DoD Manual 5000.96

Operational and Live Fire Test and Evaluation of Software

Originating Component:	Office of the Director of Operational Test and Evaluation
Effective:	December 9, 2024
Releasability:	Cleared for public release. Available on the Directives Division Website at https://www.esd.whs.mil/DD/.
Incorporates and Cancels:	Director of Operational Test and Evaluation Memorandum, "Guidelines for Operational Test and Evaluation of Information and Business Systems," September 14, 2010
Approved by:	Douglas C. Schmidt, Director of Operational Test and Evaluation

Purpose: In accordance with the authority in DoD Directive (DoDD) 5141.02 and the policy in DoD Instruction (DoDI) 5000.98, this issuance implements policy, assigns responsibilities, and provides procedures for operational test and evaluation (OT&E) and live fire test and evaluation (LFT&E) of DoD software-intensive systems and services and software embedded in systems and services (referred to in this issuance as "DoD systems") acquired via the Defense Acquisition System or via other non-standard acquisition systems.

TABLE OF CONTENTS

TABLES

FIGURES

SECTION 1: GENERAL ISSUANCE INFORMATION

1.1. APPLICABILITY.

This issuance applies to

a. The OSD, the Military Departments, the Office of the Chairman of the Joint Chiefs of Staff and the Joint Staff, the Combatant Commands, the Office of Inspector General of the Department of Defense, the Defense Agencies, the DoD Field Activities, and all other organizational entities within the DoD (referred to collectively in this issuance as the "DoD Components").

b. DoD software-intensive systems and software embedded in DoD systems acquired via the Defense Acquisition System, pursuing any adaptive acquisition framework pathway, in accordance with DoDD 5000.01 and DoDI 5000.02.

c. DoD software-intensive systems and software embedded in DoD systems under special access controls, in accordance with DoDD 5205.07.

d. Non-standard acquisition systems (e.g., Missile Defense System).

1.2. POLICY.

In accordance with DoDI 5000.98 and relevant DoD manuals and guidance, the DoD will plan, fund, execute, and report on OT&E and LFT&E of software-intensive systems or software embedded in systems to evaluate the operational effectiveness, suitability, survivability, and lethality (as applicable) of DoD systems in support of the delivery of each incremental capability.

SECTION 2: RESPONSIBILITIES

2.1. DIRECTOR OF OPERATIONAL TEST AND EVALUATION (DOT&E).

Pursuant to Sections 139, 4171, 4172, and 4231 of Title 10, United States Code and Section 223 of Public Law 117-81, the DOT&E reviews and approves exceptions to policy from this issuance for systems on the T&E Oversight List.

2.2. UNDER SECRETARY OF DEFENSE FOR RESEARCH AND ENGINEERING (USD(R&E)).

The USD(R&E):

a. For programs under T&E oversight for developmental test and evaluation (DT&E), assesses the adequacy and approves DT&E strategies documented in the Test and Evaluation Master Plan (TEMP), test and evaluation (T&E) strategy, or equivalent document, referred to in this issuance as "TEMP/T&E strategy."

b. For all other acquisition programs under DT&E oversight, advises the milestone decision authority (MDA) by conducting an independent analysis of test data, reports, modeling and simulation (M&S) results, and the adequacy of the DT&E plan in the TEMP/T&E strategy.

2.3. UNDER SECRETARY OF DEFENSE FOR ACQUISITION AND SUSTAINMENT (USD(A&S)).

The USD(A&S) enforces this issuance for DoD programs for which the USD(A&S) is the MDA.

2.4. DOD CHIEF INFORMATION OFFICER.

The DoD Chief Information Officer coordinates with the DOT&E, the USD(R&E), the USD(A&S), and the Under Secretary of Defense for Intelligence and Security (USD(I&S)) to synchronize the OT&E and LFT&E processes in this issuance with the DoD Cybersecurity Program

2.5. USD(I&S).

The USD(I&S) oversees intelligence support to programs throughout the acquisition life cycle and advises the DOT&E concerning intelligence supportability requirements that affect OT&E and LFT&E.

2.6. CHIEF DIGITAL AND ARTIFICIAL INTELLIGENCE OFFICER.

The Chief Digital and Artificial Intelligence Officer:

a. Establishes policy and issues guidance on definitions of requirements and testability for artificial intelligence (AI)-enabled systems to implement and demonstrate adherence to the DoD AI Ethical Principles established in the February 21, 2020 Secretary of Defense Memorandum and the DoD Responsible AI Strategy and Implementation Pathway.

b. Issues guidance, methodologies, and best practices on T&E for AI capabilities in DoD systems.

c. Coordinates with the USD(R&E) and the DOT&E on developing and using common tools and infrastructure for T&E and verification and validation of AI capabilities in DoD systems.

2.7. DOD COMPONENT HEADS.

The DoD Component heads follow the procedures outlined in this issuance through:

a. Component Acquisition Executives.

b. Program managers.

c. LFT&E organizations.

d. Their designated operational test agency (OTA) or operational test organization (referred to in this issuance as "OTA").

SECTION 3: SOFTWARE OT&E AND LFT&E OVERVIEW

3.1. SCIENCE- AND TECHNOLOGY- BASED SOFTWARE OT&E AND LFT&E.

a. The planning, execution, analysis, and reporting of software OT&E and LFT&E (e.g., survivability in contested cyberspace, electromagnetic spectrum attack) will be based on the latest advances in science and technology (e.g., equivalence partitioning, combinatorial testing, space-filling designs) to enable the:

(1) Determination of operational effectiveness, suitability, survivability, and lethality (as applicable) of software-intensive DoD systems or software embedded in DoD systems in the context of the over-arching DoD system with scientific rigor.

(2) Development and implementation of risk-based levels of test assessments and mission-based risk assessments (MBRAs) required to:

(a) Inform an appropriate scope of OT&E and LFT&E while considering the software development cadence and process, integration and deployment processes, and system usability.

(b) Characterize and quantify, where possible, risks to meeting OT&E and LFT&E objectives, the acquisition program, warfighter, mission engineering outcome, and DoD operations throughout the software development life cycle.

(3) Automated and manual software testing within software development including modern methods such as Agile and development, security, operations (DevSecOps).

(4) Definition, adequacy, and efficient coverage of planned OT&E and LFT&E during the various stages of software development.

b. Organizations conducting software OT&E and LFT&E planning, execution, analysis, and reporting will maximize the use of all relevant data and M&S results to inform acquisition decisions and evaluate whether the system will meet the intended operational effectiveness, suitability, survivability, and lethality (as applicable) as the software matures and adapts over time.

3.2. SOFTWARE OT&E AND LFT&E ACROSS THE ACQUISITION LIFE CYCLE.

a. Software is a critical element in the acquisition of most DoD systems pursuing any of the acquisition pathways within the Defense Acquisition System. T&E of software across the DoD system life cycle is necessary to provide decision makers with information on the technical requirements and operational effectiveness, operational suitability, survivability, and lethality (as applicable) of the DoD system. T&E of software-intensive DoD systems and software embedded in DoD systems must be included within the TEMP/T&E strategy for each DoD system.

b. The T&E organizations will leverage program and organizational sponsor user agreements to engage users, integrate operationally representative conditions across the acquisition life cycle, and enable real-time feedback throughout software development.

c. Software T&E planning, execution, analysis, and reporting will integrate into and support software development methodologies including modern methodologies (e.g., Agile and DevSecOps) and inform capability delivery throughout the entire life cycle of the software-intensive DoD system or software embedded in DoD systems. Software T&E cadence will align with the incremental software development cadence composed of a sequence of capability releases.

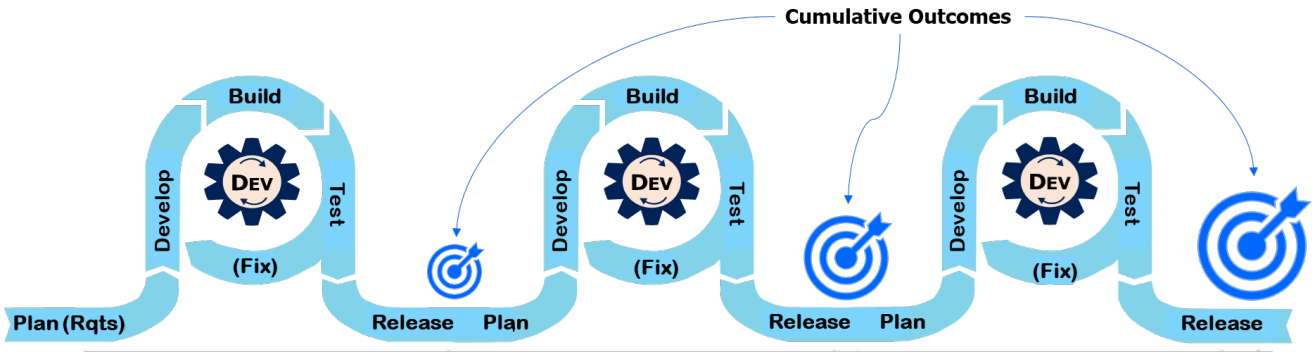
(1) Agile Development.

Figure 1 shows an example of an Agile software development cycle.

(a) Each iteration is preceded by planning activities in which the product owner, in conjunction with the operational users (referred to in this issuance as "users"), prioritizes, allocates, and decomposes the requirements for upcoming iterations to support rapid development while also defining "done" for the iteration. During each iteration, a small team will develop, build, and test the software and address any identified defects and vulnerabilities.

(b) Once sufficient quality criteria and metrics as documented in the T&E plan have been satisfied and informed by test results, the designated decision authority will "release" the software. In some circumstances, the software may be released to an external production environment for fielding. In most cases, the software will be released to an internal environment for integration or test purposes.

Figure 1. Notional Agile Iteration Cycle



(2) DevSecOps.

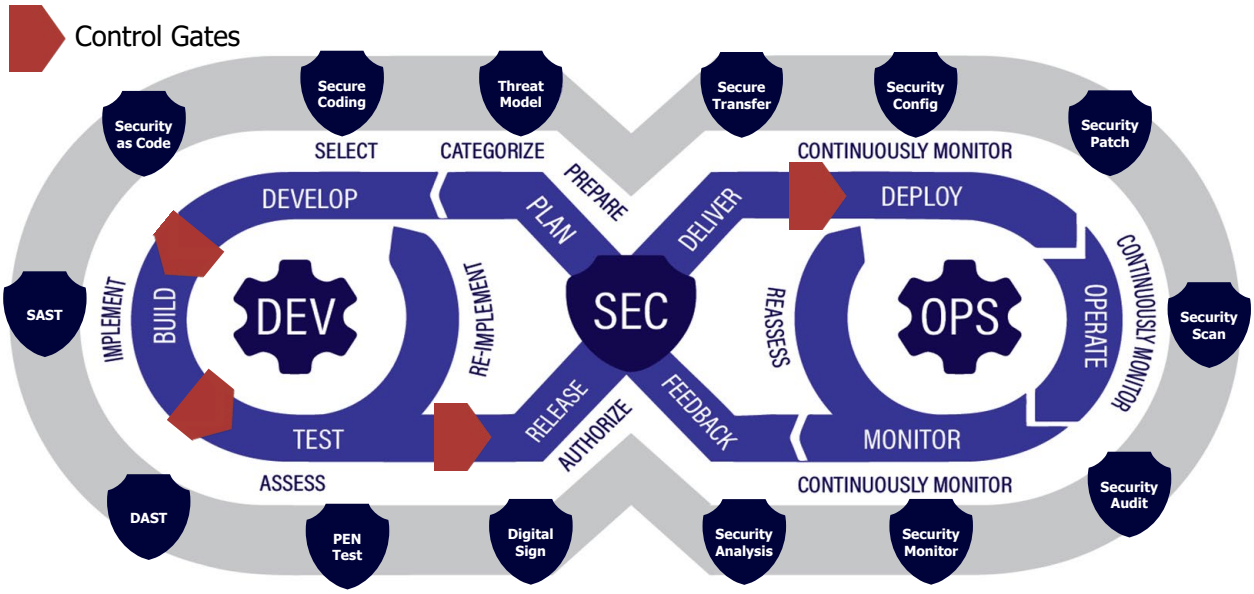
(a) Figure 2 shows the 10 phases of DevSecOps from build to deployment within which security is constantly integrated as a key component throughout engineering, development, and testing. Each control gate (displayed by the red polygons in Figure 2) represents a feedback mechanism and decision point to move forward to the next phase. Feedback generated from operations (e.g., escaped defects, help desk trouble reports) will inform

the T&E of future iterations or increments to better identify defects before the iterations or increments reach operations.

1. The development team makes lower-level decisions (e.g., decisions between develop and build; build and test), while the product owner makes other, higher-level decisions (e.g., decisions between test and release; release and deliver; deliver and deploy).

2. The T&E organizations will support the development of criteria for decision makers at software pre-deployments (as needed) and will plan and conduct testing to support decision makers at software deployment and subsequent releases (sometimes known as software development control gates).

Figure 2. DevSecOps Cycle with Control Gate Overlay

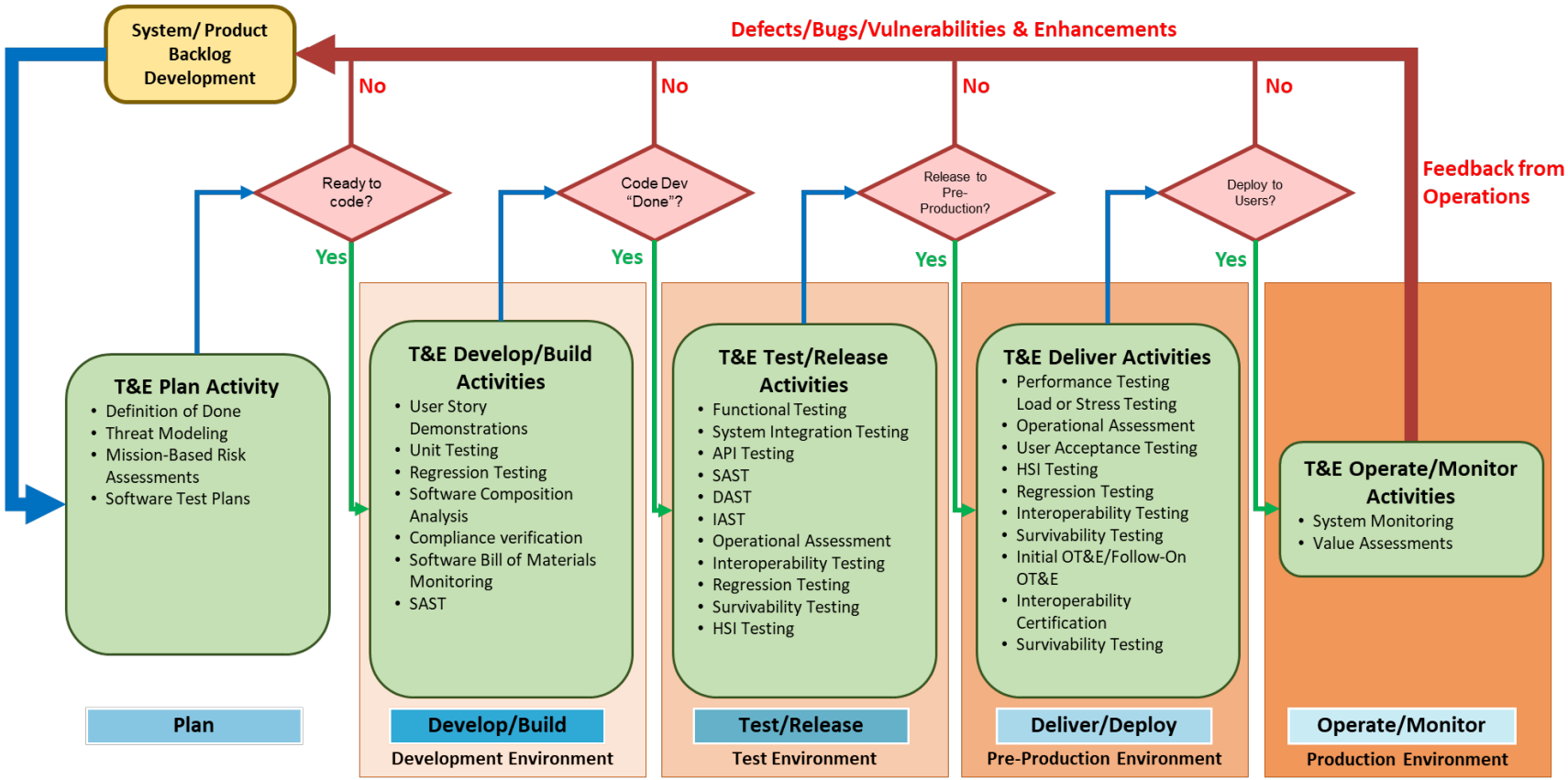


(b) Software development and operations may take place in an end-to-end architecture with separate development, test, pre-production (staging), and production environments. Figure 3 shows a notional software pipeline, noting various test activities that occur in support of software development control gate decisions. Blue arrows in Figure 3 indicate the test input to the decision. Decision criteria and decision ownership will be specified in the TEMP/T&E strategy.

(c) Software factories employ multiple pipelines, a suite of pipeline tools (for each of the steps within the development), process workflows, scripts, and environments to produce software. DoD systems may employ software factories in different configurations. These may include (but are not limited to):

1. Contractor owned and operated factories where the contractor performs testing during development and shares test data with the T&E Working-level Integrated Product Team (WIPT), also known as the integrated test team (ITT), (referred to in this issuance as "T&E WIPT/ITT").

Figure 3. Notional Software Pipeline



Acronyms: **API** – Application Programing Interface; **DAST** – Dynamic Application Security Testing; **HSI** – Human System Integration; **OT&E** – Operational Test & Evaluation; **IAST** – Interactive Application Security Testing, **SAST** – Static Application Security Testing, **SBOM** – Software Bill of Materials, **TTP** – Tactics, Techniques, & Procedures

2. Federal Government-owned, contractor-operated factories where the contractor and the Federal Government share test responsibilities.

3. Federal Government factories, where development teams and the T&E WIPT/ITT own all the processes.

(d) In all cases, the T&E WIPT/ITT will collaborate on test data from all test events to enable independent evaluation of DoD system using factory output. Factory outputs and processes may be critical to the system's survivability in contested cyberspace and the system's ability to adapt to changing mission and threat environments over time.

(e) Software development environments and DoD system test, pre-production, and production environments should be operationally representative to minimize risk and maximize confidence in test results.

(f) DevSecOps processes, tools, and testing must undergo survivability testing against operationally relevant threats throughout the software factory and its pipelines, especially when the factory connects to the operational system or operational environment.

(g) Table 1 summarizes the integrated T&E, OT&E, and LFT&E activities required across the software development life cycle. These activities include automated and manual test activities to meet OT&E and LFT&E objectives and facilitate initial and subsequent software releases or acquisition decisions throughout the acquisition life cycle. Different software factory configurations may impact organizational roles and responsibilities for executing the required test activities as explained in the footnotes in Table 1.

(h) The integrated decision support key (IDSK) will include events in Table 1, as applicable, and the data required to inform capability release or acquisition and program decisions (e.g., unique software factory control gate determinations). As the software development processes, practices, and tools mature over time, the T&E WIPT/ITT will update the IDSKs where appropriate.

Table 1. Baseline Software Test Data and Activities

Phase	Test Activities/ Data	Coordinated/ Conducted by	Supported by	Objective	Control Gate/ Decision to Move to Next Phase and Environment
Plan	Definition of "Done"	Program Manager	T&E WIPT/ITT	Ensure definition of "done" captures OT&E and LFT&E objectives	Ready to Code in Development Environment
	Threat Modeling	Program Manager	T&E WIPT/ITT	Identify threats to system	
	MBRA	Program Manager, Intelligence Community	T&E WIPT/ITT	Determine T&E priorities based on mission risk	
	Software Test Plans	Program Manager	Software Developer, Integrated T&E, OT&E, and LFT&E	Test Plans to capture integrated T&E, OT&E and LFT&E objectives	
Develop/ Build	User Story Demonstrations	Software Developer	OT&E, LFT&E	Capture and provide feedback	Ready to Test in Test Environment
	Unit Testing	Software Developer	OT&E, LFT&E	Capture code-level functionality; survivability, trace to higher level requirements	
	Regression Testing	Software Developer	OT&E, LFT&E	Ensure integrity of previously completed capabilities	
	Software Composition Analysis	Software Developer	OT&E, LFT&E	Identify open-source vulnerabilities and license issues	
	Compliance Verification	Software Developer	LFT&E	Ensure survivability risk management framework compliance standards met	
	Software Bill of Materials Monitoring	Software Developer	LFT&E	Capture software versions to identify vulnerabilities	
	Static Application Security Testing	Software Developer	LFT&E	Ensure source code quality	

Table 1. Baseline Software Test Data and Activities, Continued

Phase	Test Activities/ Data	Coordinated/ Conducted by	Supported by	Objective	Control Gate/ Decision to Move to Next Phase and Environment
Test/ Release	Functional Testing	Contractor T&E (CT&E)[1] Integrated T&E[2]	CT&E[3], OT&E[1]	Capture functionality at system level; trace to higher level requirements	Release to Ops Team/Pre-Production Environment
	System Integration Testing	CT&E[1] Integrated T&E[2]	CT&E[3], OT&E[1]	Ensure code segments combine to provide capability	
	Application Program Interface Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1] LFT&E	Evaluate software interfaces for functionality and security	
	Static Application Security Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Ensure source code quality	
	Dynamic Application Security Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Evaluate code quality during execution	
	Interactive Application Security Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Evaluate application quality through software instrumentation	
	Operational Assessment	OT&E	CT&E LFT&E	Ensure user end-to-end capability	
	Interoperability Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1] LFT&E	Ensure systems combine to provide capability	
	Regression Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Ensure integrity of previously completed capabilities	
	Survivability Testing	CT&E[1] LFT&E[2]	CT&E[3] OT&E[1] LFT&E	Identify vulnerabilities (to include software factory) for future mitigation	
	Human Systems Integration Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Initial evaluation of usability	

Table 1. Baseline Software Test Data and Activities, continued

Phase	Test Activities/ Data	Coordinated/ Conducted by	Supported by	Objective	Control Gate/ Decision to Move to Next Phase and Environment
Deliver	Performance Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Gather system level usage metrics	Release from Pre-Production/ Staging to Production Environment
	Load or Stress Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Ensure scalability of delivered capability	
	Operational Assessment	OT&E	CT&E[3] LFT&E[1]	Preliminary assessment of operational capability	
	User Acceptance Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Capture user feedback	
	Human Systems Integration Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Continued evaluation of usability	
	Regression Testing	CT&E[1] Integrated T&E[2]	CT&E[3] OT&E[1]	Ensure integrity of previously completed capabilities	
	Interoperability Testing	OT&E Joint Interoperability Test command (JITC)	CT&E LFT&E JITC	Ensure systems combine to provide capability	
	Survivability Testing	OT&E	CT&E LFT&E	Identify vulnerabilities (to include software factory) for future mitigation	
Deploy	OT&E (Initial OT&E / Follow-on OT&E)	OT&E	CT&E LFT&E	Determination of operational effectiveness, suitability, survivability, and lethality (as applicable)	Deploy to Users
	Interoperability Certification	OT&E JITC	CT&E LFT&E JITC	Certify system for interoperability	
	Survivability Testing	OT&E	CT&E LFT&E	Identify vulnerabilities for future mitigation	
Operate	System Monitoring	Operators	OT&E LFT&E	Capture operational performance	In Use and New Versions
	Value Assessment (see DoDI 5000.87)	Program Manager, User Community	T&E WIPT/ITT	Provide data to the program manager and user community	

1. Expected when software factory is contractor-owned.
2. Expected when software factory is Federal Government-owned.
3. Expected when software factory is Federal Government-owned, but the system has a contractor integrator.

3.3. OT&E.

OTAs must:

a. Plan and execute OT&E, as outlined in the TEMP/T&E strategy, to inform software releases to testing and production (made by the development team lead) and software deployment to users (made by the program manager). Table 1 highlights required OT&E activities.

b. Conduct independent OT&E when data collected within the software development cadence is insufficient to support evaluation of operational effectiveness, suitability, survivability, and lethality (as applicable) in operationally representative contested, congested, and constrained environments with trained users and maintainers.

c. Include trained users, cyber defenders (as applicable), and maintainers to evaluate whether the DoD system enables mission success.

d. Determine the scope and type of OT&E needed, depending on the acquisition pathway and related acquisition decision needs.

e. Confirm each test event incorporates the system-of-systems together with operationally representative information flows and operational users in an operationally realistic, contested, congested, and constrained environment.

f. Collect configuration management data, defect tracking data, and backlog management data to support evaluations throughout the acquisition life cycle and to support decisions for assessments on operational effectiveness and suitability while taking survivability and lethality (as applicable) into equal consideration.

g. Observe and evaluate the results of actions performed by users, cyber defenders, and maintainers in operationally representative contested, congested, and constrained conditions.

h. Adapt software OT&E designs to meet development requirement changes.

i. Integrate OT&E with the software factory timelines to assess whether the development process is reducing the risk to produce software with required operational effectiveness and suitability while taking survivability and lethality effects (as applicable) into equal consideration.

j. Confirm the presence, functionality, and efficiency of any rollback, failover, data backup, and continuity of operations procedures.

3.4. LFT&E.

a. Realistic, full spectrum survivability and lethality testing for software-intensive DoD systems and software embedded in DoD systems will follow the procedures outlined in DoD Manual (DoDM) 5000.99.

b. Full spectrum survivability T&E will use vulnerabilities and limitations identified in software survivability tests, including automatic software factory survivability testing.

(1) LFT&E practitioners will consider unique attack vectors into DoD systems and their associated supply chains from software factories. Cloud environments provide for realistic, full spectrum survivability testing.

(2) LFT&E practitioners will assess system countermeasures against adversarial attacks for software-intensive DoD systems and software embedded in DoD systems with unique defensive missions or capabilities.

(3) Risk-based level of test assessments and MBRAs that consider data collected in software factory security testing will inform full spectrum survivability testing scope and design.

c. Full spectrum lethality T&E for software-intensive DoD offensive capabilities and software embedded in DoD offensive capabilities will include:

(1) Deny, degrade, disrupt, deceive, destroy, exploit, or influence capabilities, as applicable.

(2) Known network or system defenses and adversarial tactics, techniques, and procedures designed to deter lethal action.

d. Realistic, full spectrum survivability and lethality test data for software-intensive DoD systems and software embedded in DoD systems will be acquired early and often to provide data for the evaluation of operational effectiveness and suitability.

3.5. CERTIFICATIONS.

The program manager must follow the certification procedures outlined in Paragraph 3.5. of DoDI 5000.98.

a. The risk management framework process must support gaining or retaining interim authorization to test, authorization to operate (ATO), and continuous ATO. The risk management framework does not replace the procedures identified in this issuance or DoDM 5000.99. Software capabilities and their corresponding software factories may specifically utilize continuous ATO to develop and deploy capabilities on a more frequent basis.

b. Cloud service providers must comply with the Federal Risk and Authorization Management Program (FedRAMP) established by the Office of Management and Budget.

c. Medical systems must comply with Public Law 104-191, also known as the "Health Insurance Portability and Accountability Act of 1996."

d. Financial systems must comply with the Financial Improvement and Audit Readiness Guidance and Chapter 4 of Volume 1 of DoD 7000.14-R.

e. Data collection for systems that include personally identifiable information and protected health information must comply with DoDI 5400.11; Volume 2 of DoDM 5400.11; DoDM 6025.18; and DoD 5400.11-R.

f. All information technology capabilities that have the potential to create, house, or use DoD records and related metadata must comply with DoDI 5015.02.

g. OTAs and LFT&E organizations must coordinate with interfacing system program management offices to use their test, pre-production, or staging environments to conduct interoperability certification and testing in accordance with DoDI 8330.01.

3.6. M&S.

The use of accredited M&S may be necessary to represent interfacing systems for functional performance, survivability, and interoperability testing in cases where test pre-production or staging environments are not available or not representative. M&S critical to meeting OT&E and LFT&E objectives must be verified, validated, and accredited in accordance with DoDI 5000.61 and DoDM 5000.102. Test teams will conduct appropriate human-systems integration planning for all human M&S activities pursuant to DoDI 5000.95.

3.7. MANAGEMENT OF OT&E AND LFT&E.

a. Program Manager.

The program manager must follow the procedures outlined in Paragraph 3.7.a. of DoDI 5000.98 and must:

(1) Codify user engagement needed for software development in a user agreement, which includes defining user resources for OT&E and LFT&E and training for users supporting OT&E and LFT&E.

(2) Integrate software OT&E and LFT&E planning and execution into the software pipeline and the systems engineering processes in coordination with the T&E WIPT/ITT.

(3) Incorporate automation throughout software development wherever economically or technically possible, to include software integration, testing, release, and monitoring.

(4) Ensure the OTA and LFT&E organizations have access to contractor information including test and other environments, facilities, repositories, test tools, test plans, test data, test reports, contractor support, automated test scripts, and contractor software pipelines.

(5) Ensure the OTA and LFT&E organizations will have an opportunity to review and provide input to contractor test plans to facilitate the generation of credible data that can inform formal evaluations.

(6) Budget for the cost of OT&E and LFT&E integration with software factories and DevSecOps platforms, data sharing, user participation, test tools and licenses, development of the test harness, test automation scripts and scenarios, the management of the software factory environments, and their maintenance.

(7) Develop and implement a data management strategy in accordance with Paragraph 3.8. of DoDI 5000.98 and provide the T&E WIPT/ITT access to actionable data (e.g., test metrics) through data dashboards or similar dynamic reporting mechanisms.

(8) Provide a test environment that is operationally representative of the pre-production and production environment including details on its verification and validation process sufficient for the OT&E and LFT&E community to assess the impact of such an environment on the planning and conduct of test activities.

(9) Provide all available factory test data to evaluate the security of the software factory and any of its effects on operational effectiveness and suitability.

b. T&E WIPT/ITT.

The T&E WIPT/ITT must follow the procedures outlined in Paragraph 3.7.b. of DoDI 5000.98 and must:

(1) Coordinate with the user community to confirm users are available for OT&E and LFT&E.

(2) Ensure the overall system OT&E and LFT&E incorporates software OT&E and LFT&E.

(3) Ensure that the user stories have traceability up to higher-level requirements, including required operational performance, and that the users report any discrepancies.

(4) Review development processes focused on code quality including pair-programming, test driven development, and code coverage to confirm the developer is building high quality software.

(5) Review team-oriented quality practices including acceptance criteria and the definition(s) of "done." These criteria and definition(s) will express the conditions that software must meet for a story or set of stories to proceed to the next stage in the pipeline.

(6) Review capability and system-level quality practices including acceptance-test driven development and behavior driven development to confirm the software is behaving as expected and that the developer is building the right software.

(7) Review overall software quality assurance in accordance with International Organization for Standardization/ International Electrotechnical Commission 25010.

(8) Confirm the T&E WIPT/ITT has competency in the automated test tools selected to manage the program. The T&E WIPT/ITT competency must include the ability to design automated test scripts to execute testing.

(9) Implement constant feedback mechanisms and be flexible to support the dynamic requirements and changing priorities over time, supporting planning flexibility, continuous improvement, and responsiveness to change.

(10) Review relevant code testing results within the context of vulnerabilities and relevant to user stories and operational missions.

(11) Request and make available the latest Intelligence Community threat assessment data relative to the software and operationally deployed system.

c. OTA.

The OTA must follow the procedures outlined in Paragraph 3.7.c. of DoDI 5000.98 and must:

(1) Support the development of capstone capability requirements throughout software development to ensure clarity and testability of epics, features, and user stories. OTA responsibilities are determined based on the software factory configuration.

(2) As resources allow:

(a) Integrate people, processes, and tools throughout software development cadence to embed OT&E perspectives into Agile and DevSecOps processes.

(b) Integrate test, data analysis, and evaluation processes and tools with the software factory processes to ensure timely delivery of test results and analysis to decision makers.

(c) Observe, support, and collect data from contractor iterations and support the planning and execution of any planned integrated T&E. OTAs will identify test requirements to be considered by the software development test teams and integrated T&E to enable early evaluation of operational effectiveness and suitability while taking into equal consideration survivability and lethality effects (as applicable).

(d) Provide risk assessments and evaluation to decision makers on whether the software will meet operational user requirements throughout the software development life cycle and the acquisition life cycle.

d. LFT&E Organizations.

The LFT&E organizations must follow procedures outlined in Paragraph 3.7.d. of DoDI 5000.98 and must:

(1) Identify test requirements to be considered by the software development test teams and integrated T&E to enable early evaluation of full spectrum survivability and lethality (as applicable).

(2) As resources allow:

(a) Integrate people, processes, and tools throughout software development cadence and embed LFT&E perspectives into Agile and DevSecOps processes.

(b) Integrate test, data analysis, and evaluation processes and tools with the software factory processes to ensure timely delivery of test results and analysis to decision makers.

(c) Observe, support, and collect data from contractor iterations and support the planning and execution of integrated T&E.

3.8. DATA MANAGEMENT.

a. The software factories, to include contractors' software factories, must provide software test results in near real time to a secure shared data repository to support major acquisition and program decisions across the software development life cycle.

b. Program relevant data will be visible, accessible, understandable, linked, trusted, interoperable, secure, and managed in accordance with the data management procedures outlined in Paragraph 3.8. of DoDI 5000.98.

3.9. DOT&E OVERSIGHT.

Programs on the T&E Oversight List for OT&E and LFT&E must follow the OT&E and LFT&E artifacts review and approval procedures outlined in Paragraph 3.9. of DoDI 5000.98.

SECTION 4: SOFTWARE OT&E AND LFT&E PROCESS

4.1. SOFTWARE OT&E AND LFT&E PLANNING.

Software OT&E and LFT&E planning will integrate with the software development cadence to support unique software capability decision making. OTAs and LFT&E organizations will take full advantage of automated test tools to develop and execute the plans and record the results.

a. Software Input to the TEMP/T&E Strategy.

In accordance with DoDI 5000.98 and DoDM 5000.100, the TEMP/T&E strategy must:

(1) Be informed by the risk-based level of test assessment and MBRA.

(2) Describe how the application and role of automated and manual software T&E maintain cadence with software development decision points.

(3) Describe how automated test management tools will support IDSK adjustment over time to capture both the progress against software requirements and the changes to those requirements.

(4) Describe the process for identifying and escalating software deficiencies and feedback mechanisms throughout the acquisition life cycle.

(5) Document plans for regression testing.

(6) Include software T&E with appropriate mission scenarios and user doctrine to determine operational effectiveness, suitability, survivability, and lethality (as applicable) in an operationally representative environment for each software release.

(7) Describe plans to evaluate operational suitability within the rapid software development cadence, to include human-system interactions (e.g., system usability, user workload, user trust in the system, reliability, maintainability, availability, and training).

(8) Include software performance testing (e.g., load and stress testing) to evaluate the capability to support the entire user base with provided infrastructure, computing power, and networking bandwidth.

(9) Describe software T&E across the software development life cycle, in accordance with Table 1, as applicable. Define T&E objectives aligned with capability release cadence, operational and test risk, and available resources.

b. Software OT&E and LFT&E Plans.

(1) In accordance with the procedures outlined Paragraph 4.1.b.(1). of DoDI 5000.98, OTAs and LFT&E organizations must develop detailed test plans supporting software

development efforts to meet OT&E and LFT&E objectives and ensure test analysis and evaluation supports decision making throughout the software life cycle.

(2) Test organizations may develop software OT&E and LFT&E plans for each test type (e.g., system integration, interoperability) for use in every iteration or for a single, integrated test plan per release.

(3) The TEMP/T&E strategy will outline the processes, frequency, and approval process for tests outlined in Table 1.

c. T&E Input to Acquisition Contracts.

T&E organizations must work with the program manager to inform acquisition contracts regarding:

(1) Contractor support to MBRAs.

(2) Delivery of the contractor's strategies for integrating software OT&E and LFT&E with functional contractor testing to support evaluation of operational performance impacts and potential survivability shortfalls.

(3) Delivery of data from the contractor testing listed in Table 1 to support OT&E and LFT&E.

(4) System safety analysis when testing software in physical systems.

(5) Contractor remediation of mission critical vulnerabilities and deficiencies prior to software deployment.

(6) Delivery of contractor reports demonstrating test-based validation that corrective action(s) taken by the contractor sufficiently mitigated the vulnerabilities and deficiencies identified during T&E.

4.2. SOFTWARE TEST PREPARATION.

The OTA and LFT&E organizations must conduct a test readiness review upon a capability release schedule in accordance with procedures outlined in Paragraph 4.2. of DoDI 5000.98. Test organizations will identify when to conduct and revisit test readiness reviews based on the MBRA or significant program changes.

4.3. SOFTWARE TEST EXECUTION.

The OTA and LFT&E organizations must execute software OT&E and LFT&E plans, respectively, in accordance with Paragraph 4.3. of DoDI 5000.98 throughout the software development and deployment life cycle. Integrated T&E, OT&E, and LFT&E will include test automation to the maximum extent possible to support planning and execution of functional,

regression, performance, and survivability testing along with the automated test results analysis in near real time.

4.4. SOFTWARE ANALYSIS AND EVALUATION.

a. The T&E WIPT/ITT will take full advantage of test management tools to provide visual representation of metrics, test progress, and other pertinent data via dashboards.

b. The OTA and LFT&E organizations must:

(1) Evaluate automated and manual test results for use in independent operational evaluations. They will focus on evaluating the overarching performance of the system in operationally representative conditions to supplement earlier testing focused on individual user stories.

(2) Determine preliminary and final operational effectiveness, suitability, survivability, and lethality (as applicable) including but not limited to:

(a) The effect of software defects, deficiencies, and supporting workarounds on operational effectiveness, suitability, survivability, and lethality (as applicable) including daily operations or mission completion.

(b) The ability of the system to scale and meet current and projected user loads.

(c) Adequacy of the process for updating training and documentation within each development cycle to support end users of fielded system changes.

(d) The effectiveness of rollback procedures that are performed during developmental T&E and validated by OT&E under operationally representative conditions.

(e) The incorporation of the survivability posture throughout the development and release pipeline and to support the supply chain analysis of the system.

4.5. SOFTWARE TEST REPORTING.

The OTA and LFT&E organizations must generate and deliver software OT&E and LFT&E reports in accordance with DoDI 5000.98. The test report should be streamlined to support release in synchronization with the development cadence. Available data dashboards will inform OT&E and LFT&E reports and enable the OTAs and LFT&E organizations to consolidate multiple automated tests within the reports.

SECTION 5: SOFTWARE T&E FOR ADAPTIVE ACQUISITION FRAMEWORK PATHWAYS

a. OT&E and LFT&E of software-intensive DoD systems or software embedded in DoD systems will support OT&E and LFT&E events, as appropriate, outlined for each of the adaptive acquisition framework pathways in Section 5 of DoDI 5000.98.

b. While software development may be managed as a separate line of effort, the overall acquisition program effort for any acquisition pathway will include adequate technical exchanges and ensure alignment with the hardware development (as applicable). Software development, including its OT&E and LFT&E, will conform to the weapon system architecture, design, and timelines.

(1) The program office will capture supporting software capability needs in a dynamic backlog of user stories. Backlog grooming and prioritization of these capabilities will be conducted with significant user input.

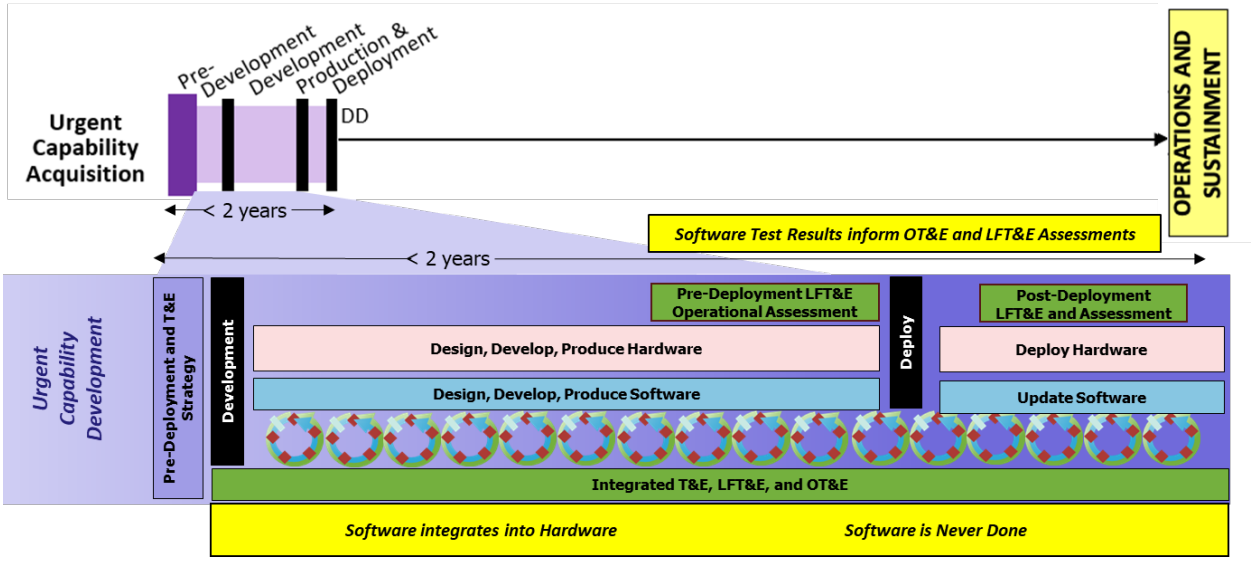
(2) As the overall program planning matures, the program manager must ensure that processes that deliver timely and adequate integration of hardware and software are included in the overall acquisition program and appropriately resourced.

c. The T&E WIPT/ITT must develop a TEMP/T&E strategy, merging test activities from the relevant acquisition pathway and the software development and must define the operationally realistic test criteria and environments for OT&E and LFT&E in support of software milestones while taking into equal consideration the objectives of the over-arching acquisition or program decisions.

d. Software OT&E and LFT&E will integrate with the software design approach and inform control gate decisions, which will influence acquisition and program decisions unique to each of the adaptive acquisition pathways. Software updates and relevant OT&E and LFT&E may be required throughout the life cycle of the program.

e. Figure 4 describes the applicability of the procedures in this issuance to the development of software embedded in a DoD system pursuing the urgent capability acquisition pathway.

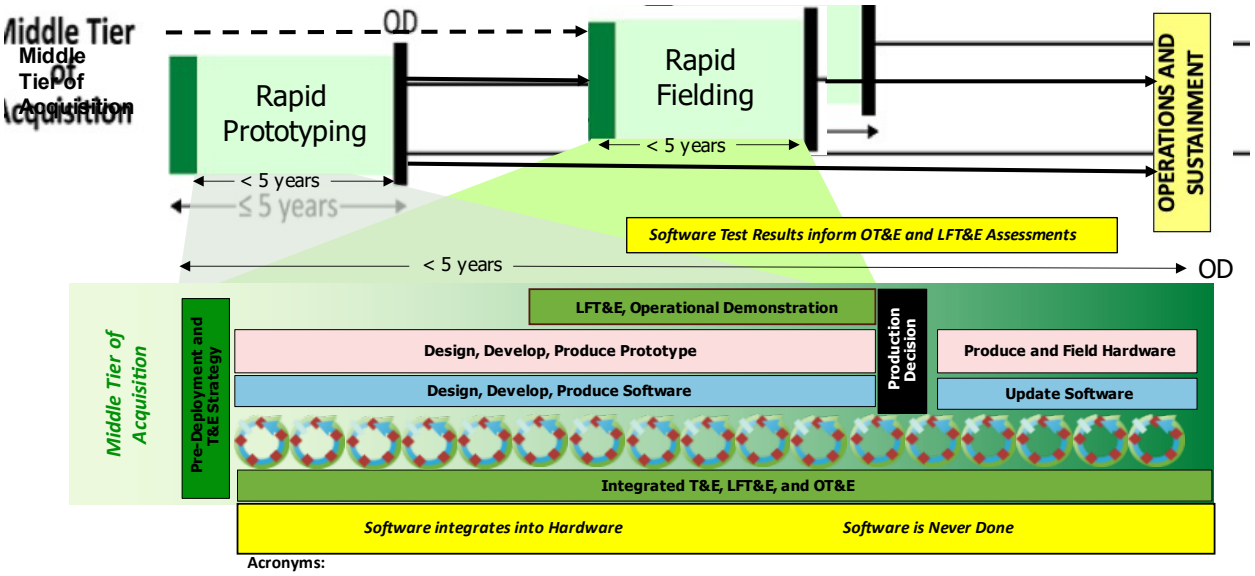
Figure 4. Software OT&E and LFT&E Considerations for the Urgent Capability Acquisition Pathway



Acronyms:
DD – Disposition Decision; **LFT&E** – Live Fire Test and Evaluation; **OT&E** – Operational Test & Evaluation

f. Figure 5 describes the applicability of the procedures in this issuance to the development of software embedded in a DoD system pursuing the middle tier of acquisition pathway.

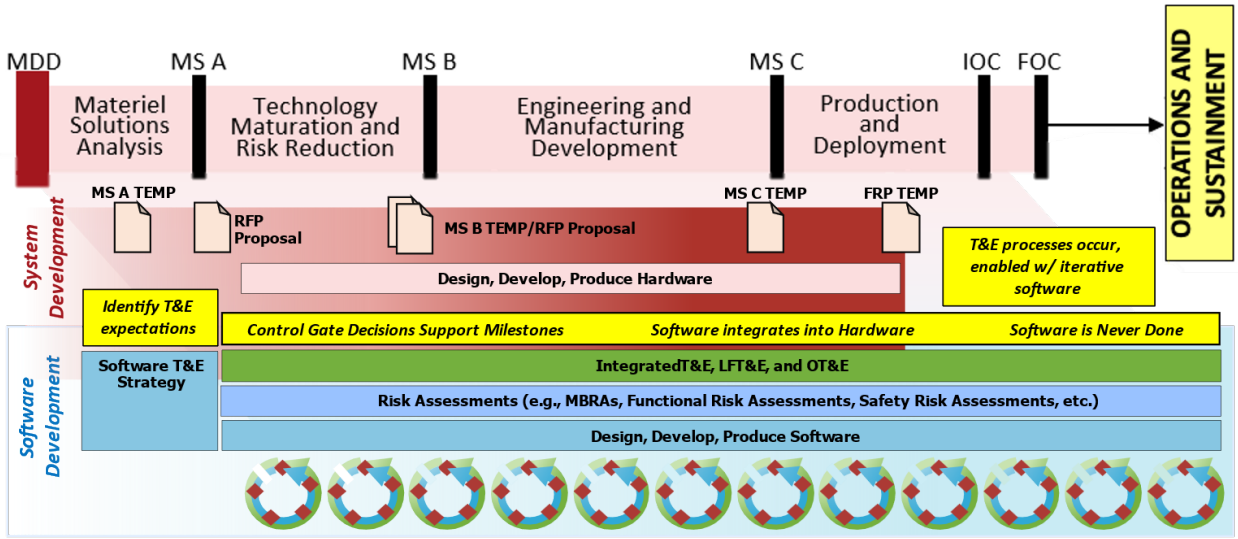
Figure 5. Software OT&E and LFT&E Considerations for the Middle Tier of Acquisition Pathway



Acronyms:
LFT&E – Live Fire Test and Evaluation; **OD –** Outcome Determination; **OT&E** – Operational Test & Evaluation

g. Figure 6 describes the applicability of the procedures in this issuance to the development of software embedded in a DoD system pursuing the major capability acquisition pathway.

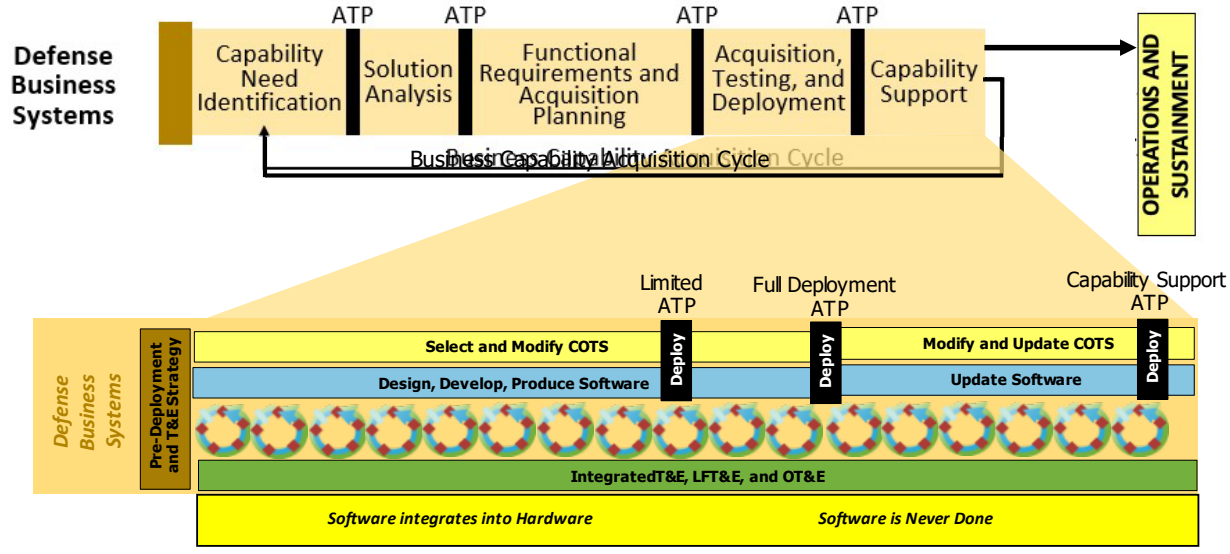
Figure 6. Software OT&E and LFT&E Considerations for the Major Capability Acquisition Pathway



Acronyms:
FOC – Full Operational Capability; **IOC** – Initial Operational Capability; **LFT&E** – Live Fire Test and Evaluation; **MBRA**s-- Mission-based Risk Assessments; **MS A** – Milestone A, **MS B** – Milestone B, **MS C** – Milestone C, **MDD** – Material Development Decision; **OT&E** – Operational Test & Evaluation; **RFP** – Request for Proposal

h. Figure 7 describes the applicability of the procedures in this issuance to the development of software embedded in a DoD system pursuing the defense business system pathway.

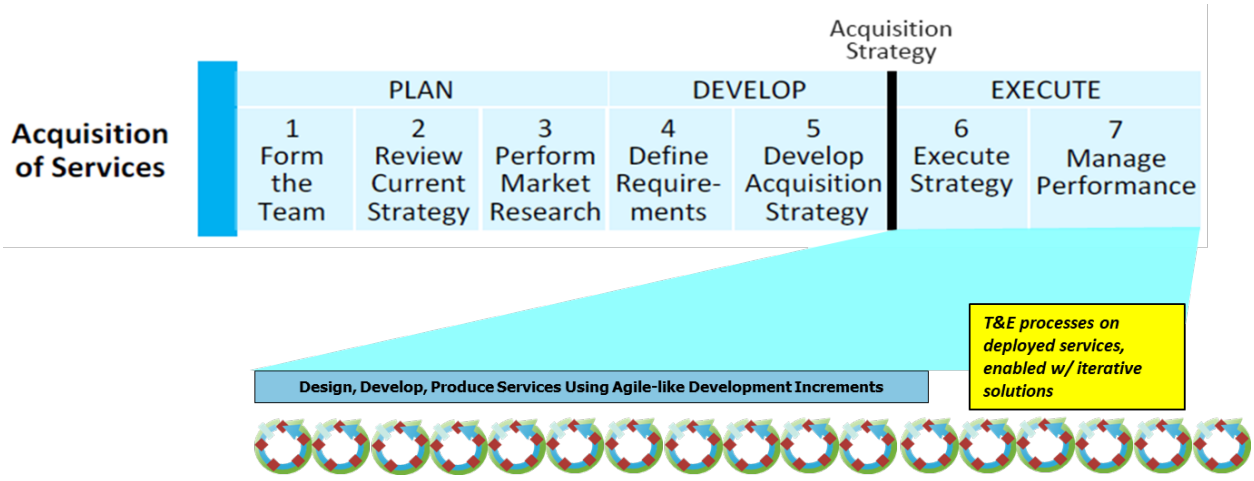
Figure 7. Software OT&E and LFT&E Considerations for the Defense Business System Pathway



Acronyms:
ATP – Authority to Proceed; **COTS** – Commercial Off-the-Shelf; LFT&E – Live Fire Test and Evaluation; **OT&E** – Operational Test & Evaluation

i. Figure 8 describes the applicability of the procedures in this issuance to the development of software embedded in a DoD system pursuing the acquisition of services pathway.

Figure 8. Software OT&E and LFT&E Considerations for the Acquisition of Services Pathway



GLOSSARY

G.1. ACRONYMS.

ACRONYM	MEANING
AI	artificial intelligence
ATO	authorization to operate
CT&E	contractor test and evaluation
DevSecOps	development, security, operations
DoDD	DoD directive
DoDI	DoD instruction
DoDM	DoD manual
DOT&E	Director of Operational Test and Evaluation
DT&E	developmental test and evaluation
FedRAMP	Federal Risk and Authorization Management Program
IDSK	integrated decision support key
ITT	integrated test team
JITC	Joint Interoperability Test Command
LFT&E	live fire test and evaluation
M&S	modeling and simulation
MBRA	mission-based risk assessment
MDA	milestone decision authority
OT&E	operational test and evaluation
OTA	operational test agency
T&E	test and evaluation
TEMP	test and evaluation master plan
USD(A&S)	Under Secretary of Defense for Acquisition and Sustainment
USD(I&S)	Under Secretary of Defense for Intelligence and Security
USD(R&E)	Under Secretary of Defense for Research and Engineering
WIPT	working-level integrated product team

G.2. DEFINITIONS.

Unless otherwise noted, these terms and their definitions are for the purpose of this issuance.

TERM	DEFINITION
Agile software development	A framework where a collaborative team of developers, test teams, and customer representatives divides requirements into smaller units of work, using terms like epics, features, and user stories. These smaller units of work, such as user stories, form the foundation for Agile software development. Agile frameworks may focus on the efforts of a single team (e.g., Scrum, Extreme Programming) or on multi-team efforts (e.g., Scaled Agile Framework®, Large-Scale Scrum, Scrum of Scrums) referred to as scaled Agile methods.
application program interface testing	A type of software testing performed by submitting requests to an application program interface to that analyzes and subsequently verifies that it fulfills its expected functionality, security, performance, and reliability. The tests occur as part of integration testing.
backlog	The single authoritative source that defines the complete set of work for an agile team or collection of teams. They may be associated with different levels of development (e.g., program or system level, release level, iteration level). Backlog items may encompass new features, feature enhancements, architecture activities, or infrastructure modifications. Backlog items are refined on an ongoing basis to include information such as priority, acceptance criteria, and estimates of effort.
code coverage	An analysis method that determines which parts of the software the test suite executed (covered) and which parts it did not execute (e.g., statement coverage, decision coverage, condition coverage).
combinatorial testing	A testing technique using multiple combinations of the input parameters to perform testing of the software product. The aim is to ensure that the product can handle different combinations or cases of the input configuration.
congested environment	Defined in DoDI 5000.98.
constrained environment	Defined in DoDI 5000.98.

TERM	DEFINITION
contested environment	Defined in DoDI 5000.98.
control gate	Mandatory checks within the DevSecOps process that provide explicit transparently understood exit criteria tailored to meet authorizing official and program risk tolerance. They can be automated by software pipeline processes or require human intervention. Integrated T&E and OT&E assessments also inform control gate determinations.
cyber defender	Defined in DoDI 5000.98.
definition of done	Defines the required conditions for a set of stories or work items to proceed to the next stage in the process or pipeline. Although "definition of done" originated as a Scrum term, all iteration-oriented Agile methods typically use this terminology. It may include criteria related to work item quality, process step completion, or tool and environment readiness. Each definition of done defines acceptance criteria for a handoff or promotion at a specific point in the overall lifecycle or pipeline.
DevSecOps	A software engineering culture and practice that builds on the Agile principles and aims at unifying software DevSecOps. DevSecOps focuses on culture, highlighting roles that emphasize responsiveness and collaboration, while combining security (also termed survivability) as a fundamental part of these transformations.
dynamic application security testing	Testing performed to analyze a running application dynamically and to identify runtime vulnerabilities and environment related issues.
embedded software	Software with a dedicated function within a larger mechanical or electrical system, often with real-time computing constraints, or software applications embedded in a platform (e.g., air vehicle, ground vehicle, space system, or ship). In the context of this issuance, embedded software does not apply to firmware or software dedicated to controlling devices.
epic	A term used in Agile development to represent a body of work that is too large to be completed in a single iteration.

TERM	DEFINITION
equivalence partitioning	A software testing technique that divides the input data of a software unit into partitions of equivalent data from which test cases can be derived to cover each partition at least once. This technique tries to define test cases that uncover classes of errors, thereby reducing the total number of test cases and test time.
FedRAMP	A government wide program that provides a standardized approach to security assessment, authorization, and continuous monitoring for cloud products and services.
functional testing	Testing that verifies the system's expected behavior given a set of inputs or actions. These tests verify delivery of user value by exercising the entire system.
human-systems integration testing	Testing to validate human factors engineering, manpower, personnel, training, safety and occupational health, force protection and survivability, and habitability. These domains are intimately and intricately interrelated and interdependent and must be among the primary drivers of effective, efficient, affordable, and safe system designs.
integration testing	Testing that ensures all separate functions, components, or systems are working together properly.
integrated T&E	Defined in DoDI 5000.98.
interactive application security testing	Analyzing code for vulnerabilities while running the application via an automated test, human tester, or any interacting activity.
interoperability certification	A formal statement of adequacy provided by the responsible interoperability certification authority agency that a system met its interoperability requirements.
interoperability testing	A structured event used to assess the technical exchange of information, data, and services and the end-to-end operational effectiveness of those exchanges.
LFT&E	Defined in DoDI 5000.98.
LFT&E organizations	Defined in DoDI 5000.98.

TERM	DEFINITION
load or stress testing	This testing executes the system at levels of performance at or above the level expected in operational missions. This testing may require virtualization or simulation to represent multiple users and a high throughput.
minimum viable capability release	The initial set of features suitable to be fielded to an operational environment that provides value to the warfighter or end user in a rapid timeline. It delivers initial warfighting capabilities to enhance some mission outcomes. It is analogous to a minimum marketable product in commercial industry.
minimum viable product	An early version of the software to deliver or field basic capabilities to users to evaluate and provide feedback on. Insights from minimum viable products help shape scope, requirements, and design.
non-functional test	A test to ensure the system meets quality characteristics that functional testing does not capture. Examples include performance, security, and usability tests.
OT&E	Defined in DoDI 5000.98.
product owner	Agile team member primarily responsible for maximizing the value delivered by the team by ensuring that the team backlog is aligned with customer and stakeholder needs.
performance testing	Testing that verifies the system meets functional and non-functional performance requirements, such as response time, latency, and throughout, specified in the requirements documentation.
pipeline	A sequence of orchestrated, automated tasks implementing the software delivery process for a new application version.
program manager	Defined in DoDI 5000.98.
regression testing	A type of testing, triggered by a modification to a component or system, that detects if defects have been introduced or uncovered in unchanged areas of the software or system under test.
rollback	Returning a hardware product or software program back to an earlier version after encountering issues with a later version.

TERM	DEFINITION
software assurance	The level of confidence that the software functions as intended and is free of vulnerabilities, either intentionally or unintentionally designed or inserted as part of the software throughout the lifecycle.
software bill of materials	A nested inventory or a list of ingredients that make up software components. It is a key building block in software security and software supply chain risk management.
software capabilities	Software-intensive systems or software-intensive components or sub-systems derived from high-level requirements.
software composition analysis	The performance of dependency vulnerability checking to identify vulnerabilities in open source-dependent components.
software factory	An organized software development and delivery approach that contains multiple pipelines, which are equipped with a set of tools, process workflows, scripts, and environments to produce a set of software deployable artifacts with minimal human intervention. It automates the activities in the develop, build, test, release, and deliver phases.
software-intensive system	A system in which software represents the largest segment in one or more of the following criteria: system development cost, system development risk, system functionality, or development time.
software quality assurance	The means and practice of monitoring all software engineering processes, methods, activities, and work items to ensure compliance against defined standards.
space-filling design	A method of arranging experimental points to evenly cover the experimental space, ensuring representative sampling across the entire region of interest. It is used in statistics and experimental design to provide a comprehensive understanding of complex systems.
sprint demonstration	An end-of-iteration demonstration providing users an opportunity to see delivered capability in the iteration and revise their priorities. It gives them a chance to say, "That's what I said, but it's not what I meant." This is a form of critiquing the product.
static application security testing	A testing technique to scan source code, binary, or byte code and components to identify potential vulnerabilities in software architecture.

TERM	DEFINITION
system testing	Testing performed on a complete system to evaluate its compliance with specified requirements.
test harness	Tools, libraries, and software components designed to facilitate the automated testing of software applications. Libraries, drivers, test data, logging and reporting structure, and setup and teardown procedures are components of the test harness and support the execution of automated test cases to allow developers to assess the functionality and performance of their code.
T&E	Defined in DoDI 5000.98.
unit testing	Testing, commonly performed while program code is under development, that verifies the behavior of a small part of the overall system. It may be as small as a single object or method that is a consequence of one or more design decisions.
user acceptance testing	A type of testing performed by the end user or the client to verify/accept the software system before moving the software application to the production environment. This testing is done in the final phase of testing after functional, integration, and system testing is done.
user story	A short natural language description of requirements expressed from a user perspective. It specifies the type of user, what they want to be able to do, and what value they want to obtain. User stories are the smallest unit of work in agile development and must be sized to be completed within a single iteration. Prior to implementation, user stories must be elaborated to incorporate acceptance criteria to enable the development team to accurately understand and meet user needs.
user story demonstrations	These tests guide coders as they write the code and help the team know when it has met the customers' conditions of satisfaction. After users have a chance to see what the team is delivering, they might have different ideas about how they want it to work.

REFERENCES

DoD 5400.11-R, "Department of Defense Privacy Program," May 14, 2007

DoD 7000.14-R, Volume 1, "Department of Defense Financial Management Regulation (DoD FMR): General Financial Management Information, Systems and Requirements," current edition

DoD Directive 5000.01, "The Defense Acquisition System," September 9, 2020, as amended

DoD Directive 5141.02, "Director of Operational Test and Evaluation (DOT&E)," February 2, 2009

DoD Directive 5205.07 "Special Access Program Policy," September 12, 2024

DoD Instruction 5000.02, "Operation of the Adaptive Acquisition Framework," January 23, 2020, as amended

DoD Instruction 5000.61, "DoD Modeling and Simulation Verification, Validation, and Accreditation," September 17, 2024

DoD Instruction 5000.87, "Operation of the Software Acquisition Pathway," October 2, 2020

DoD Instruction 5000.95, "Human Systems Integration in Defense Acquisition," April 1, 2022

DoD Instruction 5000.98, "Operational Test and Evaluation and Live Fire Test and Evaluation," December 9, 2024

DoD Instruction 5015.02, "DoD Records Management Program," February 24, 2015, as amended

DoD Instruction 5400.11, "DoD Privacy and Civil Liberties Programs," January 29, 2019, as amended

DoD Instruction 8330.01, "Interoperability of Information Technology, Including National Security Systems," September 27, 2022

DoD Manual 5000.99, "Realistic Full Spectrum Survivability and Lethality Testing," December 9, 2024

DoD Manual 5000.102, "Modeling and Simulation Verification, Validation and Accreditation in Test and Evaluation," December 9, 2024

DoD Manual 5000.100, "Test and Evaluation Master Plans and Test and Evaluation Strategies," December 9, 2024

DoD Manual 5400.11, Volume 2, "DoD Privacy and Civil Liberties Programs: Breach Preparedness and Response Plan," May 6, 2021

DoD Manual 6025.18, "Implementation of the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule in DoD Health Care Programs," March 13, 2019

DoD Responsible AI Working Council, "U.S. Department of Defense Responsible Artificial Intelligence Strategy and Implementation Pathway," June 2022

Office of the Under Secretary of Defense (Comptroller)/Chief Financial Officer, "Financial Improvement and Audit Readiness (FIAR) Guidance," April 2017

Public Law 104-191, "Health Insurance Portability and Accountability Act of 1996," August 21, 1996

Public Law 117-81, "National Defense Authorization Act for Fiscal Year 2022," December 27, 2021

Secretary of Defense Memorandum, "Artificial Intelligence Ethical Principles for the Department of Defense," February 21, 2020

United States Code, Title 10