REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01/01/2023	Conference Briefing	

4. TITLE AND SUBTITLE

DoD Enterprise T&E of AI-Enabled Systems Future State and Strategic Goals

5a. CONTRACT NUMBER

W56KGU-18-D-0004/W56KGU20F0009

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The MITRE Corporation
7515 Colshire Drive
McLean, VA 22102

8. PERFORMING ORGANIZATION REPORT NUMBER

PRS-22-4193

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of the Director for Developmental Test, Evaluation, and Assessments, Office of the Undersecretary of Defense

10. SPONSOR/MONITOR'S ACRONYM(S)

OUSD DTE&A

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Information Session Purpose: § Continue communication about the DoD T&E of AI* Future State and Enterprise Goals § Begin convening a DoD T&E of AI Stakeholder Community § Prepare DoD T&E of AI stakeholders for survey. The DoD Enterprise T&E of AI goals were developed by a group of representatives from across DoD: DTE&A, DOT&E, CDAO, TRMC, USD(R&E), OTAs (AFOTEC), MITRE, IDA, CMU SEI. The goals were published in the July CAPE Study Report.

15. SUBJECT TERMS

Artificial Intelligence; Systems Development; test; evaluation; Artificial Intelligence; AI

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Susan Carpenito
				23	19b. TELEPHONE NUMBER *(Include area code)* 781-271-7646

MITRE | SOLVING PROBLEMS FOR A SAFER WORLD®

DoD Enterprise T&E of AI-Enabled Systems Future State and Strategic Goals

January 2023

Approved for Public Release; Distribution Unlimited.

Public Release Case Number 22-4193

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

DoD Enterprise T&E of AI-Enabled Systems Future State & Goals Information Session Purpose and Topics

Information Session Purpose:

- **Continue communication about the DoD T&E of AI* Future State and Enterprise Goals**
- **Begin convening a DoD T&E of AI Stakeholder Community**
- **Prepare DoD T&E of AI stakeholders for survey**

Co-Sponsors: DTE&A, DOT&E, and CDAO

Invitees:

- **DoD personnel and contractor/FFRDC support who have a stake in T&E of AI-enabled systems or would benefit from the outcomes of these goals**

Topics:

- **Background: Brief review of DoD T&E of AI Future State**
- **Review DoD Enterprise T&E of AI Goals, with focus on the Near-Term Goals**

* Note: "AI" here means AI-enabled systems and does not include autonomous systems.

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

DoD T&E of AI-Enabled Systems Future State

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

Future State – Current Challenge

We are projecting a future state for the test and evaluation (T&E) of Artificial Intelligence (AI) to better handle the tidal wave of AI-enabled capabilities.

- Statistical measures are useful, but don't address how the AI-enabled system enables the mission
- Data to train AI models are in short supply, critical gap that limits robust evaluation
- Infrastructure to build and test AI-enabled systems not readily available when/where needed
- Cybersecurity of AI-enabled systems can only be achieved by processes that extend the whole AI lifecycle
- DoD does not yet put urgency on T&E of AI policy development, tide is turning
- User Engagement is critical for AI-enabled systems to ensure trustworthiness in the deployed system

Developmental Test, Evaluation, and Assessments (DTE&A) is working to help the Department of Defense (DoD) community get ahead of this challenge through investigation and collaboration.

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

Current State of T&E of AI-Enabled Systems in DoD

- **SMEs in AI are** joining the workforce to build AI but unclear path forward for ensuring the T&E community is well qualified to enable T&E.
- The **method to train and evaluate stand alone AI models** is maturing, but the T&E **of AI-enabled software** is less defined. Unclear how to account for the variable and unique behavior of AI in their deployed environment.
- **Infrastructure** that can provide a safe and secure shared development, training, and test environment is clearly needed. Infrastructure to serve the DoD services is emerging (e.g., CDAO/JAIC's Joint Common Foundation became operational in 2021). More work is needed to provide the necessary HW/compute, tools, to DoD's AI developers *where they work*.
- AI enabled systems must prove that it can meet the **needs of users** and that users will **trust** and adopt AI for use in their missions. Traditional AI developers work with limited user access or engagements. This limited access often produces AI-enabled systems that are not readily adopted despite its apparent value. Standard methods to provide access to users to shape AI are not yet consistently applied nor well defined.
- **Data** to support the training, development, and testing of AI is being collected and managed at the local or project level today. Today, data is often not managed and retained in a way that enables it to be used for a local project or for broader T&E benefit. Protecting the data is not yet understood as a threat vector that needs to be guarded.
- **Measures** currently exist to assess AI Model's performance in a constrained setting. Approaches that incrementally assess AI-enabled system's performance within representative operational conditions driven by the mission yield the most promise. While there is much information about the qualitative aspects of AI performance, broader measures that inform the level of trustworthiness that one should place when it is developed with certain infrastructure, data, SME, user involvement, and model training do not yet exist.

DTE&A TEM Participants:

OUSD R&E
OUSD DTE&A
TRMC
JAIC T&E, JCF
DOT&E
ATEC
AFOTEC
JITC
NCRC
NAVY
USAF
STAT COE
IDA
MITRE
MIT LL
Oak Ridge National Lab
DARPA
DISA
Army Futures Command
Army PEO C3T
Virginia Tech
Carnegie Mellon Univ.
ARL, CCDCUSAF
AFMC
DUSA-TE

In FY21 MITRE conducted a study to understand enabling efforts, methods, and resources to build and assess AI capabilities. TEMs (FY21/22) were conducted to share findings and gather information on needs and gaps.

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

Focus Areas to Achieve the Future State of T&E of AI

Artificial Intelligence (AI) is a critical technology that the DoD must use to meet certain mission needs to keep pace with its near-peer adversaries. The Undersecretary for Research and Engineering cited "trusted AI" as among the top research priorities.* DOD must define holistic T&E methods for AI models and AI enabled systems that assure trust. Coordinated efforts are required to achieve this future state.

"AI, Networks, Hypersonics Are the Pentagon's Top Research Priorities," Tirpak, John A., Air Force Magazine, Jan 2022

Policy
- Standardize and modernize, informed by DOD guidance and priorities.
- New paradigms for AI-enabled systems acquisition and T&E
- Further guidance and policy to drive practitioners

Users
- Human AI Teaming practices are in place throughout the lifecycle.
- Trained users providing feedback on systems through continuous feedback loops
- User adoption increased though informed trust

Cybersecurity
- Legal requirement to self declare when AI is present
- Decisionmakers are continually informed of risks and mitigations
- AI as part of acceptance (e.g., RMF)
- Tools and methods to conduct threat informed T&E

Measures
- Standard set of clear repeatable metrics mapped to AI model, user and mission needs
- Integrated metrics to facilitate trust and inform on mission relevance
- AI as a system in a system

Data
- Data set planned, generated (e.g., synthetic) and curated (labeling)
- T&E Practitioners have data sets mapped to operational conditions
- Test data and operational data captured to drive feedback loops

Infrastructure
- Provides end to end AI lifecycle from training to deployment in and independently verifiable system
- Integration to DE products and processes
- Curates, partitions and protects data
- Integrated with: M&S, test stim, instrumentation

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

DoD T&E of AI-Enabled Systems Strategic Goals

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

DoD Enterprise T&E of AI Goals Formation Approach

The DoD Enterprise T&E of AI goals were developed in a hands-on, interactive workshop by a group of representatives from across DoD: DTE&A, DOT&E, CDAO, TRMC, USD(R&E), OTAs (AFOTEC), MITRE, IDA, CMU SEI. The goals were published in the July CAPE Study Report.

Inputs:

- DoD AI T&E Current State Report (2021)
- DoD AI T&E TEMs
- DoD AI T&E Future State (13 Apr 2022)
- Demand & Supply Survey Results:
  - T&E requirements and demand forecast, based on surveying DoD AI programs
  - T&E Supply Baseline: Catalog of existing and planned DoD AI T&E efforts

Review DoD AI T&E Current State, Recent TEMs Takeaways

Refresh DoD AI T&E Future State
- Note goals during presentations

Ideate a "First Pass" Set of Goals

Understand Demand & Supply Survey Results (Environmental Scan)
- Note gaps & enablers during presentations

Identify Gaps
- Collect & organize gaps

Summarize Major Strategic Gaps

Crystallize Top-Level Goals: Near-, Medium-, and Long-Term

Map Goals to High-Level Timeline

Review & Adjust Goals Timeline

Present Goals Timeline to SESs

Outcomes:

DoD AI T&E:
- **Future State (refreshed)**
- **Gaps**
- **Enterprise Goals**
- **High-level Timeline of Enterprise Goals**

All published in the CAPE Study Report

Critical Success Factors: Participants took a DoD enterprise perspective and were empowered to make decisions.

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

Long-Term (LT) DoD Enterprise T&E of AI Goals

1. LT1 - Workforce: Develop a DoD T&E workforce that has knowledge of AI as a capability and is skilled in the tools, methods, and best practices for testing and evaluating AI models and AI-enabled systems.
2. LT2 - Datasets: Provide sufficient, operationally relevant, usable datasets that DoD AI developers and T&E personnel can find and access to independently develop and test AI models and AI-enabled systems.
3. LT3 - Tools: Provide interoperable tools and reusable measures in repositories accessible across DoD to develop, train, test, evaluate, manage, and sustain AI models and AI-enabled systems across the AI lifecycle.
4. LT4 - User/Operator Involvement: Involve users and operators early and often throughout the AI lifecycle to ensure effective generation of AI use cases, requirements, T&E plans, AI-enabled work processes with human-machine teaming, and AI trustworthiness assurance methods, all resulting in suitable, verified, and validated AI capabilities and outcomes.
5. LT5 - Interoperability: Establish a DoD T&E of AIES ecosystem in place of the current stove-piped paradigm that provides integrated and interoperable infrastructure, processes, and practices for AI T&E and development across DoD, including safety, cybersecurity, test range capabilities, measures, and AI trustworthiness assurance.

The enterprise T&E of AI long-term goals provide a usable end-state for the DoD Enterprise T&E of AI Roadmap

MITRE

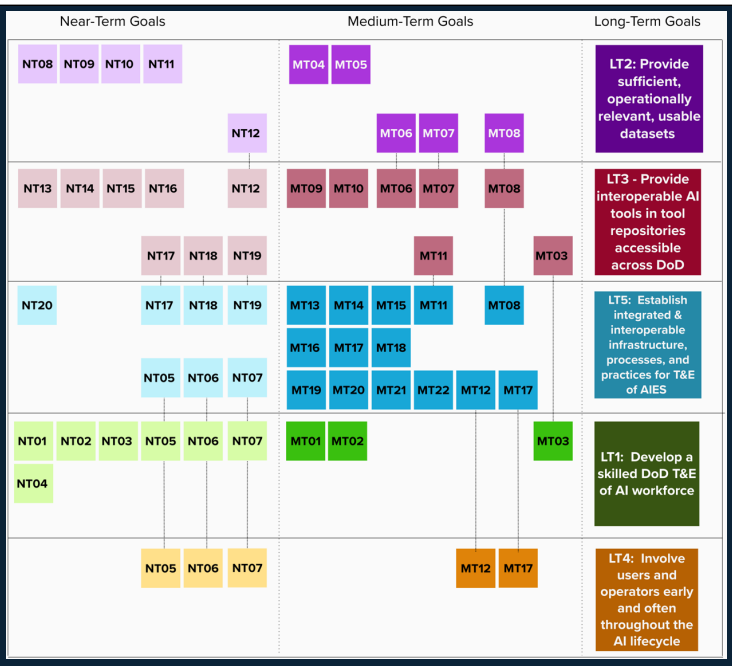
© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

This set of LT goals is the result from the cross-DoD Enterprise T&E of AI Goals Formation Workshop and is included in the CAPE Study Report.

DoD Enterprise T&E of AI Goals Big Picture

- 5 long-term goals, 22 medium-term goals, and 20 near-term goals.
- Each long-term (5+ year) goal has contributing near-term (1 - 1.5 year) and medium-term (2 – 5 year) goals.
- Some near-term and medium-term goals contribute to more than one long-term goal (shown with grey lines across the LT goal swimlanes)

Roadmapping process starts with defining activities to achieve the near-term goals



MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

10

This NT, MT, and LT goals set and alignment is the result from the cross-DoD Enterprise T&E of AI Goals Formation Workshop and is included in the CAPE Study Report.

DoD Enterprise T&E of AI Near-Term Goals

Near-Term Goal Short Phrase	Near-Term Goal Description
NT01: DoD Ecosystem Roles & Responsibilities for T&E of AIES	Define organizational roles, responsibilities, dependencies, & authorities across DoD for T&E of AIES
NT02: DoD T&E Community AI Literacy	Build the DoD T&E community's literacy about AI and T&E of AIES
NT03: Communication Methods for Threats & Vulnerabilities of AIES	Create communication methods to provide information on threats to and vulnerabilities of AIES to requirements development and T&E organizations
NT04: Cyber Test Teams Capacity Building for AIES	Identify missing skillsets, tools, and capabilities of cyber testing teams for AIES
NT05: Guidance on Involving T&E in Requirements, Contracts, & RFPs/RFIs for AIES	Establish and communicate guidance & practices to involve the T&E of AI community in defining specifications for AIES in requirements documents, RFPs/RFIs, contracts, and similar documents
NT06: Guidance on User Involvement in T&E of AIES	Develop & communicate guidance to facilitate end-user, operator, and unit involvement in T&E of AIES
NT07: Guidance on CONOPS, Process Engineering, and HMT for AIES	Create and disseminate guidance for AI development teams to document/refine a CONOPS and process designs with human-machine teaming (HMT)
NT08: Stakeholder Agreement on Data Discoverability	Obtain DoD stakeholder agreement on data being discoverable and corresponding controls
NT09: Data Access & Availability	Increase availability of and access to high-quality training, validation, and test data, including synthetic data
NT10: Data Architecture for Tailorable Tech Stacks & Pipelines	Build data architecture to support tailorable tech stacks for AI pipelines
NT11: Practices for Data Management, Characterization, Taxonomies	Develop frameworks/practices for data management, characterization, and taxonomies for AIES
NT12: Site Connectivity	Identify needs/gaps and expand connectivity between sites and platforms/systems for data sharing, tools, etc. for T&E of AIES
NT13: Tools Visibility & Access	Provide increased visibility and access to T&E tools for AI developers and T&E personnel
NT14: M&S Into T&E Practices	Integrate Modeling and Simulation (M&S) for T&E into DoD T&E of AIES practices early in the AI lifecycle
NT15: Cyber T&E Tools Requirements/Gaps and Pilots	Identify gaps and requirements for cyber T&E tools, capabilities, and repeatable processes for AIES
NT16: Tools Research, Prototyping, & Build/Acquisition	Research, prototype, pilot, and build and/or acquire T&E of AIES tools and capabilities that are cloud-agnostic and are usable across varying environments
NT17: Exemplar Reusable Measures	Define exemplar reusable measures for T&E of AI models and AIES
NT18: Sharing Practices, Tools, Measures Guidance, Etc.	Iteratively create, collect, share, and refine T&E best practices, tools, and measures specific to AI models and applications
NT19: Data, Model, & Process Characterization	Clarify data, model, and process characterization and reporting across program types and lifecycles
NT20: Policies & Rqmts for T&E-Developer Interoperability	Develop policies advancing interoperability between DoD T&E of AIES personnel and model developers

AIES: AI-enabled systems

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

Next Steps

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

Way Forward to the Future State

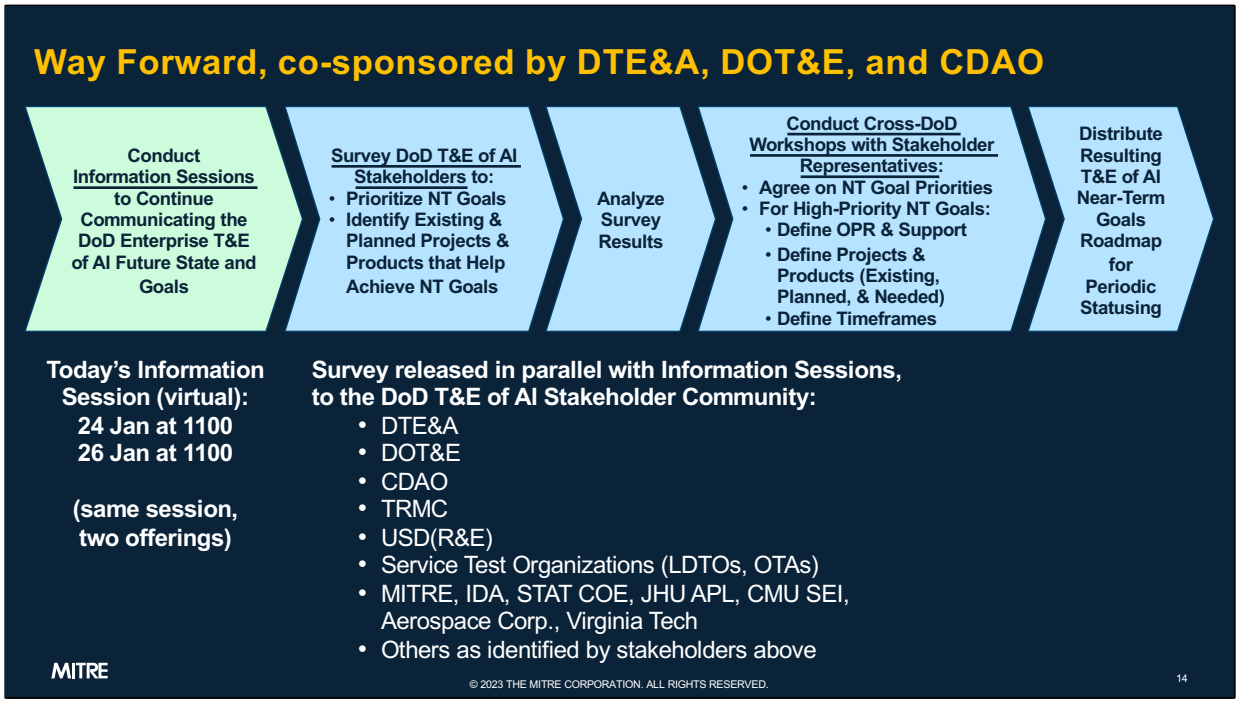
Work across the broader DoD, commercial, FFRDC, and academic communities to launch specific initiatives that enable these efforts. The T&E community will need to navigate competing ideas to quickly move towards the future state, while staying in step with planned and future AI capabilities.

Engage stakeholders across DoD who are involved in T&E of AI-enabled systems and leverage goals to inform a roadmap

- **Start with the near-term goals:**
  - Prioritize the near-term goals
  - Identify current and future planned projects, activities, products that help achieve the near-term goals
- **Develop a Near-Term Goals Roadmap that "connects the dots" across existing, planned, and new projects, activities, and products to show line-of-sight activities and OPRs to achieve the high-priority near-term goals**
- **Review and refresh the Near-Term Goals Roadmap at least quarterly with status, new priorities, and updates**
- **Review progress in 1.5 years and develop Medium-Term Goals Roadmap**

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.



Brown-Bag:
Invite people who have a stake/interest in either executing on the T&E of AI strategic goals or would be a beneficiary/recipient of outcomes from the goals.

Why a DoD Enterprise T&E of AI Near-Term Goals Roadmap?

Roadmap Purpose: The Roadmap **"connects the dots"** and **provides DoD line-of-sight** on how the Department is achieving the **DoD enterprise T&E of AI near-term goals** which, in turn, **advances DoD's achievement of the future state** for T&E of AI-enabled systems.

Roadmap Benefits:

- Builds on the comprehensive set of long-, medium-, and near-term goals developed by a cross-DoD group to drive to the T&E of AI future state
- Provides line-of-sight from near-term goals to longer-term goals and the DoD T&E of AI Future State
- Identifies existing and planned projects, activities, and products, gaps to be filled, organizational OPRs, and timelines to achieve the near-term goals
- Allows for prioritization of near-term goals

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

Survey to Get Input from the DoD T&E of AI Community

Survey Questions:

- How would you **prioritize** these T&E of AI near-term goals?
- What **current and planned projects, activities, and products** help achieve the near-term goals? What organizations are the **OPRs** for these efforts?
- Would you like to be included in a **"DoD T&E of AI Community Directory"** that can be used to begin building a DoD-wide stakeholder community?

Survey Logistics: Please provide your input *by 1700 on Friday 3 Feb*

→ Download survey Excel file attachment (in Info Session invitations). Complete survey (instructions within).
Rename your survey response Excel file as: ORGANIZATION-FIRSTNAME-LASTNAME.xlsx

→ *People with DoD CAC*: Upload your renamed survey response Excel file to Intelink folder here: link

- Note: Blank survey Excel file can also be downloaded from this Intelink folder.

→ *People without DoD CAC (or unable to access Intelink)*:

- Exchange digital certs through signed emails to/from Jacqui Lee (jacquelinelee@mitre.org).
- Send encrypted email, with your survey response Excel file attached, to Jacqui Lee.

Please help get the survey to people in your organization who are involved in T&E of AI.

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

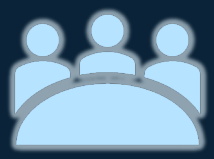
16

Seeking input from anyone in DoD involved in or with a stake in T&E of AI-enabled systems. More input enriches the identification of current and planned projects/products/activities that contribute to the goals.

Roadmapping Workshops with DoD T&E of AI Representatives

Roadmapping Workshops Purpose:

- Gain consensus on priorities for the near-term goals
- For each high-priority near-term goal, define:
  - A DoD "owner" to steward this goal
  - Existing and planned projects, products, and activities that contribute to this goal, with timeframes (by quarter)
  - Gaps and additional projects/products/activities needed
  - OPR and support organization(s) for each of the projects/products/activities
- Bring the above information together into a roadmap for the high-priority near-term goals



Timing: Feb – Mar 2023

Participants:

- TBD representatives from DoD and partner organizations who are working on efforts related to the high-priority goals (related efforts and OPRs to be identified in the survey)
- Facilitators (MITRE)
- Note: May need separate workshops by goal or topic area, depending on survey results

MITRE

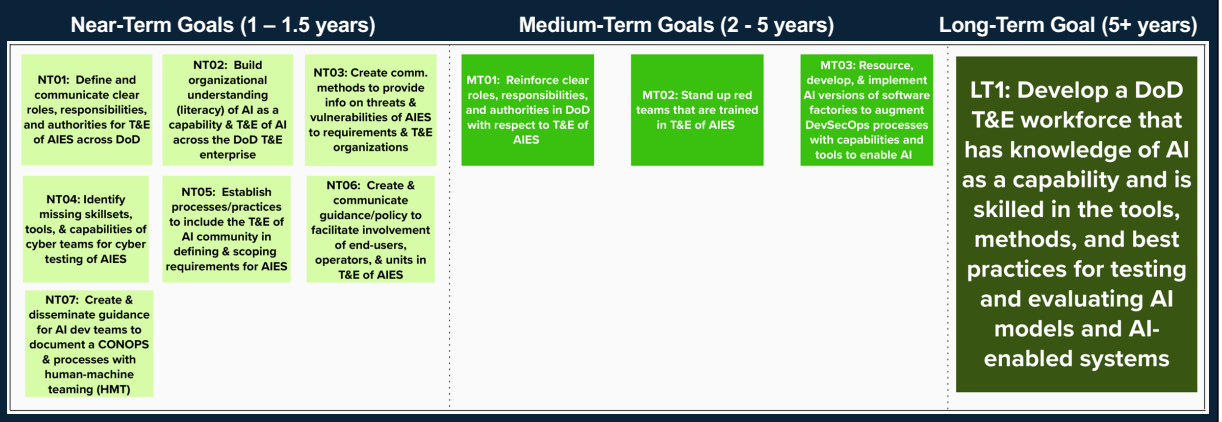
© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

BACK-UP:
DoD Enterprise T&E of AI Goals Swimlane Diagrams

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

Goals Swimlane for Long-Term Goal 1 (LT1): Workforce



TOPICS

- DoD organizational roles, responsibilities, authorities
- T&E community literacy
- methods to inform about threats & vulnerabilities
- cyber team skillsets, tools, capabilities in AIES
- practices to include T&E in requirements, contracts, RFPs/RFIs
- guidance to involve users
- guidance to define CONOPS, processes, HMT

- reinforcement of roles & responsibilities across DoD
- red teams trained in T&E of AIES
- AIES versions of software factories

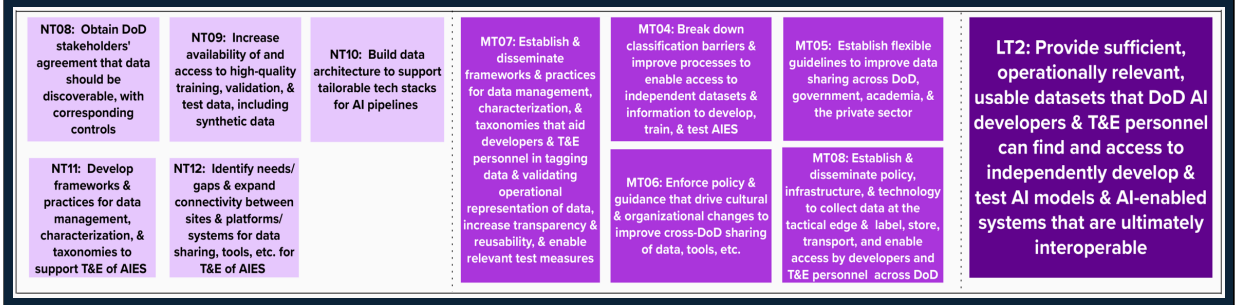
AIES: AI-enabled systems

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

19

Goals Swimlane for Long-Term Goal 2 (LT2): Datasets



TOPICS

- data discoverability
- data availability & access
- data management & characterization
- site connectivity needs/gaps for data sharing
- data architecture
- classification barriers
- data sharing across DoD, govt, academia, & private sector
- cross-DoD data sharing
- data management & characterization
- collecting data at the tactical edge

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

AIES: AI-enabled systems

Goals Swimlane for Long-Term Goal 3 (LT3): Tools

Near-Term Goals (1 – 1.5 years)			Medium-Term Goals (2 - 5 years)			Long-Term Goal (5+ years)
NT12: Identify needs/ gaps & expand connectivity between sites & platforms/ systems for data sharing, tools, etc. for T&E of AIES	NT13: Provide increased visibility and access to T&E tools for AI developers and T&E personnel, define tool reqmts, & establish standards for what the tools are used for	NT14: Integrate modeling and simulation (M&S) for T&E into DoD T&E of AI practices early in the AI lifecycle & ensure LVCA environments & M&S strategies account for program needs & technical limitations	MT03: Resource, develop, & implement AI versions of software factories to augment DevSecOps processes with capabilities and tools to enable AI	MT06: Enforce policy & guidance that drive cultural & organizational changes to improve cross-DoD sharing of data, tools, etc.	MT07: Establish & disseminate frameworks & practices for data management, characterization, & taxonomies that aid developers & T&E personnel in tagging data & validating operational representation of data, increase transparency & reusability, & enable relevant test measures	LT3: Provide interoperable tools & reusable measures in repositories across DoD to develop, train, test, evaluate, manage, & sustain AI models and AI-enabled systems across the AI lifecycle
NT15: Identify gaps, define requirements, and begin development of cybersecurity T&E tools, capabiliites, & repeatable process for AIES, informed by threats, across the whole lifecycle	NT16: Build & acquire cloud-agnostic T&E of AI tools & capabilities that can be used across a variety of environments, including the data and AI development environments	NT17: Define exemplar reusable measures for T&E of AI models & AIES that can be tailored or augmented for specific operational T&E requirements & user trustworthiness throughout the full AI lifecycle.	MT08: Establish & disseminate policy, infrastructure, & technology to collect data at the tactical edge & label, store, transport, and enable access by developers and T&E personnel across DoD	MT09: Expand & equip test range infrastructure to accommodate T&E of AIES, including T&E range safety & assurance use case development for more realistic testing of AI on DoD ranges		
NT18: Iteratively collect, create, share, & refine T&E best practices, tools, measures, guidance, policies, & other content specific to AI models & applications	NT19: Clarify data, model, and process characterization and reporting across program types and lifecycles.		MT10: Provide the right tools to enable use of improved cybersecurity practices for T&E of AIES	MT11: Build, establish, & disseminate interoperable T&E of AIES tools & AI model repository(ies) that shift T&E upstream, are platform-agnostic, can be used to evaluate AIES post-fielding, etc.		

TOPICS

- site connectivity needs/gaps for tools
- increased visibility and access to tools
- M&S into T&E of AI
- cybersecurity T&E tool/process gaps & requirements
- cloud-agnostic T&E of AI tools
- exemplar reusable measures
- iterative collection & sharing of tools, guidance, etc.
- data, model, and process characterization

- AIES versions of software factories
- Policy/guidance to improve cross-DoD sharing
- data management, characterization, & taxonomies
- collecting data at the tactical edge
- test range infrastructure
- tools to improve cybersecurity practices for T&E
- interoperable tools & model repositories

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

AIES: AI-enabled systems

Goals Swimlane for Long-Term Goal 4 (LT4): User Involvement

Near-Term Goals (1 – 1.5 years)			Medium-Term Goals (2 - 5 years)		Long-Term Goal (5+ years)
NT05: Establish processes/practices to include the T&E of AI community in defining & scoping requirements for AIES	NT06: Create & communicate guidance/policy to facilitate involvement of end-users, operators, & units in T&E of AIES	NT07: Create & disseminate guidance for AI development teams to document a CONOPS & processes with human-machine teaming (HMT)	MT12: Create a more rigorous, repeatable T&E of AIES assurance framework & assurance cases that: explicitly address user concerns, involve users, & use multiple sources of test evidence collection	MT17: Inform PMOs about relevant threats to AI models and AIES to drive choices & risks assessments	LT4: Involve users & operators early & often throughout the AI lifecycle to ensure effective generation of AI use cases, requirements, T&E plans, AI-enabled work processes with human-machine teaming

TOPICS

- practices to include T&E in requirements, contracts, RFPs/RFIs
- guidance to involve users
- repeatable assurance framework & cases
- informing PMOs about threats

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

AIES: AI-enabled systems

Goals Swimlane for Long-Term Goal 5 (LT5): Interoperability

Near-Term Goals (1 – 1.5 years)			Medium-Term Goals (2 - 5 years)				Long-Term Goal (5+ years)
NT05: Establish processes/ practices to include the T&E of AI community in defining & scoping requirements for AIES	NT06: Create & communicate guidance/policy to facilitate involvement of end-users, operators, & units in T&E of AIES	NT07: Create & disseminate guidance for AI dev teams to document a CONOPS & processes with human-machine teaming (HMT)	MT08: Establish & disseminate policy, infrastructure, & technology to collect data at the tactical edge & label, store, transport, and enable access by developers and T&E personnel across DoD	MT11: Build, establish, & disseminate interoperable T&E of AIES tools & AI model repository(ies) that shift T&E upstream, are platform-agnostic, can be used to evaluate AIES post-fielding, etc.	MT12: Create a more rigorous, repeatable T&E of AIES assurance framework & assurance cases that: explicitly address user concerns, involve users, & use multiple sources of test evidence collection	MT13: Develop, establish, & disseminate repeatable processes & practices for measuring human-machine teaming & user interaction with AIES	LT5: Establish a T&E of AIES ecosystem across DoD that provides integrated and interoperable infrastructure, processes, and practices for development and T&E of AIES to streamline evaluations of effectiveness, safety, suitability, cybersecurity, & trustworthiness & improve test range capabilities
NT17: Define exemplar reusable measures for T&E of AI models & AIES that can be tailored or augmented for specific operational T&E requirements & user trustworthiness throughout the full AI lifecycle	NT18: Iteratively collect, create, share, & refine T&E best practices, tools, measures, guidance, policies, & other content specific to AI models & applications	NT19: Clarify data, model, and process characterization and reporting across program types and lifecycles	MT14: Create & provide end-to-end infrastructure with adequate storage & bandwidth that supports developers & T&E personnel across the entire AI lifecycle	MT15: Define & regularly update standards for interoperability of data, models, & tests	MT16: Create a "playbook" of reusable exemplar measures for T&E of AI models & AIES, based on type of model & application, that can be tailored for specific operational requirements	MT17: Inform PMOs about relevant threats to AI models and AIES to drive choices & risks assessments	
NT20: Develop policies on ensuring interoperability between DoD AI T&E personnel and model developers to gain deeper insights on model behavior, including predictability and interpretability			MT18: Ensure safety policies provide flexibility & rigor for AI development & T&E without generating unnecessary safety incidents	MT19: Define & disseminate exemplar requirements for cybersecurity in AIES for development & T&E of AI, & integrate these rqmts into the Risk Mgmt Framework & DoD software acquisition	MT20: Craft effective T&E policy that accounts for whether a given system uses "artificial intelligence"	MT21: Develop threats information & characterize cybersecurity threats distinctive to AIES types (including ML & fixed-code approaches) to enable relevant, informed AI adoption	
			MT22: Refine DoD AI development & T&E processes used throughout the AI lifecycle & "certify" what the processes are used for				

TOPICS

- practices to include T&E in requirements, contracts, RFPs/RFIs
- guidance to involve users in T&E
- guidance to define CONOPS, processes, HMT
- exemplar reusable measures
- iterative collection & sharing of tools, guidance, etc.
- policies for interoperability between T&E & developers

- collecting data at the tactical edge
- interoperable tools & model repositories
- AI assurance framework and cases
- practices for measuring HMT & user interaction
- infrastructure, storage, bandwidth
- interoperability standards for data, models, & tests

- playbook of exemplar reusable measures
- informing PMOs about threats
- flexible yet rigorous safety policy
- exemplar cybersecurity requirements
- cybersecurity threats for types of AIES
- development & T&E processes

MITRE

© 2023 THE MITRE CORPORATION. ALL RIGHTS RESERVED.

AIES: AI-enabled systems