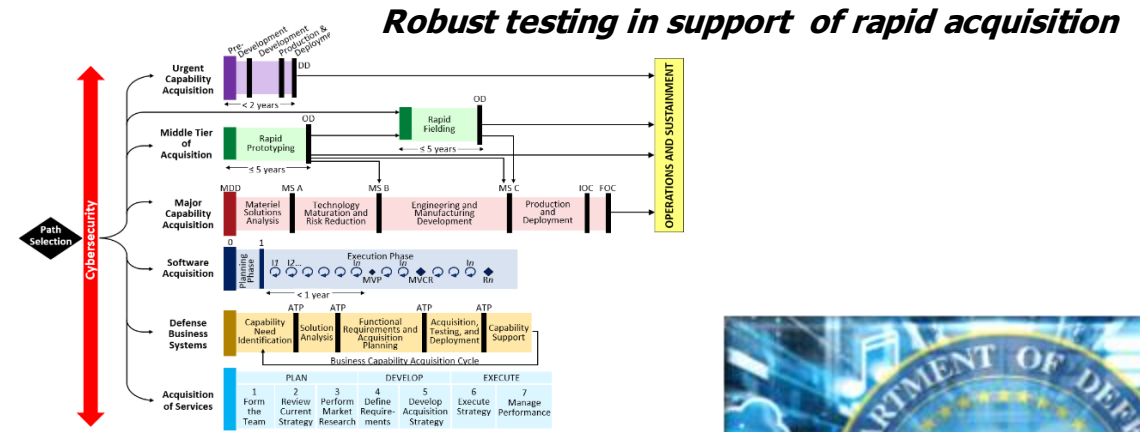
**Goal: Eliminate failure on first use in combat**

# Accelerate the Delivery of Weapons That Work

*Develop enterprise level T&E data management & analysis solutions that maximize the power of data*



**PILLAR 2**  
*Accelerate The Delivery of Weapons That Work*



**Government cloud data access**



**Accessible, validated, & secure data for all stakeholders**

**Goal: Increase confidence that warfighters have the best available capabilities**

# Improve the Survivability of DoD in a Contested Environment



*Satellite in an anechoic chamber*

*Enable a dynamic response to cyber & electromagnetic threats, & tests for full-spectrum survivability*



**PILLAR 3**  
*Improve the Survivability of DoD in a Contested Environment*



*Known & Expected Cyber Survivability Limits*



*Stryker during Army Mounted Assured PNT System (MAPS) operational testing*

**Goal: Provide freedom of maneuver throughout the modern battlespace**

# T&E of Weapon Systems Built to Change Over Time

*Automated evaluation of software-reliant system operational performance*



*Multi-Domain / Multi-System Operations*



**PILLAR 4**  
*T&E of Weapon Systems Built to Change Over Time*

*Increase the use of digital twins & responsible AI in T&E & tools to support effective software T&E*

**Goal: Counter rapid threat evolution to remain current & relevant**

# Foster an Agile & Enduring T&E Enterprise Workforce

*Automation & cyber-security skillsets*



*Augmented intelligence (AI+) & generative AI skillsets*



*Organize, train, & equip the next-generation workforce*

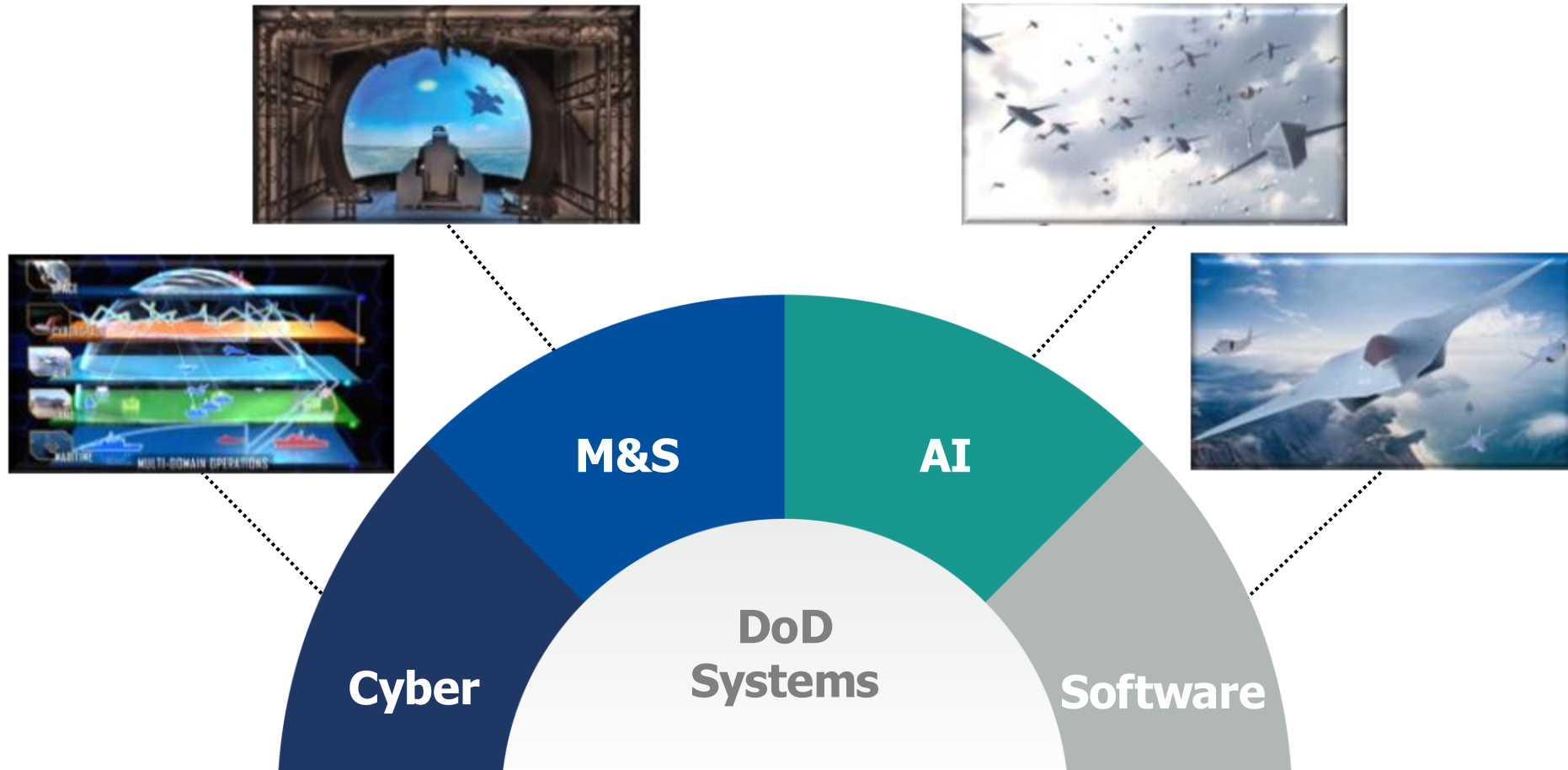


**PILLAR 5**  
Foster an Agile & Enduring  
T&E Enterprise Workforce

**Goal: Learn & embrace emerging technologies to leverage their benefits**

# Key Challenge Areas for Testing & Evaluation

Faster & more effective testing, dynamic adaptation to new scenarios & threats, & adequate testing in environments that are physically hard to replicate



The advantage in future conflicts will accrue to whichever side can fix & improve their software most rapidly & reliably



# DoD Polices for OT&E & LFT&E

## DoDI 5000.XF

Establishes policy, assigns responsibilities, & prescribes procedures for operational test & evaluation (OT&E) & live fire test & evaluation (LFT&E).

## DoDM 5000.UX

OT&E & LFT&E input to the test & evaluation master plan (TEMP), a test & evaluation (T&E) strategy, or an equivalent artifact.

## DoDM 5000.UW

Verification, validation, & accreditation (VV&A) of modeling & simulation (M&S) tools critical to meeting OT&E & LFT&E objectives.

## DoDM 5000.96

OT&E & LFT&E of DoD software-intensive systems & services, & software embedded in systems & services.

## DoDM 5000.UT

Realistic full spectrum survivability & full spectrum lethality testing of DoD systems & services.

## DoDM 5000.UZ

OT&E & LFT&E of artificial intelligence (AI)-enabled & autonomous systems & services.

CREATE

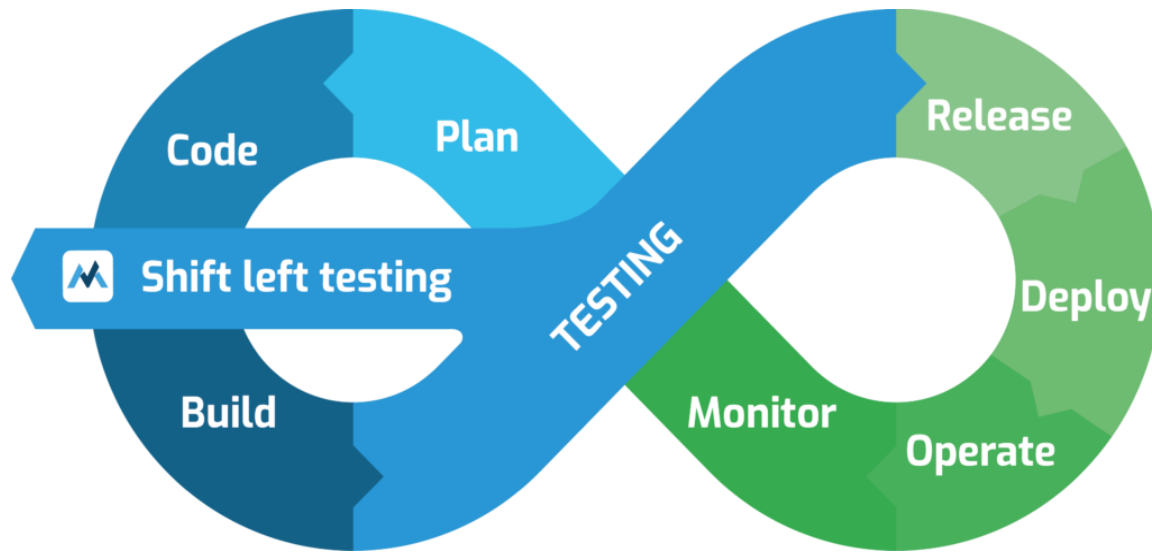
APPLY

AMPLIFY

These policies should be released soon

# Opportunities for Generative AI in T&E

- Augment various activities across the T&E enterprise



**Generative Augmented Intelligence (AI+) Presentations**  
Douglas Schmidt  
44 videos • 1,316 views • Last updated on Aug 18, 2024

Play all Shuffle

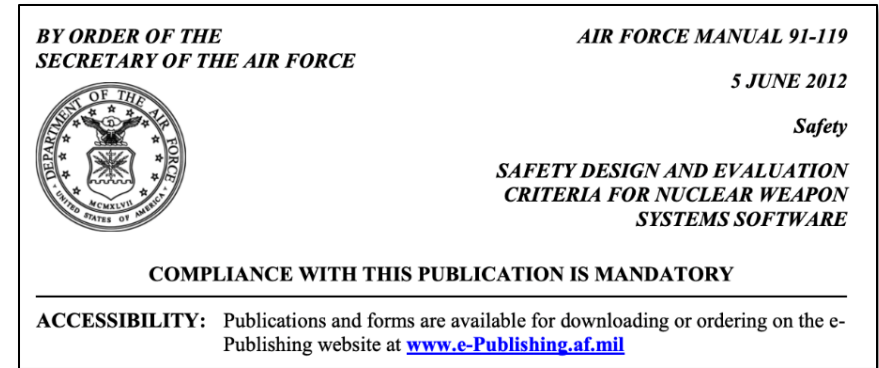
This playlist contains video presentations I've given either by myself or with my colleagues at Vanderbilt related to generative augmented intelligence and ChatGPT et al.

- 1 Technology Innovations and Their Ethical Implications  
Douglas Schmidt • 333 views • 5 months ago
- 2 Overview of Generative Augmented Intelligence (AI+)  
Douglas Schmidt • 192 views • 5 months ago
- 3 Applying Generative AI to Computer Science Courses at Vanderbilt  
Douglas Schmidt • 174 views • 5 months ago
- 4 Strategies for Using AI+ Effectively and Ethically  
Douglas Schmidt • 69 views • 5 months ago
- 5 Wrapping Up and Looking Ahead  
Douglas Schmidt • 72 views • 5 months ago
- 6 Ask Us Anything: Generative AI Edition  
Software Engineering Institute | Carnegie Mellon University • 1.8K views • Streamed 11 months ago
- 7 The Future of Software Engineering and Acquisition with Generative AI  
Software Engineering Institute | Carnegie Mellon University • 4.1K views • Streamed 7 months ago
- 8 Navigating Our AI-Augmented Future in National Security & Other High-Stakes Domains  
Vanderbilt University • 306 views • 6 months ago
- 9 ChatGPT is Just the Beginning: Generative AI will Transform Computing  
Douglas Schmidt • 1.5K views • 6 months ago
- 10 Generate a Retrofit API Client Automatically via ChatGPT-4  
Douglas Schmidt • 278 views • 6 months ago
- 11 Navigating Our AI-augmented Future (P3): Impact on Engineering AI-augmented Software Systems  
Douglas Schmidt • 125 views • 7 months ago
- 12 Navigating Our AI-augmented Future (P2): Impact on AI-augmented Software Development  
Douglas Schmidt • 145 views • 7 months ago

See [www.youtube.com/playlist?list=PLZ9NgFYEMxp72Zo0yrTNS6utAXxYpqNGI](https://www.youtube.com/playlist?list=PLZ9NgFYEMxp72Zo0yrTNS6utAXxYpqNGI)

# Opportunities for Generative AI in T&E

- Augment various activities across the T&E enterprise
- Derive test cases from relevant design, policy, & requirement documents
  - e.g., use LLMs to analyze documents written in natural language
  - Ensure tests align w/specifications, policies, etc. from the outset



Objectives	Using LLMs
Ensure document is clear & complete to ensure nuclear surety for key software components	Check for discrepancies <ul style="list-style-type: none"><li>• within 91-119</li><li>• between it &amp; other documents</li></ul>

- **Ambiguity in Safety Certification Components:** On pages 32-33, 91-119 discusses safety certification for software components, suggesting to list them separately or combined with other safety-certified components. However, combining safety-certified components with non-safety-certified ones could complicate change tracking. Clearer guidelines are needed to avoid inconsistencies in how components are combined and tracked.

See [insights.sei.cmu.edu/blog/applying-large-language-models-to-dod-software-acquisition-an-initial-experiment](https://insights.sei.cmu.edu/blog/applying-large-language-models-to-dod-software-acquisition-an-initial-experiment)

# Opportunities for Generative AI in T&E

- Augment various activities across the T&E enterprise
  - Derive test cases from relevant design, policy, & requirement documents
- Use large language models (LLMs) to simulate diverse usage patterns & environments to test systems under various conditions
  - e.g., apply the *Persona* pattern

Act as a senior security engineer. You will help me investigate potential threats to my organization.

We will work together to investigate threats. I can run tools and software to gather information for us. I can cut/paste the outputs here for you to analyze.

You can ask me to do the following things:

1. Run a Linux command-line tool that I have access to and provide the output from the tool.
2. Run a Python program that you create to collect information and print it out to the terminal so that I can cut / paste it her for you to look at.
3. Write a Python script that I can run to query the NIST CVE database for known vulnerabilities related to the host, OS, services, etc. on a device and cut/paste the results for you to look at.

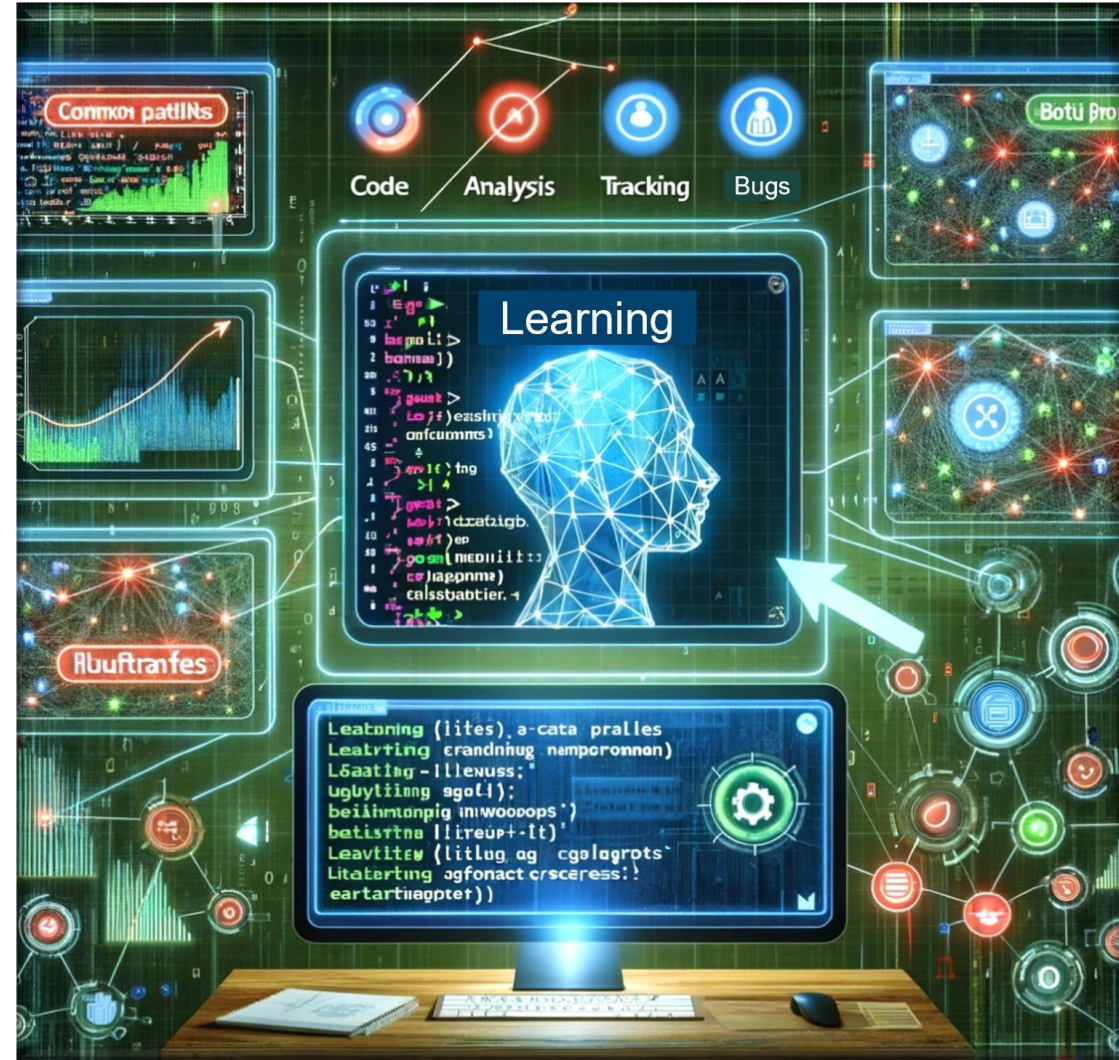
You will keep asking me to perform operations until you have enough information to recommend a plan of action. After each task you ask me to perform, remind me of what we are doing in a paragraph and then ask me for the input from the last task that you asked me to perform.

Ask me for the threat to investigate.

See [www.dre.vanderbilt.edu/~schmidt/PDF/PLoP-patterns.pdf](http://www.dre.vanderbilt.edu/~schmidt/PDF/PLoP-patterns.pdf)

# Opportunities for Generative AI in T&E

- Augment various activities across the T&E enterprise
  - Derive test cases from relevant design, policy, & requirement documents
  - Use large language models (LLMs) to simulate diverse usage patterns & environments to test systems under various conditions
- Help testers & T&E organizations learn from prior efforts
  - Continuously improve the T&E process over time by analyzing test data to identify common pitfalls & best practices



See [www.linkedin.com/pulse/5-ways-ai-disrupting-traditional-software-testing-process-sheldon](https://www.linkedin.com/pulse/5-ways-ai-disrupting-traditional-software-testing-process-sheldon)

# Opportunities for Generative AI in T&E

- Augment various activities across the T&E enterprise
  - Derive test cases from relevant design, policy, & requirement documents
  - Use large language models (LLMs) to simulate diverse usage patterns & environments to test systems under various conditions
  - Help testers & T&E organizations learn from prior efforts
- Identify trends in DoD acquisition processes over time
  - e.g., analyze DOT&E annual reports for the past three decades



See [www.dote.osd.mil/annualreport](http://www.dote.osd.mil/annualreport)

# How to Help Deliver Weapons that Work... Faster



# How to Help Deliver Weapons that Work... Faster



**Enhanced Cyber Resilience**



**Operationally Representative Data**



**Digital-Physical Fusion**



**AI-/ML-enabled Systems & Workforce**



# How to Help Deliver Weapons that Work... Faster



Enhanced Cyber Resilience



Operationally Representative Data

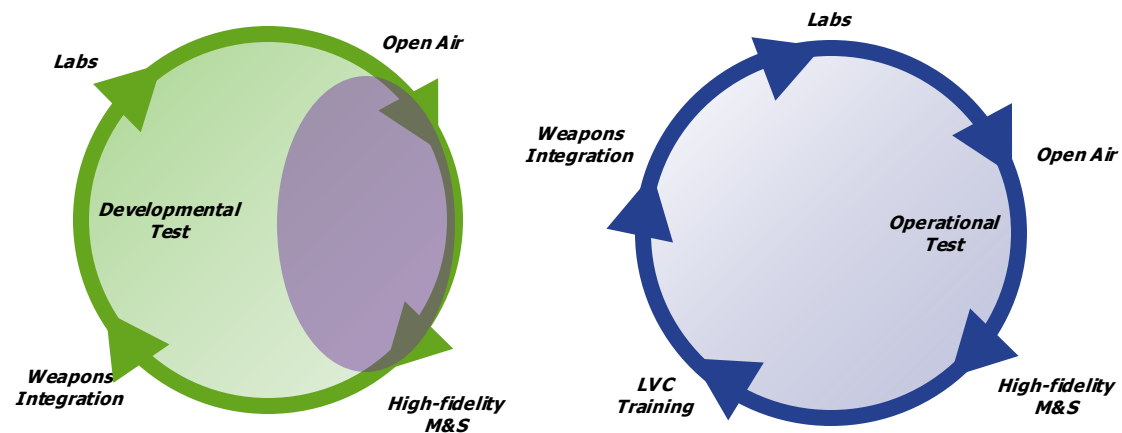


Digital-Physical Fusion



AI-/ML-enabled Systems & Workforce

## Integrated Test



← Shift left, look right →



# Questions & Answers

For more information please contact:

**HON Dr. Douglas C. Schmidt**

Director, Operational Test & Evaluation

Office of the Secretary of Defense

Tel. (703) 697-3655

See [www.dote.osd.mil](http://www.dote.osd.mil)