International Council on Systems Engineering
A better world through a systems approach

Digital Engineering Testbed for Test and Evaluation: Operation Safe Passage Status and Lessons Learned

Short Title: ***DE for T&E Testbed***

Presenter: Dr. Paul Wach, PhD

First author: Brandt Sandman, PhD student

Co-authors:

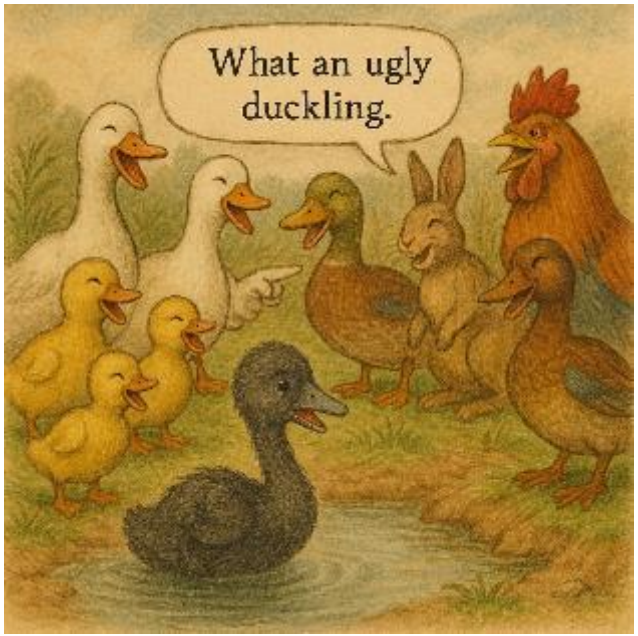
Drs. Alejandro Salado and Joe Gregory, University of Arizona

Dr. Taylan Topcu, Mr. Geoff Kerr, Virginia Tech

INCOSE International Symposium 2025 | Ottawa, Canada

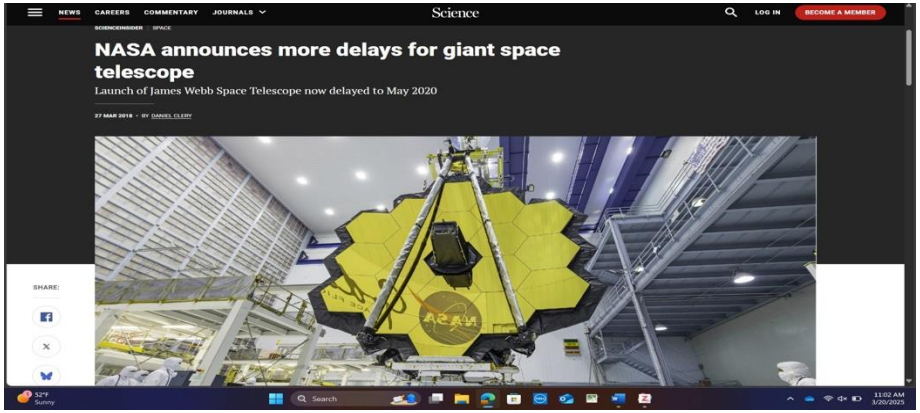
…before we get into the technical details…

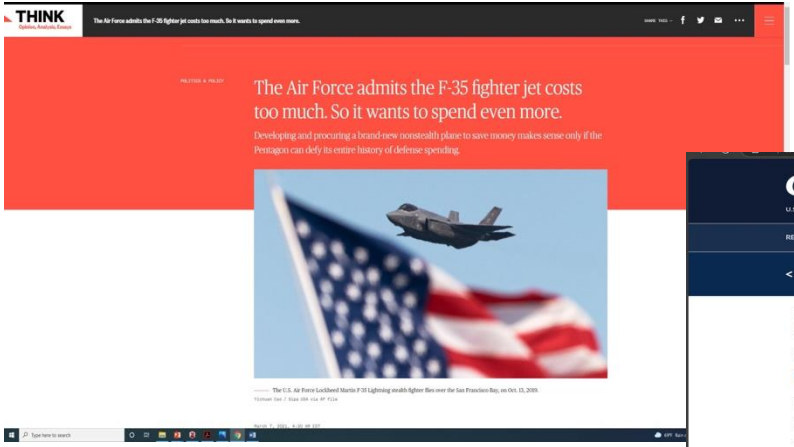
Reactions to digital engineering…



Are we there yet?

- Difficult to use real programs/systems
- 10+ years, classification, etc.
- Changes in technology over time and struggle to stay adapted



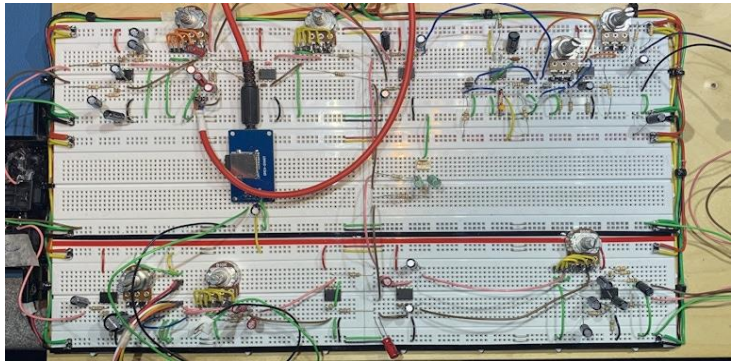


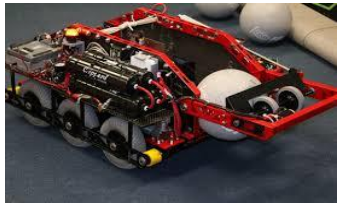
What do we mean when we say “test and evaluation” (T&E)?

- Government owned and executed verification and validation (V&V) activities
  - Research team perspective (may not be sponsor’s perspective)
- From a US DOD perspective:
  - Read DoDI.5000.89
  - Developmental T&E
    - Conducted while the systems is under development
    - Department of Developmental Test, Evaluation, and Analysis (DTE&A)
  - Operational T&E
    - Conducted before and while the system is in operation
    - Department of Operational Test and Evaluation (DOT&E)

Use of proxies/analogies is regular practice

- Modeling & Simulation (M&S)
- Breadboards
- Prototypes
- Wind tunnels (reduced scale)
- Test ranges (reduced scale)

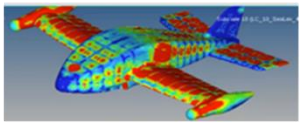




Recognized need to start new

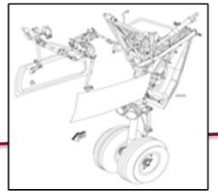
Desire deliberate elegance rather than organic, momentary success

Deep Dives by Phases

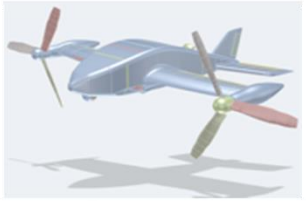


P1: Multi-physics

Graphical CONOPS Scenario: Search & Rescue



P2: Airworthiness



P3: Cost Modeling

Performance constraints force Multi-physics Design considerations – similar to Bell Eagle Eye

*Skyzer **Doing Everything in Models to Demonstrate Art-of-the-Possible**

Meet the team

Part of the **Systems Engineering Research Center** (SERC) and Acquisition Innovation Research Center (AIRC)

THE UNIVERSITY OF ARIZONA

VIRGINIA TECH

Dr. Alejandro Salado, UA

Dr. Joe Gregory, UA

Dr. Paul Wach, VT

Geoffrey Kerr, VT

Brandt Sandman, VT

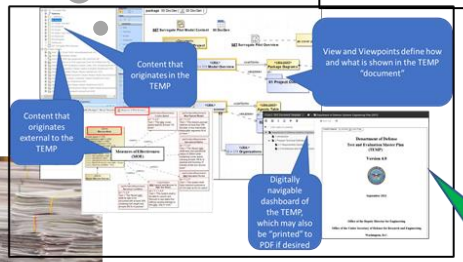
Taylan Topcu, VT

Phases of T&E Transformation

Current measure of goodness?

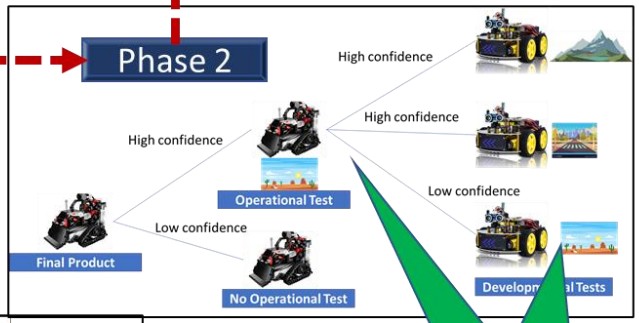
Phase 0

Phase 1



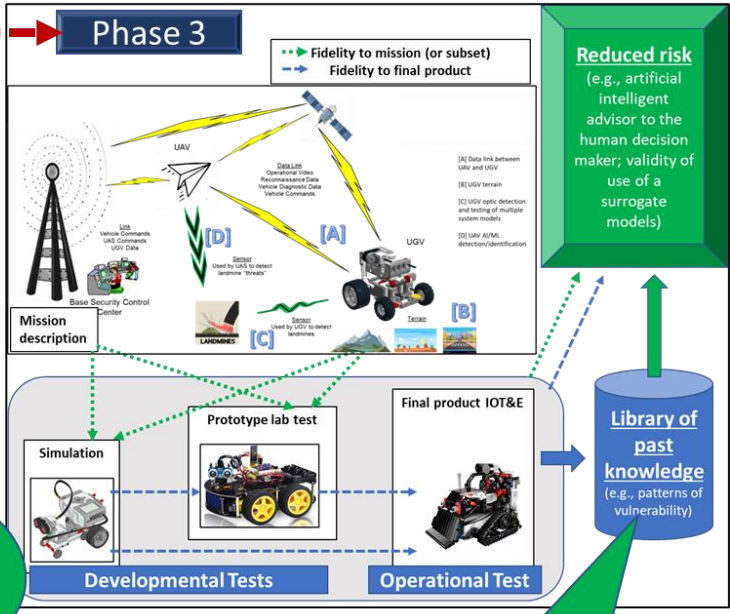
Begins conversion to a dynamic data continuum

Phase 2



Decision dependence (e.g., DT>>OT) and metrics

Phase 3



Reduced risk (e.g., artificial intelligent advisor to the human decision maker; validity of use of a surrogate models)

Library of past knowledge (e.g., patterns of vulnerability)

Interoperability, reuse, artificial intelligence and machine learning (AI/ML), known fidelity from inception to retirement

Historical Fiction?!

Mission proxy

- Operation Safe Passage

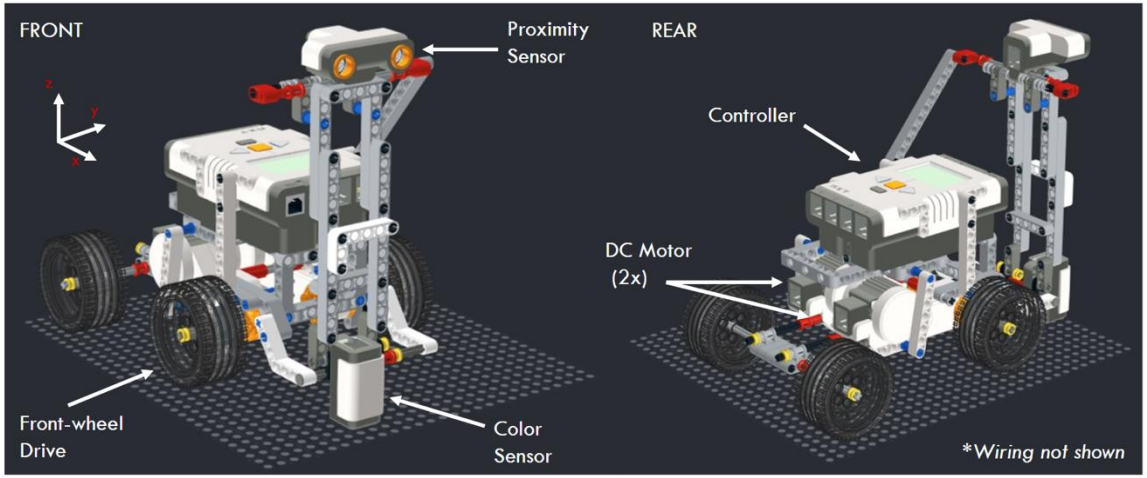
- Operation Safe Lego

Historical Fiction?!

System proxy

- Operation Safe Passage

- Operation Safe Lego

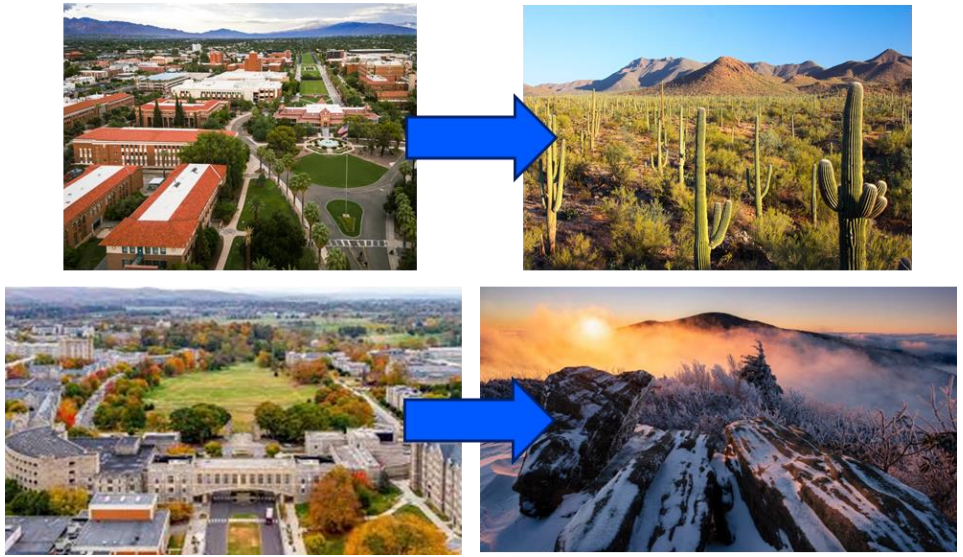


Historical Fiction?!

Terrain proxy

- Operation Safe Passage

- Operation Safe Lego



Historical Fiction?!

Threat proxy

- Operation Safe Passage

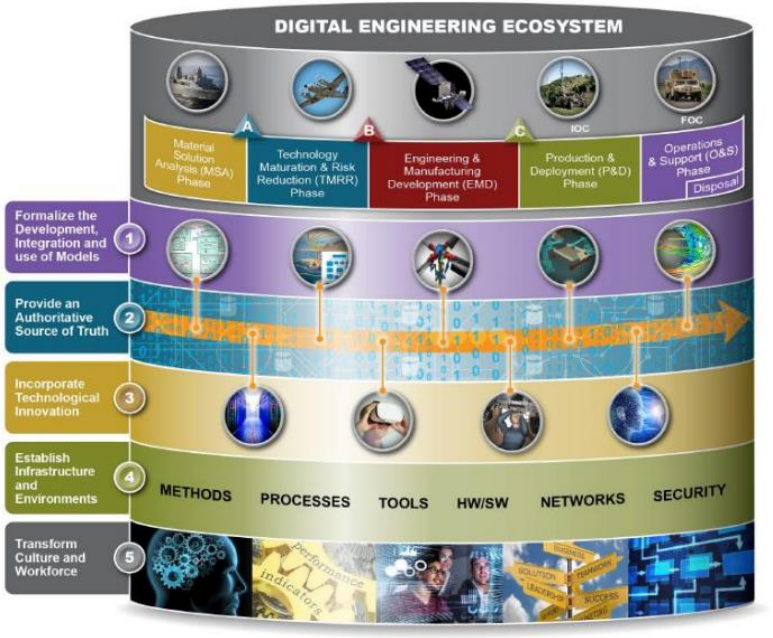
- Operation Safe Lego

Historical Fiction?!

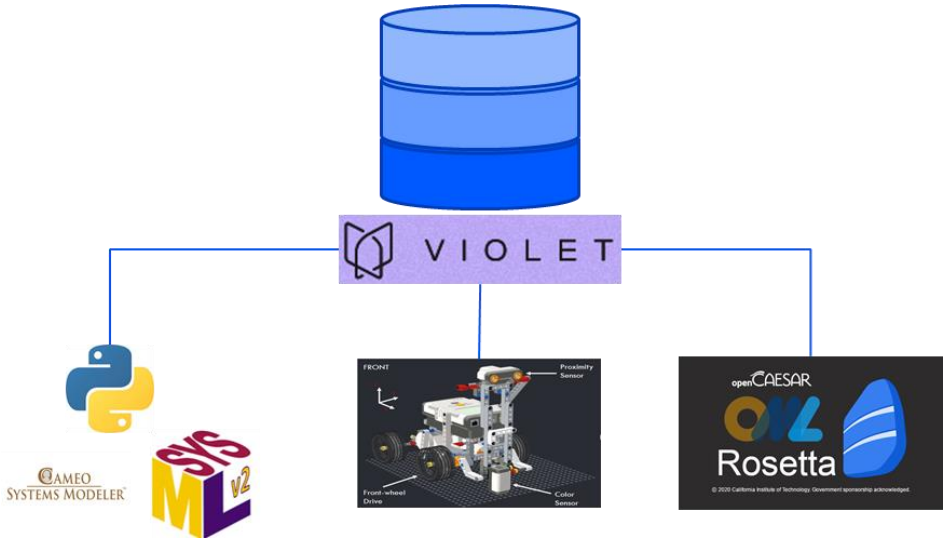
https://www.violetlabs.com/

System proxy

- Operation Safe Passage



- Operation Safe Lego



Key Performance Parameters (KPPs)

Measures	Definition
Detection Rate*	% mines detected correctly
False Alarm Rate, Detection*	% mine detection false alarms
Classification Rate*	% mines classified correctly
False Alarm Rate, Classification*	% mine classification false alarms
Pathway Creation Time*	Time taken to clear a pathway from start to finish
Safety Rating	Adherence to safety standards and protocols
Adaptability Rating	Ability to operate in various terrains and conditions

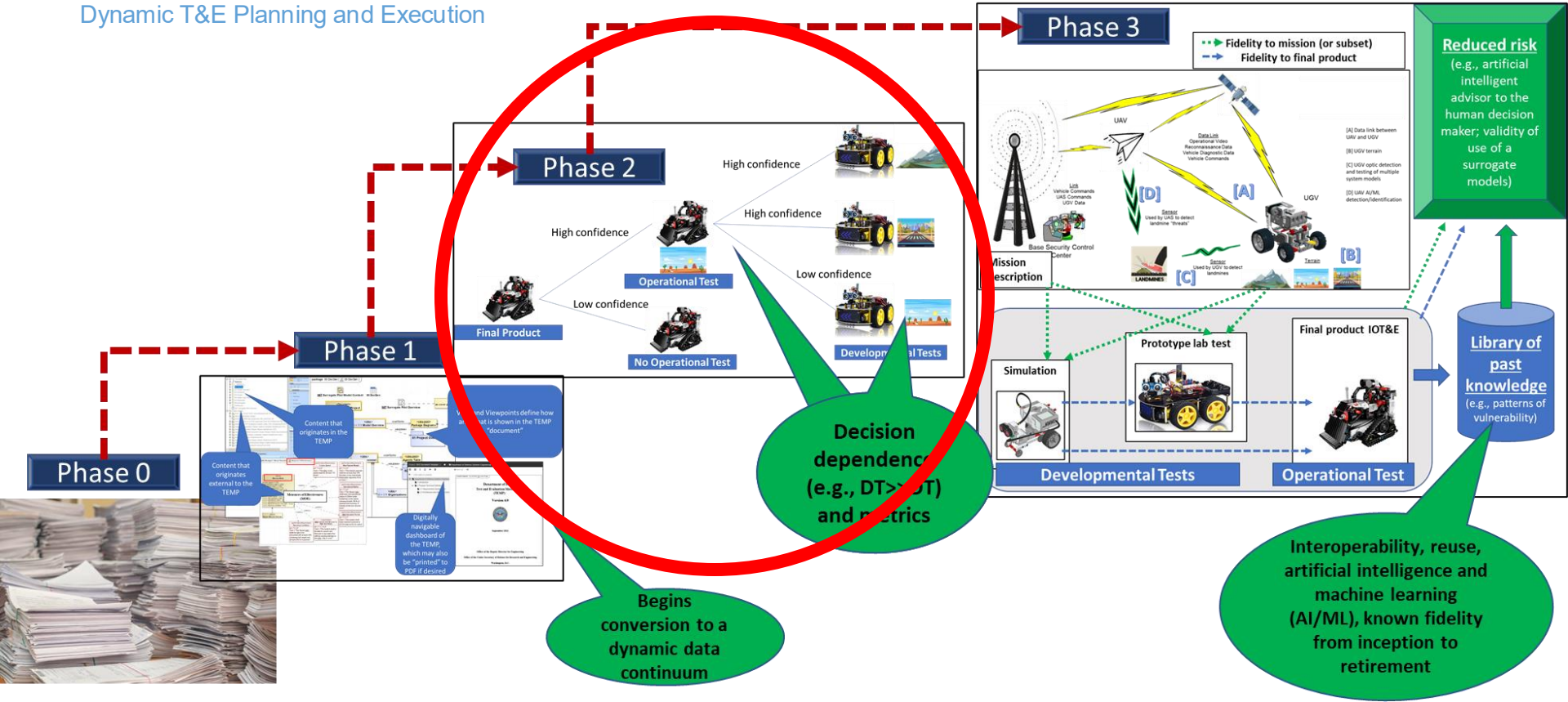
*"mine" ~= "sticky note"

Mock interim design review" (IDR)

Category	Description
Entities involved and role	1. DOT&E personnel with a role as the T&E decision authority 2. AIRC researchers with a role to conduct the T&E activities and provide artifacts to the decision authority
Supporting artifacts	1. PowerPoint slides as a read ahead 2. Lego mission model 3. Lego (SysML) system model 4. Lego CAD models 5. MBTEMP 6. Modeling and Simulation (M&S) models (Python-based) 7. Analytical results 8. IDSK (decision dashboard)
Means of conducting the review	1. Virtual (MS Teams) 2. Use of PowerPoint to guide the review 3. Use of real-time presentation of models and data (e.g., MBTEMP and IDSK)
T&E use cases	1. Developmental tests resulting in design change after T&E showed the inability to achieve traction at lower weights 2. Developmental tests resulting in design and performance trade after T&E showed the tire diameter and terrain are co-dependent 3. Operational tests resulting in change in employment due to inability to employ to UGV and causing the need to utilize a UAV as an alternative
Review phases	1. Initial IDR (June 2024) 2. Delta IDR (Aug 2024)

Phase of T&E Transformation

Dynamic T&E Planning and Execution



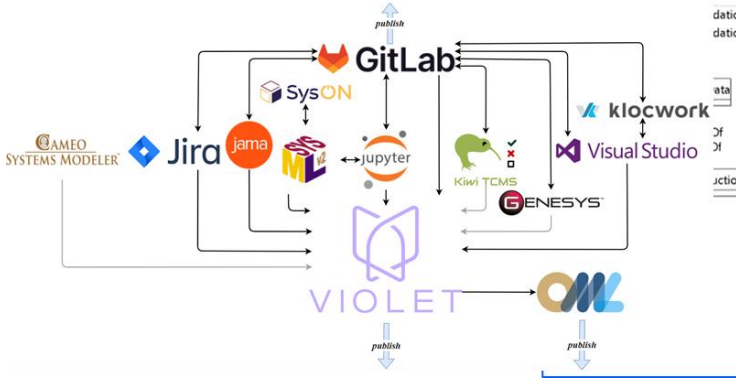
Digital Test and Evaluation Master Plan (dTEMP)

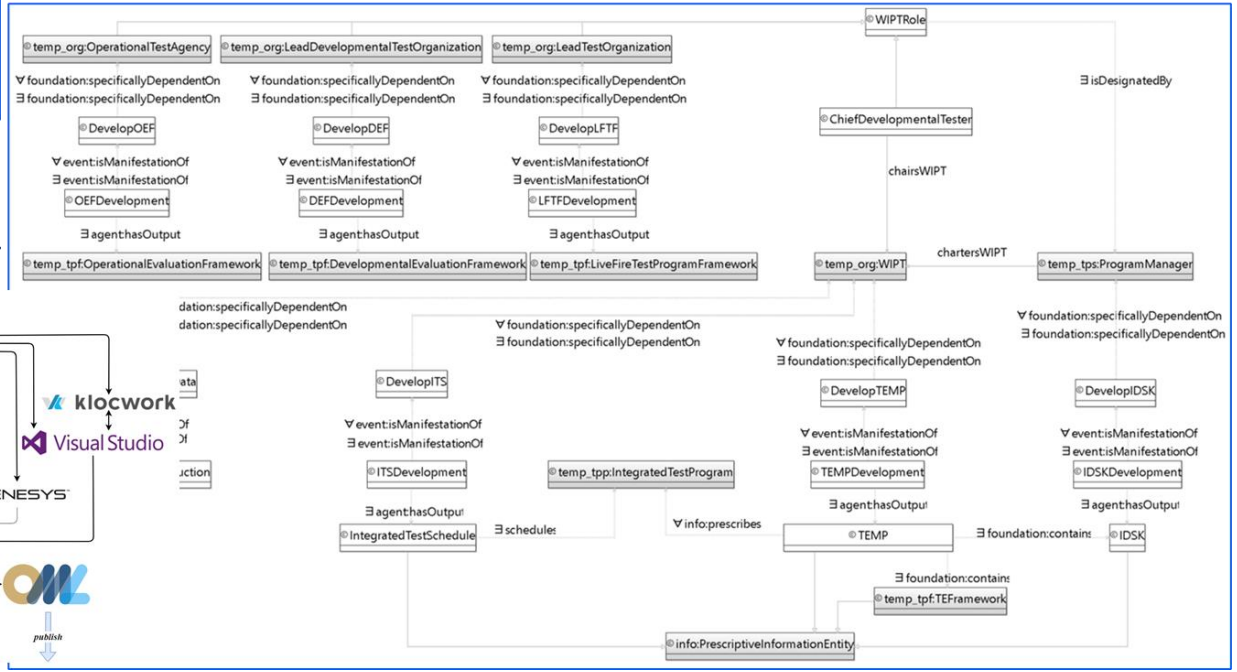
DoD Instruction 5000.89

Test and Evaluation

Originating Component:	Office of the Under Secretary of Defense for Research and Engineering
Originating Component:	Office of the Director, Operational Test and Evaluation
Effective:	November 19, 2020

Mapping the DoDI Into Foundational Tooling





Digital Test and Evaluation Master Plan (dTEMP)

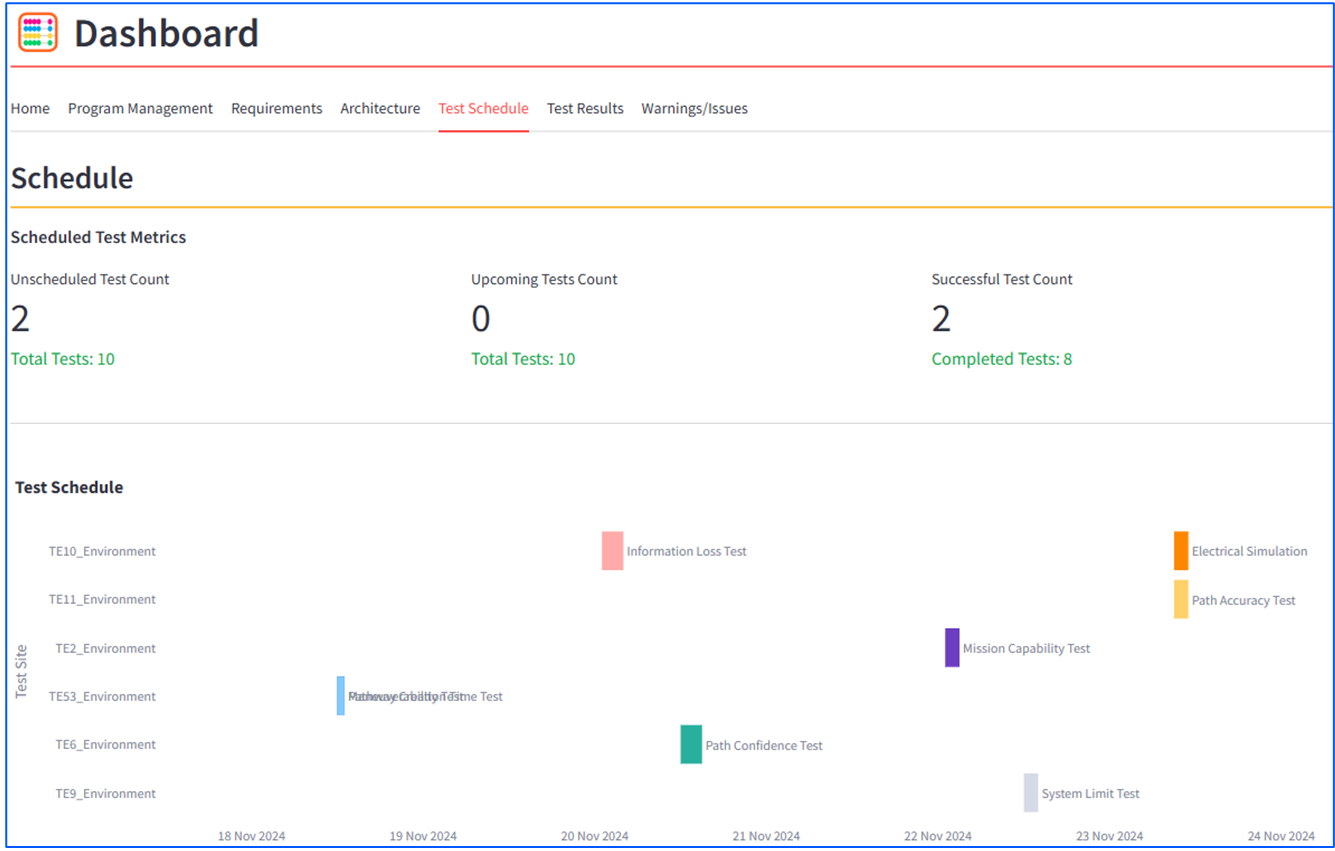
⚠ Four tests have overlapped scheduling (find more info on Issues tab)

❗ Electrical Test has potential schedule conflict with other tests

❗ Path Accuracy Test has potential schedule conflict with other tests

❗ Maneuverability Test has potential schedule conflict with other tests

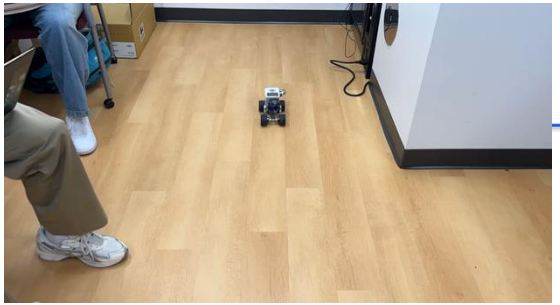
❗ Pathway Creation Time Test has potential schedule conflict with other tests



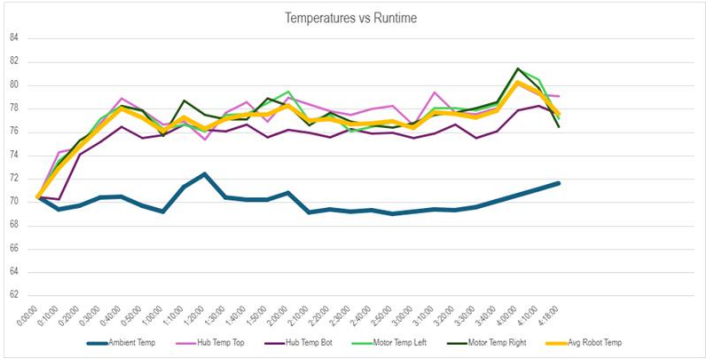
A job fit for students: Battery testing

Spike Robot

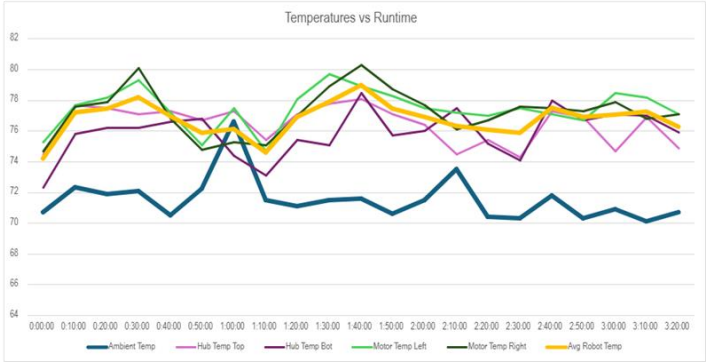
EV3 Robot



EV3
Test 1

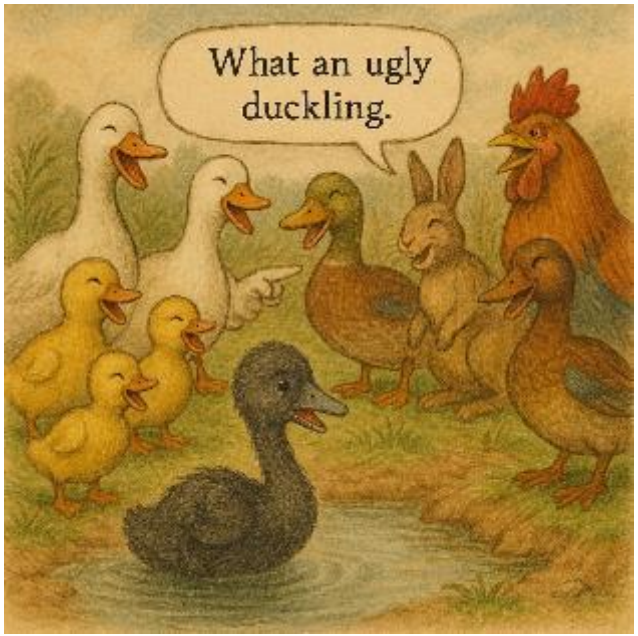


EV3
Test 2



Reactions to our work???

We hope not



Questions?

Paul Wach, paulw86@vt.edu
Alejandro Salado, alejandrosalado@arizona.edu
Joe Gregory, joegregory@arizona.edu

Current trajectory:

- ✓ Continued injection into classroom
- ✓ Scaling and expanding
- ✓ Phase 3 (mathematically underpinned models/data)
- ✓ Expanding to other systems
- ✓ Expanding to system of systems
- ✓ Expanding to mission/capability level

Backup

What do I really need to test?

Developmental Tests

High confidence

High confidence

Low confidence

Operational Test

High confidence

Final Product

No Operational Test

Low confidence

A job fit for students

Brandt Sandman

Trevor Ierardi

Robert Sharpe III

Walter Lin

Allegra Oledibe

Liam Superville

Austin Lane

Emmanuel Oteng

Winston Mensah

Test procedures

Test execution

Actions	Measurements
Measure the room temperature using the thermometer.	
Simultaneously initiate Rover movement and activate stopwatch.	
Check the Rover's wheels are moving by inspection.	
Measure the rpm of the Rover's wheels.	
Continue observation of wheel movement.	
Stop the stopwatch once the Rover's wheels stop moving.	
Record the time indicated on the stopwatch.	

Test preparations

Actions
Place Rover on the Workbench.
Load SD Card 1 into the Rover Computer Module.
Place Rover in the Disruption Lab.
Plug the Rover Battery Charger into the Rover.
Wait until the battery indicator indicates the battery is fully charged.
Unplug the Rover Battery Charger from the Rover.
Place the Rover in the Hallway with no surface in contact with the Rover's wheels.
Reset the stopwatch to 0.

Check Preconditions

Name	Calibration Status
Old Engineering University of Arizona	Valid
Thermometer	Valid
Stopwatch	Valid
Rover Battery Charger	Valid
Disruption Lab	Valid
Hallway	Valid
Workbench	Invalid
Rover Computer Module	Invalid

Notional Case: Legacy Data/Systems Used to Inform T&E for New System