AUTONOMOUS SYSTEMS DT

Enabling effective T&E of autonomous systems using
Digital Modeling and Simulation and **Safety-Enabled Testing**

Paul J. **Kwashnak**
U.S. Army Test and Evaluation Command
paul.j.kwashnak.civ@mail.mil

Purpose

Provide an informational brief on “**Enabling Effective Test and Evaluation (T&E) of Autonomous Systems using Digital Modeling and Simulation (DMS)**” in support of the 37th International T&E Symposium.

Agenda

- Background
- Test Approach
- M&S Test Services
- Test Design
- Emerging Program (ASTC)

Outcomes

- Define a T&E approach using DMS to test autonomous ground systems
- Recognize various digital test services being deployed as a part of an emerging ecosystem
- Identify personnel in the T&E community to enhance collaboration and further investigation

T&E of Autonomous Systems

Autonomy and Artificial Intelligence (AI) have been identified as critical military technologies by the Office of the Under Secretary of Defense Research and Engineering (OUSDR&E)

Challenge

- Current methods rely on observing developer testing or conducting a minimal number of live scenarios; insufficient to **fully validate and "build trust"** that the system will operate as intended
- To achieve a **statistical confidence for safe operation**, the autonomy software can be stressed in a **virtual test environment** in a wide array of scenarios to calculate the probability of anomalies by varying test factors such as sensor inputs, weather conditions, obstacles, and degraded communications

"Decision engine testing will primarily rely on M&S as a strategy"
- *FY 2018-FY 2028 Strategic Plan for DoD T&E Resources*

"M&S must mature [as] a test resource at the MRTFB that enables cost-effective and increasingly comprehensive autonomy evaluation of perception and comprehension."
- *Autonomy Test and Evaluation Infrastructure Gap Identification Report, GTRI*

"Robust M&S is essential for T&E of these future systems"
- *Analytical and Technical Support to the U.S. Army Evaluation Center (AEC) Study, IDA*

"The development of more extensive modeling and simulation will be key to accelerating testing and conserving resources."
- *Technical Assessment: Autonomy, OTI*

"M&S will become increasingly important ... to be able to precisely stimulate the system under test and extract the internal states of key autonomous functions/components such that perception, reasoning, decision making and learning can be rigorously tested"
- *Workshop Report: Test and Evaluation of Autonomous Systems, STAT T&E COE*

Documented **Need**

"Current DoD infrastructure and capabilities are inadequate to safely test and assess Autonomous Systems performance." [1]

"T&E Exec Need: FY18 (IC4) Test instrumentation capable of recording and analyzing Autonomous and Semi-autonomous systems sensory inputs and decision logic, and test controls that prevent such systems from leaving prescribed safety boundaries or creating a hazard to others." [2]

- ❑ **Significant Gap Finding 1 – Effective T&E Modeling and Simulation (M&S)** [3]

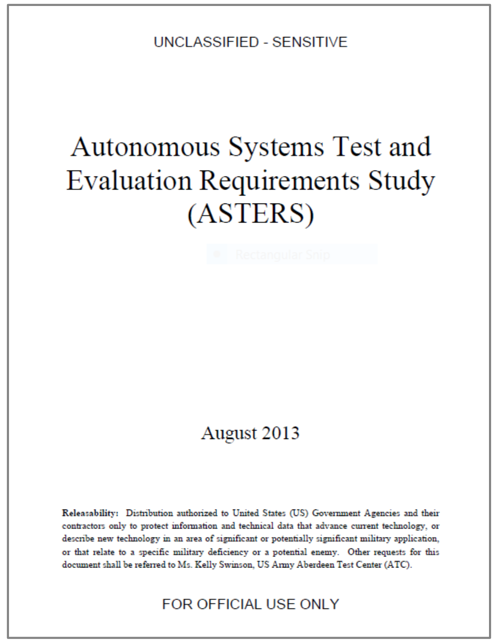
  "M&S must mature from a tool used during development, or limited HWIL support, to a test resource at the MRTFB that enables cost-effective and increasingly comprehensive autonomy evaluation of perception and comprehension."

- ❑ **A validated soft-stop capability is critical for system and personnel safety** [3]

  "While emergency-stop technology is widely available, those systems bring vehicles to an immediate uncontrolled stop without regard for potential vehicle damage. Having a convoy of large, high-speed vehicles being brought to an uncontrolled stop is not the optimal technique for aborting a test scenario."

System complexity results in an inability to test under all known conditions, difficulties in objectively measuring risk, and failures in the autonomous processing and decision-making software that are not discovered until dedicated OT may be very costly, and perhaps even impossible to fix

1. TRMC. "Autonomous Systems Test and Evaluation Requirements Study (ASTERS)", August 2013.
2. DoD T&E Infrastructure Investment Guidance for 2015 Memorandum. December 2014.
3. TRMC. "Autonomy Test and Evaluation Infrastructure Gap Identification Report", August 2017



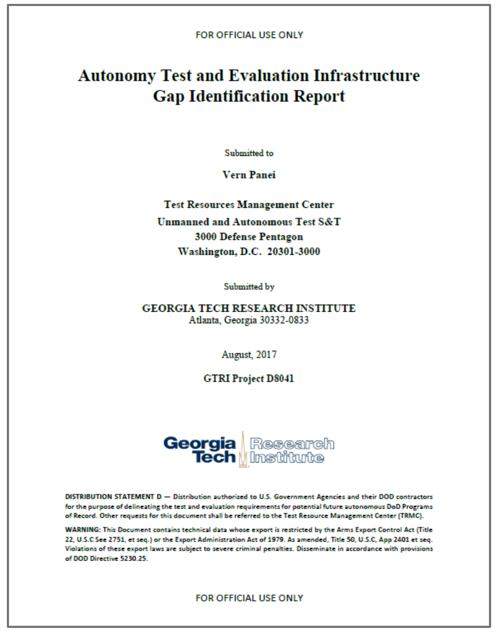
UNCLASSIFIED - SENSITIVE

Autonomous Systems Test and Evaluation Requirements Study (ASTERS)

August 2013

Releasability: Distribution authorized to United States (US) Government Agencies and their contractors only to protect information and technical data that advance current technology, or describe new technology in an area of significant or potentially significant military application, or that relate to a specific military deficiency or a potential enemy. Other requests for this document shall be referred to Ms. Kelly Swinson, US Army Aberdeen Test Center (ATC).

FOR OFFICIAL USE ONLY



FOR OFFICIAL USE ONLY

Autonomy Test and Evaluation Infrastructure Gap Identification Report

Submitted to

Vern Panei

Test Resources Management Center
Unmanned and Autonomous Test S&T
3000 Defense Pentagon
Washington, D.C. 20301-3000

Submitted by

GEORGIA TECH RESEARCH INSTITUTE
Atlanta, Georgia 30332-0833

August, 2017

GTRI Project D8041

Georgia Tech Research Institute

DISTRIBUTION STATEMENT D — Distribution authorized to U.S. Government Agencies and their DOD contractors for the purpose of delineating the test and evaluation requirements for potential future autonomous DoD Programs of Record. Other requests for this document shall be referred to the Test Resource Management Center (TRMC).

WARNING: This Document contains technical data whose export is restricted by the Arms Export Control Act (Title 22, U.S.C See 2751, et seq.) or the Export Administration Act of 1979. As amended, Title 50, U.S.C, App 2401 et seq. Violations of these export laws are subject to severe criminal penalties. Disseminate in accordance with provisions of DOD Directive 5230.25.

FOR OFFICIAL USE ONLY

VISION

Utilize **model based techniques** threaded together by digital practices, **integrated data management and analysis** tools and computing infrastructure to enable rapid testing and data insights of autonomy and AI at the ***speed of relevance***

Digital Test Ecosystem

Mr. Radar [Online Image]. https://spaceballs.fandom.com/wiki/Mr._Radar

By characterizing autonomous behavior, virtual testing aids in:

- Identifying economical test scenarios
- Understanding indicators of catastrophic or critical anomalies
- Reducing risk of failure

Stakeholders are able to:

- Conduct a safety review based on the findings
- Identify issues that require corrective action
- Recommend additional scenarios that need concentrated testing

Test Approach

Integrated, progressive methodology

DMS

Develop early trust using Digital Modeling and Simulation (DMS)

- Test autonomous systems in a simulated environment
- Lab-based system autonomous controller safety and performance assessments
- Simulation of system sensors
- Expose any safety concerns, security or performance deficiencies

Integration

System level integration testing via Hardware in the Loop (HWIL)

- Test autonomous systems in a controlled environment
- Simulation and stimulation of system sensors
- Test stability and controllability of the autonomous system

Live Test

Test in a safety-enabled live test environment

- Execute live testing with safety controls
- Safety monitoring and anomaly detection
- Bring systems under control in the event of undesired behavior

Modeling and Simulation

Comprehensive M&S serves as a crucial enabler to address this intractably large test space to facilitate effective and affordable test and evaluation (T&E) of autonomy perception and comprehension.[1]

An integrated M&S approach provides:

- Enhanced lab-based test environments
- Tools and reference libraries
- Local models and synthetic effects
- Range connectivity to core test facilities

Advanced modeling and simulation tools and techniques will:

- Accelerate safety certification for technology demonstration and Warfighter use
- Expedite acquisition and fielding by reducing developmental test iterations in live environments
- Achieve statistical confidence in safe autonomous behaviors

1. TRMC. “Autonomy Test and Evaluation Infrastructure Gap Identification Report”, August 2017

Use Case

Logistics Transport

Large, heavy haul UGVs that transport supplies between ports and bases. Considerations are explored for mixed manned and unmanned systems up to fully autonomous convoys.

Mission Tasks

Base-to-base cargo haul
- Navigate // Route Planning
- Road Perception
- Traffic negotiation

Hazardous environment operations
- Dynamic Routing
- Fail-safe recovery
- Combat Evasion

Convoy operations
- Structured Verbal Interaction
- Collision Avoidance
- Formation Keeping

M&S
EXECUTION PLAN

To achieve this **vision**, the U.S. Army must undertake a holistic approach to the development of autonomous and intelligent systems. This approach cannot consider **acquisition, development, testing, and simulation** independently, but must develop an approach by which each of these **mutually support** each other.

Multiple organizations must collaborate to include the Program Manager (PM), the Original Equipment Manufacturer (OEM), U.S. Army Test and Evaluation Command (ATEC) and Combat Capabilities Development Center (CCDC).

Science and Technology

AGR STO/TWV LF DR HIL SIL

Acquisition

Continuous Autonomy Simulation Test Laboratory Environment (CASTLE)

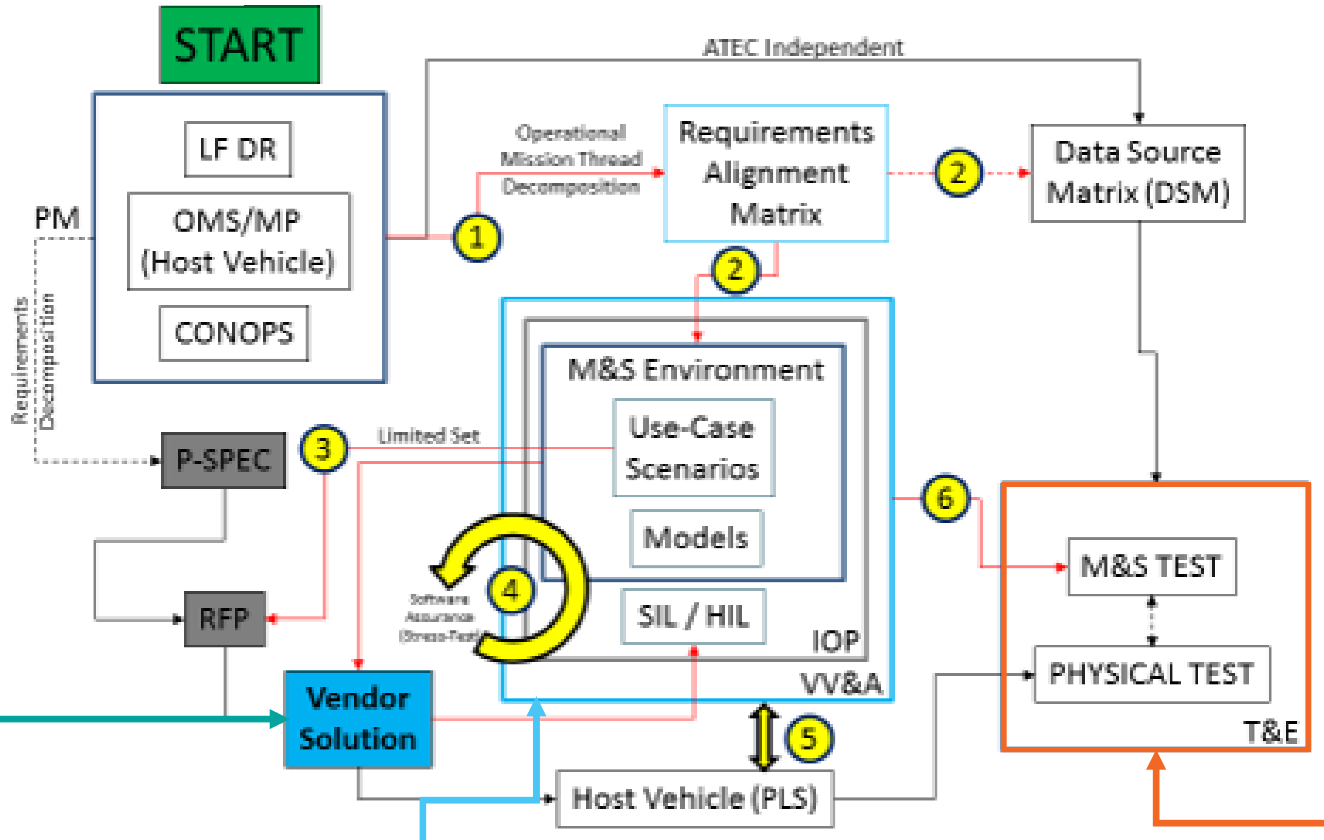
Test and Evaluation

Autonomous System Test Capability (ASTC)

Test Process

Current test chain

Modeling & Simulation (M&S) Process Plan for Robotic Systems
13 June 2017; Brian McVeigh, PM Force Projection



Focusing Improvements

Development
- Testability
- M&S Integration

Simulation
- Automation
- High-volume
- Scenario/environment
- Real-world data

Testing
- Safety confirmation and validation
- Controlled experiments
- Data collection
- Edge cases

Developers

Development Test Environment
- Unit Level SW Testing (White-box)
- Integration Testing
- Component Interface Testing
- System Acceptance Testing

AFC

Pre-Production Test Environment
- Model Verification
- HIL/SIL (Component level) Testing
- Model Validation
- Plug-in Testing

ATEC

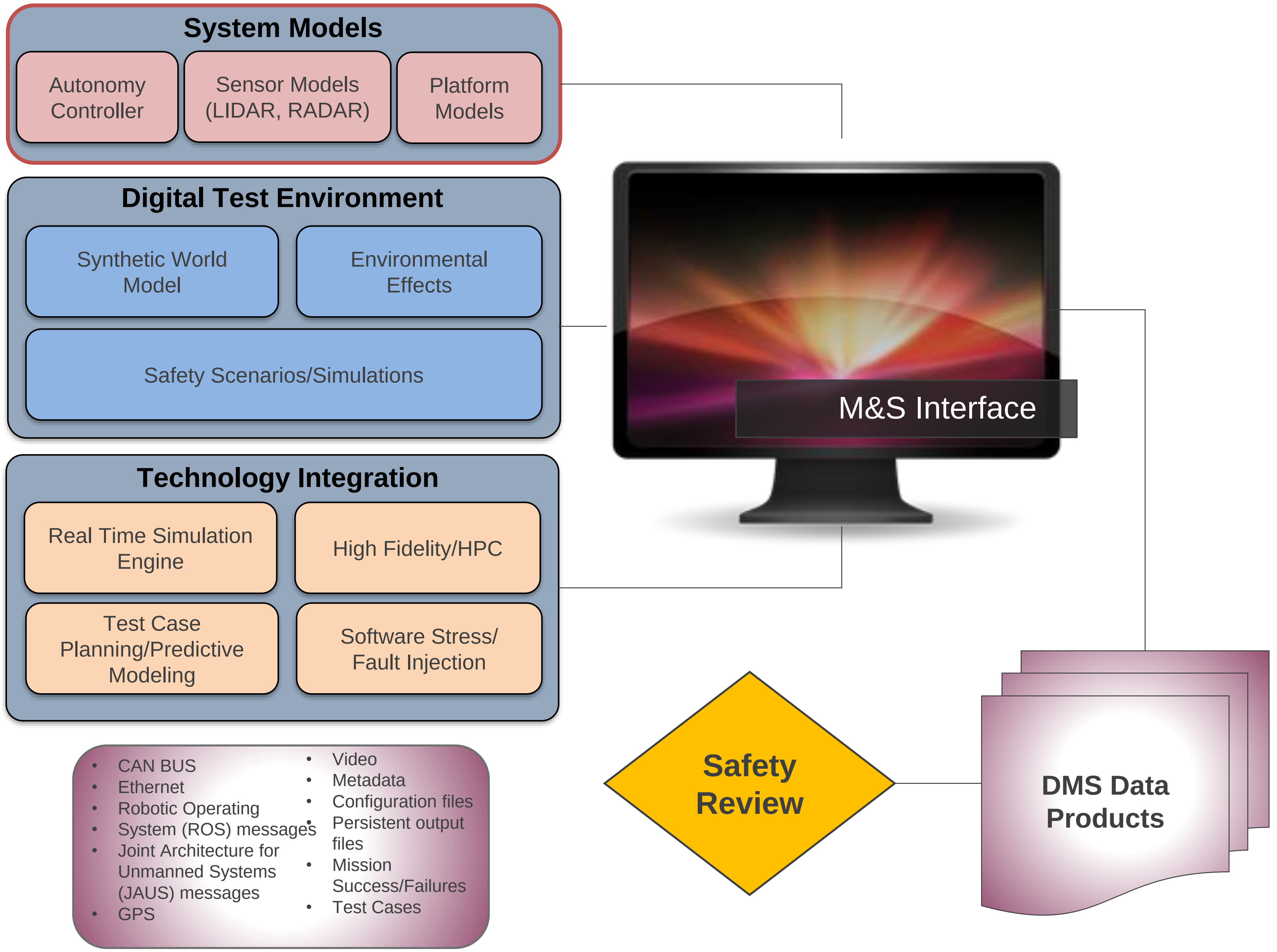
Production Test Environment
- Safety/Developmental Testing
- Software Assurance (Stress-Test)
- Software-in-the-Loop (System Level)
- Hardware-in-the-Loop (System Level)

Digital **M&S**

Modeling and Simulation

Digital Test Environment

- Physics-based simulation environment to characterize autonomy behavior while executing pre-defined scenarios to facilitate safety analysis
  - Varying road surfaces and grades
  - Buildings
  - Dynamic vehicles
  - Obstacles
  - Vegetation



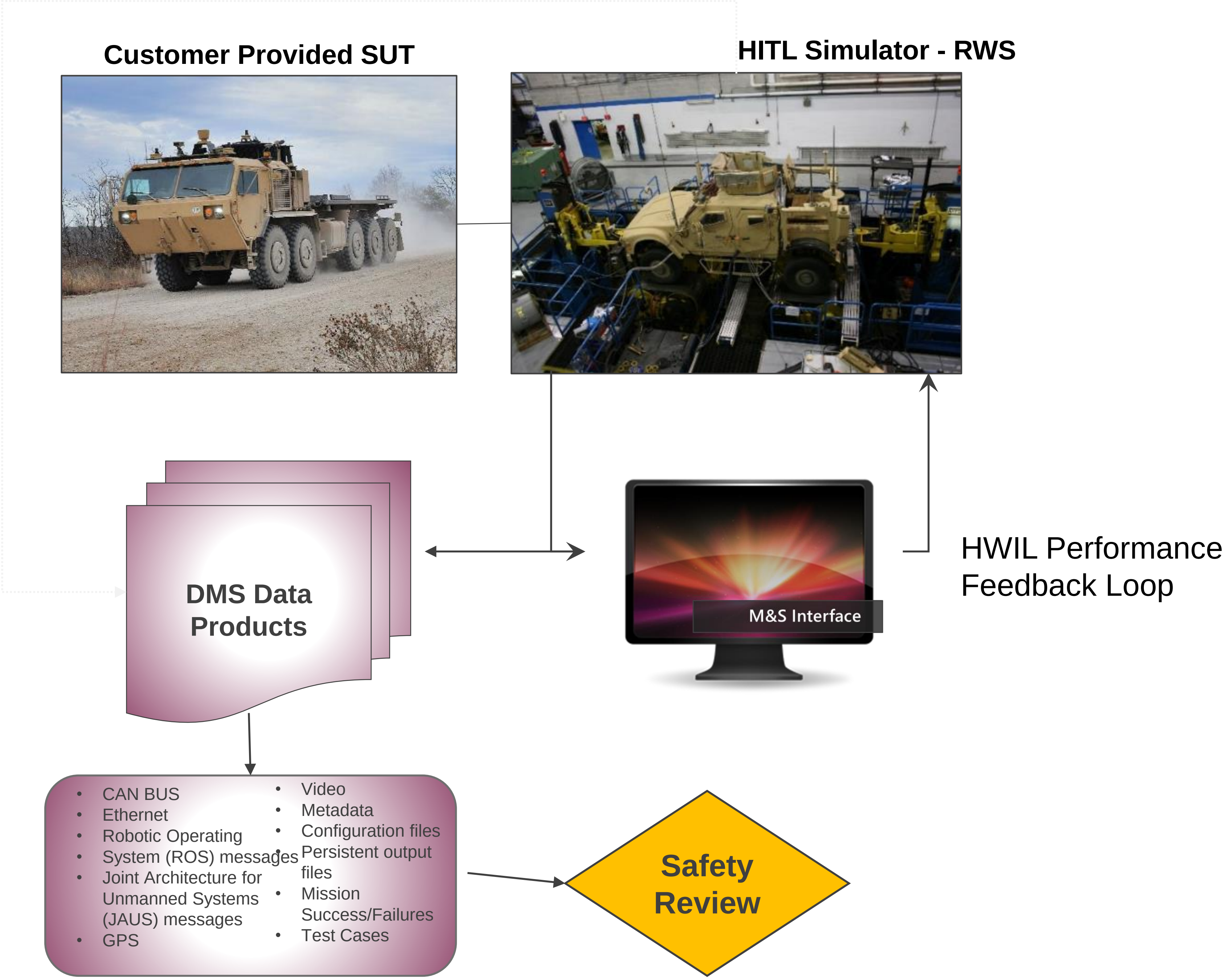
Open air ranges (OAR) are digitally constructed using data collected from Light Detection and Ranging (LIDAR) scans and photographic images to replicate test infrastructure and conditions such as existing Major Range Test Facility Bases (MRTFB).

Digital to **Physical**

Hardware in the Loop

HWIL Test Environment

- Test environment synchronized with a virtual test environment, integrated with automated data processing tools, capable of characterizing autonomous system physical components and system behaviors while executing pre-defined scenarios to facilitate safety analysis



Testing Workflow

Progressive T&E sequence

Autonomy Controller

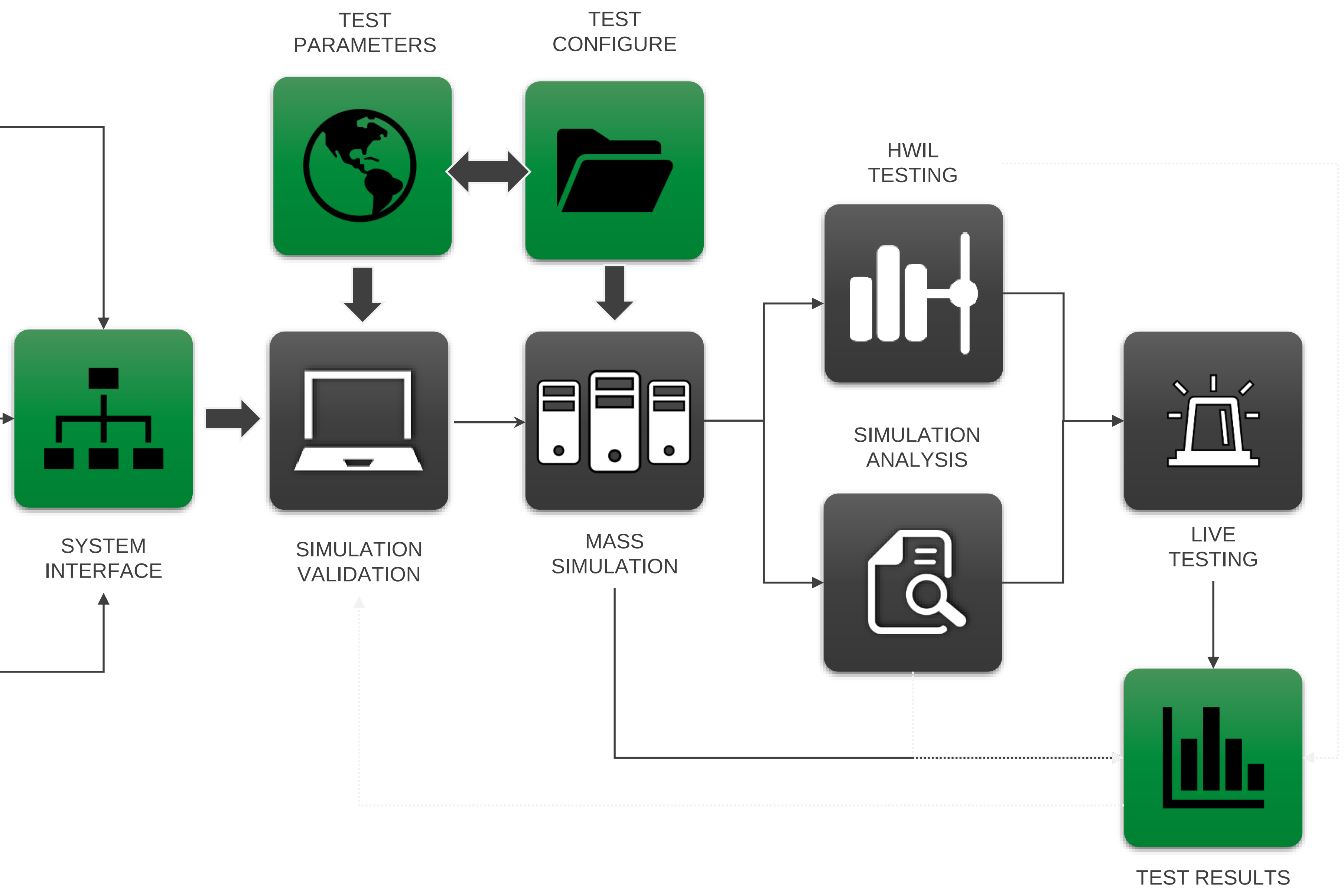
ROSIntegration Plugin for Unreal Engine 4 was used to support the integration with AGR/ExLF and RTK Autonomy SW

Platform Model

Palletized load system (PLS) model developed using AGS PLS 3D Model and Off World Industries (OWI) N-Wheeled plugin

Sensor Models

Plug-in development and conversion for LiDAR, radar, ultra-wide band (UWB)



DMS Test Services

System Components

Simulation and Visualization

Real time digital modeling and simulation visualization tool and physics engine

Perception Modeling

High fidelity simulator to characterize perception system performance in operational settings and environments

Unified Digital Data

Enhanced data management and analysis tools and frameworks

Intelligent Test Selection

Automated test case planning via optimization algorithms (OA) and clustering to rapidly generate salient test scenarios

Fault Injection and Analysis

Software test tool that learns conditions for activating vulnerabilities deep within an autonomous system

Data Analytics and Reporting

System software diagnosis tools and algorithms (e.g., metrics to identify anomalous or undesired states of behavior)

Simulation Core

Real-time Simulation and Visualization Engine

Real time digital modeling and simulation visualization tools and physics engine

- Interface software under test to simulated environments or diagnosis algorithms
- Connecting glue between System Under Test (SUT) and simulation framework
- Continuous improvements to vehicle dynamics and sensor performance
- Open API / Free Source Code

https://www.unrealengine.com/en-US/spotlights/offworld-industries-brings-realistic-infantry-training-to-the-simulation-community

Synthetic Test Environment. Synthetic inputs are precisely controlled and known; perception stimuli to assess autonomous software logic.

- Easy, expedient, and precise variability of many test factors to stimulate sensor inputs to decision engine
- Develop an immersive environment for complex terrain
- Pursue issue and exact initial conditions of repeated tests

UNREAL ENGINE

Unreal Engine is an **open**, user-centric, game engine.

Backbone of our immersive simulation architecture

Model V&V

Structured Test Design for Continuous Feedback

Verification and Validation

V&V of the models and simulations are required to ensure the results are accurate and provide confidence to the test community

- Real-world testing will establish a feedback loop to validate the simulated events
- Design and integrate data inputs and formats to synchronize virtual and live test environments
- Optimize use of HWIL and live testing to validate the perception components and virtual test results rather than relying on physical components to explore the test space[1]

	Autonomous Convoy Operations	Test Area	Function	Test Case	Configuration Variables
1	Vehicles marshal in depot and prepare for road march	ATEF Bomb Ramp	Basic turns, stop/start	Turning Radius	Vehicle Speed 65 km/hr. (paved) & 40 km/hr. (gravel) 80 kilometer (50 miles) road march Leader-Follower Mode Time of day – varying (sunrise, mid-morning, noon, mid-afternoon, evening, sunset, and night) Obstacles – Vehicle/Pedestrian Sun/Partly Cloudy/Fog/Night Sensor Degradation: - Reduced Visibility - Sensor Outage
2	Convoy on-the-move and departs depot	ATEF Bomb Ramp – Entry control point install	Maneuver through an entry control point avoiding obstacles and following leader waypoints	Entry Control Point	
3	Convoy road march on paved highway traveling 40 kilometers at 65 km/hr	ATEF Paved	Maneuver by accelerating, braking and changing lanes	Emergency Lane Change	
4	Convoy road march on paved road with obstacles	ATEF Paved – Brake Test area	Maneuver by steering, braking, changing lanes	Emergency Lane Change; Obstacle Detection and Avoidance (ODA)	
5	Convoy road march on paved road with communications loss or blowing dust across road	ATEF – Smith Turn area	Maneuver by steering, braking, changing lanes	UGV Communications Test	
6	Convoy road march on gravel road traveling 40 kilometers at 40 km/hr	ATEF - Gravel	Maneuver by accelerating, braking and changing lanes	Emergency Lane Change	
7	Convoy road march on gravel road with obstacles	ATEF Gravel – Brake Test area	Maneuver by steering, braking, changing lanes	Emergency Lane Change and ODA Test	
8	Convoy road march on gravel road with communications loss or blowing dust across road	ATEF – Smith Turn area	Maneuver by steering, braking, changing lanes	UGV Communications Test	
9	Convoy arrives at forward operating base	ATEF Bomb Ramp – Entry control point install	Maneuver through an entry control point avoiding obstacles and following leader waypoints	Entry Control Point	

Fault Injection

Software Assurance

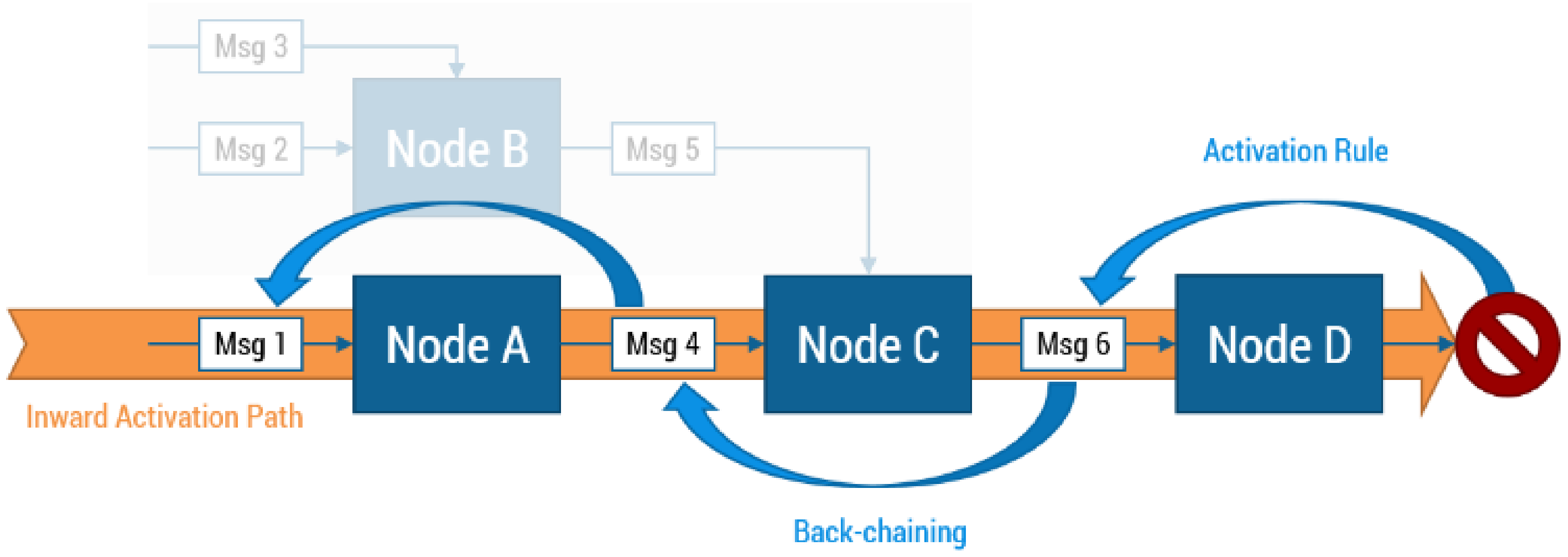
Conduct robustness testing of the autonomy software to find unsafe conditions that may map to software bugs, requirements deficiencies, or design flaws.

- Efficiently detect vulnerabilities in autonomy
- Understand the safety implications and risks posed by the vulnerabilities
- Design tests to exercise and software resiliency

Edge Case Research

Robust Inside-Out Testing

Enables software stress testing and fault injection



Test Design

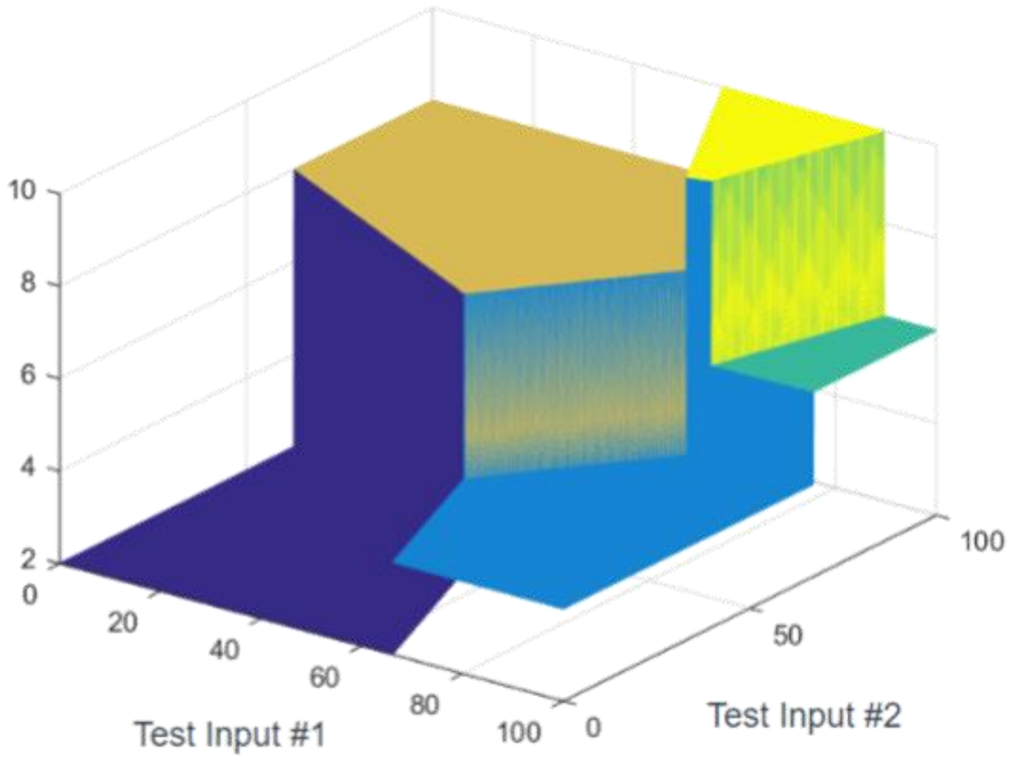
Preparing the right question

TEST PARAMETERS | VALIDATED MODELS | TEST CASES

https://hitchhikers.fandom.com/wiki/Great_and_Hopefully_Enlightening_Day_(Day_of_the_Answer)

Test Selection Methodology

Edge Case Discovery



Range Adversarial Planning Tool (RAPT)

Enables edge case discovery to down select scenarios

JHU Applied Physics Laboratory

Test case analysis to systematically identify edge cases and sequences of input conditions that can trigger emergent or unanticipated behavior over numerous iterations of a given test scenario

Automated test case planning via optimization algorithms (OA) and clustering to rapidly generate salient test scenarios[1]

- Automated test plan development
- Identifies performance boundaries
- Identifies needed re-testing when mission, hardware or software changes occur

Analyze the potential for emergent behavior in order to avoid it will be central to providing assured dependability for autonomous systems.

- Identify criticality-ranked, performance-stressing scenarios for an autonomy drawn from the environment, mission, and vehicle state spaces
- Identify discrete pass/fail boundaries for autonomy performance

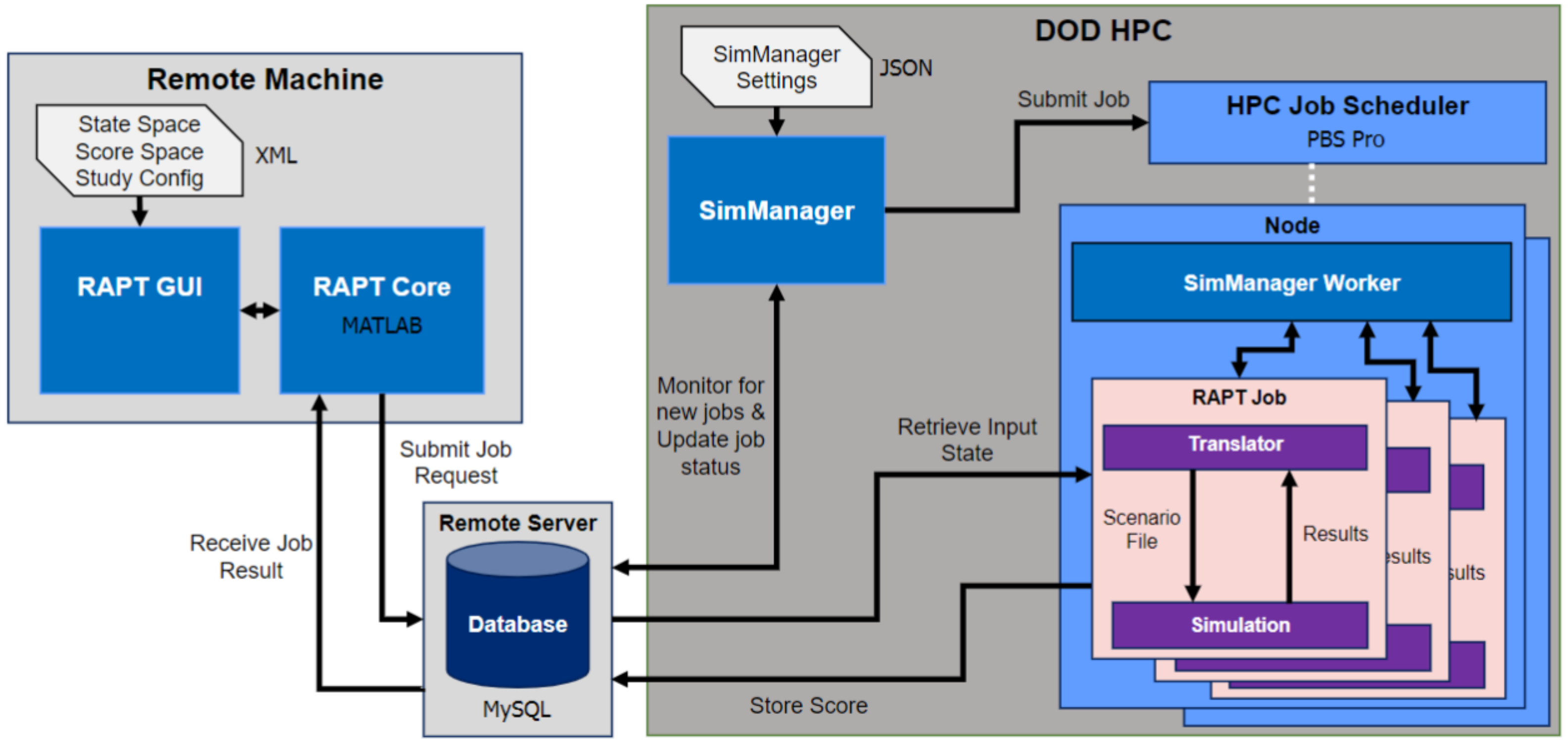
1. Delivering Test and Evaluation Tools for Autonomous Unmanned Vehicles to the Fleet, *Johns Hopkins APL Technical Digest*, Volume 33, Number 4 (2017), www.jhuapl.edu/techdigest

Mass Simulation

High Throughput through HPC Resources

Analysis achieved by automating execution of simulations with the appropriate level of fidelity based on configured combinations of inputs

- High volume of test simulations conducted to estimate probabilities of mission and safety failures
- Template simulations and variables are loaded and executed in parallel to reduce test time and using a batch mode to reduce operator involvement



Horris, J & Stankiewicz, P. Range Adversarial Planning Tool (RAPT) Briefing, Mar 2020.

Time-managed algorithmic software will accelerate this process by implementing a faster than real-time framework

High performance computing (HPC) and cloud-enabled infrastructure are key in accomplishing this simulation framework

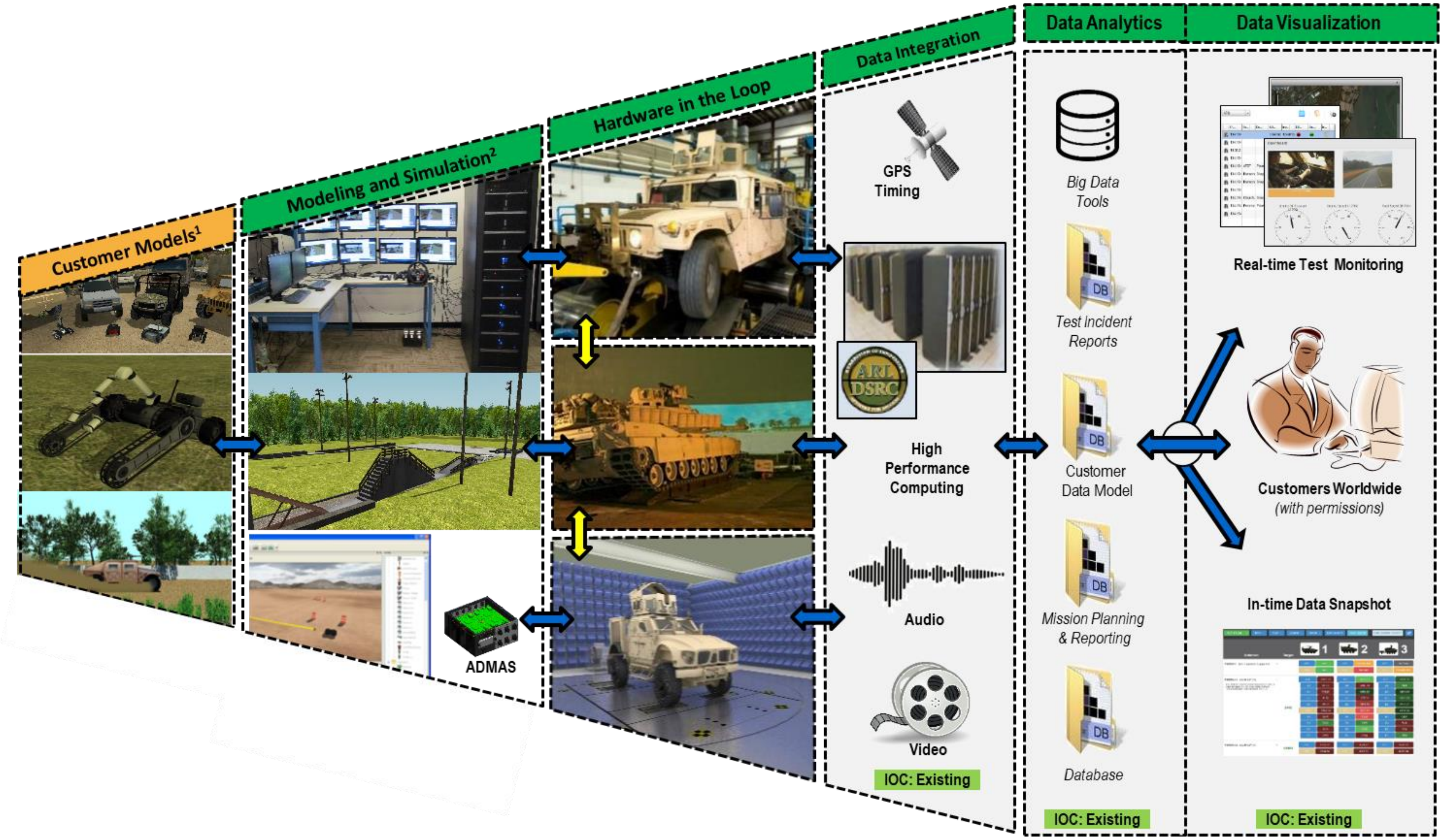
U.S. Army Research Laboratory
DoD Supercomputing Resource Center

Excalibur is a Cray XC40 system located at the ARL DSRC. It has 3,098 standard compute nodes, 32 large-memory compute nodes, and 32 GPU compute nodes (a total of 3,162 compute nodes or 101,184 compute cores). It has 421 TBytes of memory and is rated at 3.77 peak PFLOPS

https://centers.hpc.mil/systems/unclassified.html#Excalibur

Unified Digital Data

Data Acquisition and Analysis



Implementation of data acquisition tools will ensure data from all sources are read and interpreted consistently

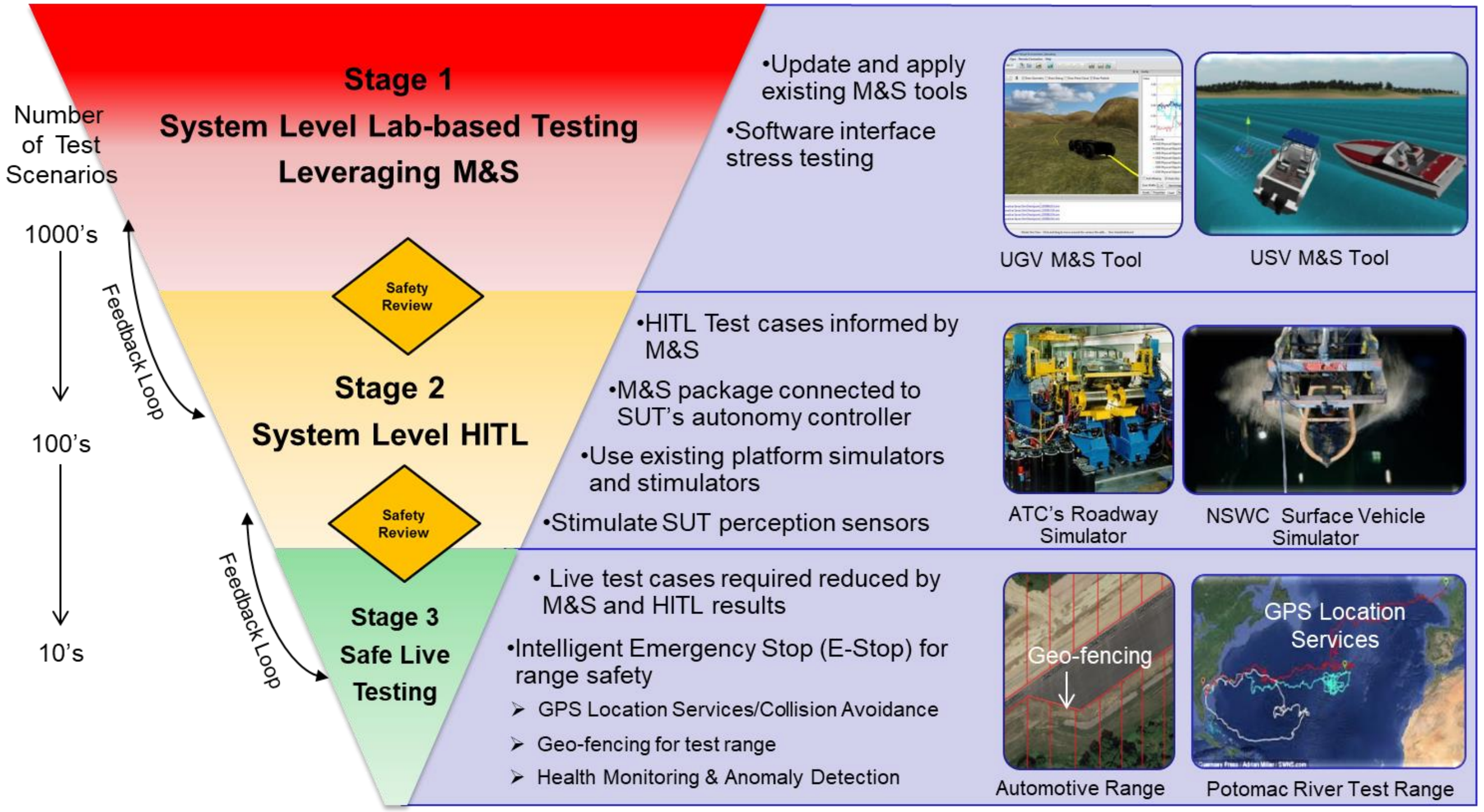
Simulation Configuration
Configuration of virtual environment and autonomy software interfaces
- XML Files
- Dynamic Link Libraries (DLL)

Scenario Generation
Develop mission, tests to be conducted, and faults to determine autonomy is functioning correctly
- Mission
- Test Cases
- Faults

Results
Collect results and develop reports and visualizations to assist in post test analysis
- Mission Failures
- Safety Failures

DMS in Action

Autonomous System Test Capability (ASTC)



ASTC **DRIVE**

<u>D</u>igital <u>R</u>obotic and Autonomous Systems (RAS) <u>I</u>ntegrated <u>V</u>irtual <u>E</u>nvironment

DRIVE represents digital testing ecosystem to enable testing autonomous systems.

U.S. Army Aberdeen Test Center

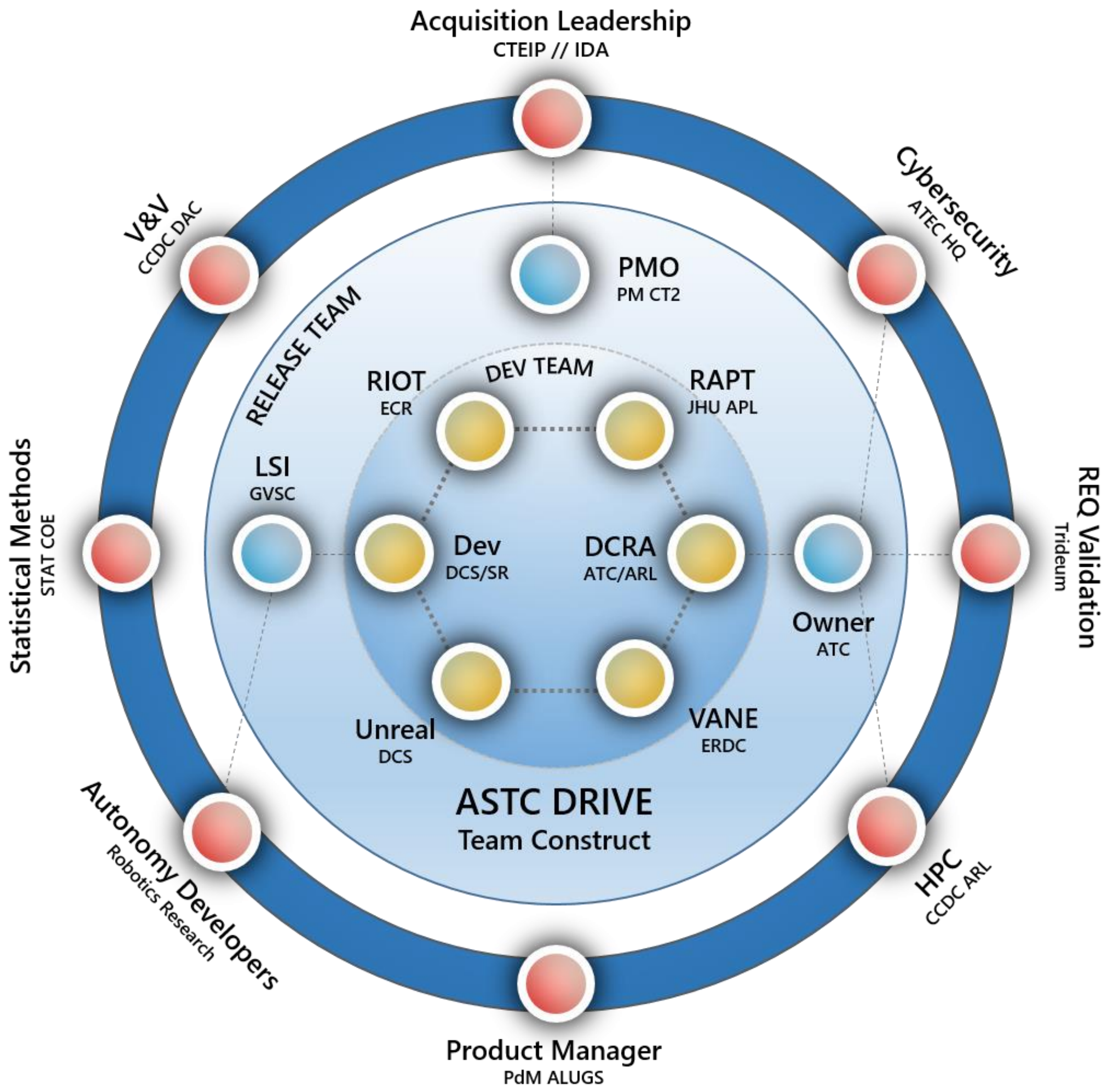
For more details on the program, contact: Mr. John Whitt, john.m.whitt4.civ@mail.mil

Deploying Expertise

Agile Teams – Optimize Collaboration

Agile Management

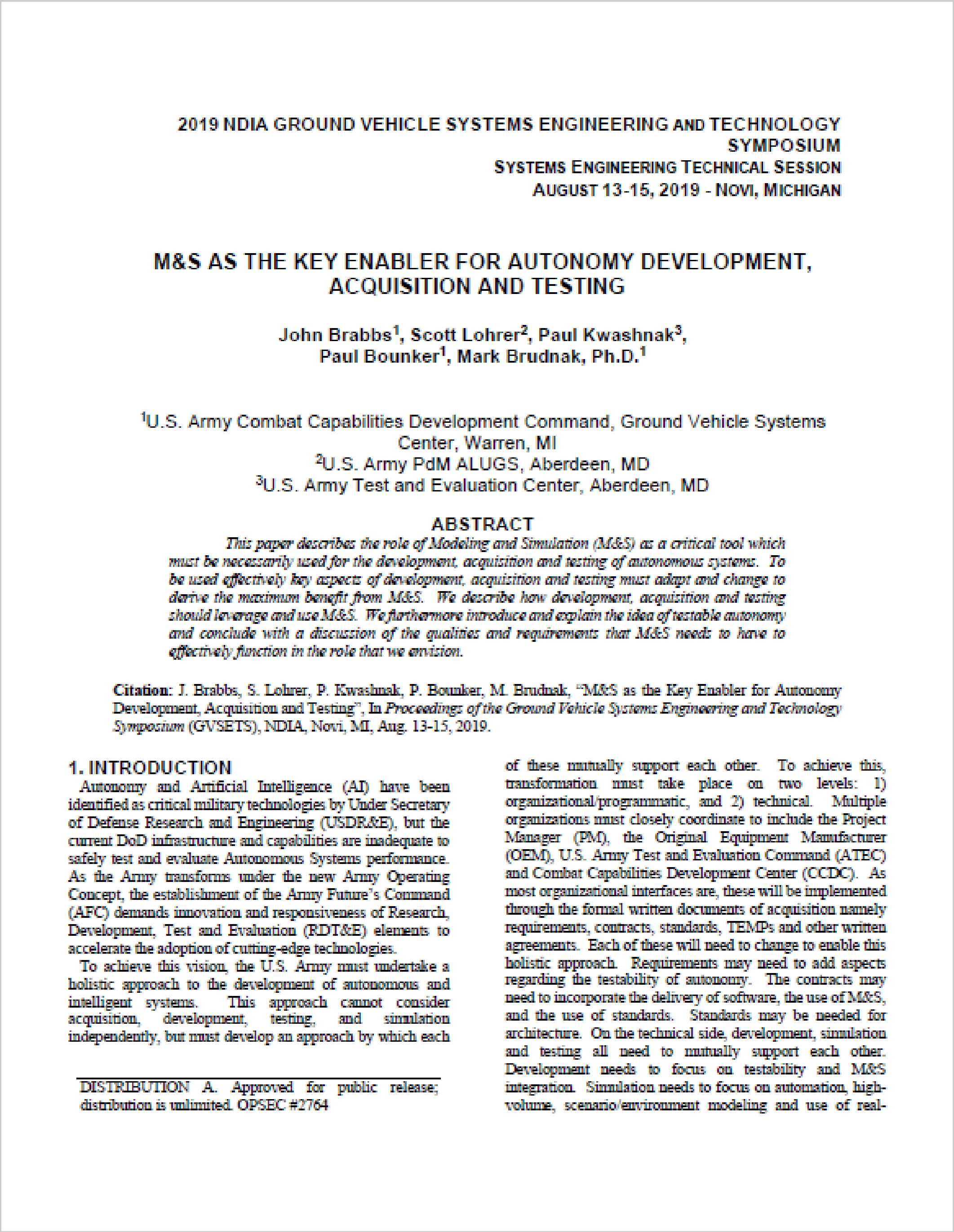
- Incremental development for software enhancements and integration using iterative releases
  - Focuses on continuous capability releases
  - Releases provide internal milestones and metrics to meet performance objectives per technical specification
  - Approach delivers capability throughout life cycle; improves responsiveness to changes in operations, technology, and budget
  - Actively involve users throughout development to ensure operational value; embed user training and promote RDT&E collaboration
  - Practice currently executed by proposed integrators



Further Reading

M&S As The Key Enabler For Autonomy Development, Acquisition And Testing

J. Brabbs, S. Lohrer, P. Kwashnak, P. Bounker, M. Brudnak, “M&S as the Key Enabler for Autonomy Development, Acquisition and Testing”, In Proceedings of the Ground Vehicle Systems Engineering and Technology Symposium (GVSETS), NDIA, Novi, MI, Aug. 13-15, 2019



2019 NDIA GROUND VEHICLE SYSTEMS ENGINEERING AND TECHNOLOGY SYMPOSIUM
SYSTEMS ENGINEERING TECHNICAL SESSION
AUGUST 13-15, 2019 - NOVI, MICHIGAN

M&S AS THE KEY ENABLER FOR AUTONOMY DEVELOPMENT, ACQUISITION AND TESTING

John Brabbs[1], Scott Lohrer[2], Paul Kwashnak[3], Paul Bounker[1], Mark Brudnak, Ph.D.[1]

[1]U.S. Army Combat Capabilities Development Command, Ground Vehicle Systems Center, Warren, MI
[2]U.S. Army PdM ALUGS, Aberdeen, MD
[3]U.S. Army Test and Evaluation Center, Aberdeen, MD

ABSTRACT

This paper describes the role of Modeling and Simulation (M&S) as a critical tool which must be necessarily used for the development, acquisition and testing of autonomous systems. To be used effectively key aspects of development, acquisition and testing must adapt and change to derive the maximum benefit from M&S. We describe how development, acquisition and testing should leverage and use M&S. We furthermore introduce and explain the idea of testable autonomy and conclude with a discussion of the qualities and requirements that M&S needs to have to effectively function in the role that we envision.

Citation: J. Brabbs, S. Lohrer, P. Kwashnak, P. Bounker, M. Brudnak, “M&S as the Key Enabler for Autonomy Development, Acquisition and Testing”, In *Proceedings of the Ground Vehicle Systems Engineering and Technology Symposium* (GVSETS), NDIA, Novi, MI, Aug. 13-15, 2019.

1. INTRODUCTION

Autonomy and Artificial Intelligence (AI) have been identified as critical military technologies by Under Secretary of Defense Research and Engineering (USDR&E), but the current DoD infrastructure and capabilities are inadequate to safely test and evaluate Autonomous Systems performance. As the Army transforms under the new Army Operating Concept, the establishment of the Army Future’s Command (AFC) demands innovation and responsiveness of Research, Development, Test and Evaluation (RDT&E) elements to accelerate the adoption of cutting-edge technologies.

To achieve this vision, the U.S. Army must undertake a holistic approach to the development of autonomous and intelligent systems. This approach cannot consider acquisition, development, testing, and simulation independently, but must develop an approach by which each of these mutually support each other. To achieve this, transformation must take place on two levels: 1) organizational/programmatic, and 2) technical. Multiple organizations must closely coordinate to include the Project Manager (PM), the Original Equipment Manufacturer (OEM), U.S. Army Test and Evaluation Command (ATEC) and Combat Capabilities Development Center (CCDC). As most organizational interfaces are, these will be implemented through the formal written documents of acquisition namely requirements, contracts, standards, TEMPs and other written agreements. Each of these will need to change to enable this holistic approach. Requirements may need to add aspects regarding the testability of autonomy. The contracts may need to incorporate the delivery of software, the use of M&S, and the use of standards. Standards may be needed for architecture. On the technical side, development, simulation and testing all need to mutually support each other. Development needs to focus on testability and M&S integration. Simulation needs to focus on automation, high-volume, scenario/environment modeling and use of real-

DISTRIBUTION A. Approved for public release; distribution is unlimited. OPSEC #2764

Truth in Testing!

Paul J. Kwashnak
U.S. Army Test and Evaluation Command
Capabilities Division, Strategic Initiatives
paul.j.kwashnak.civ@mail.mil