



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

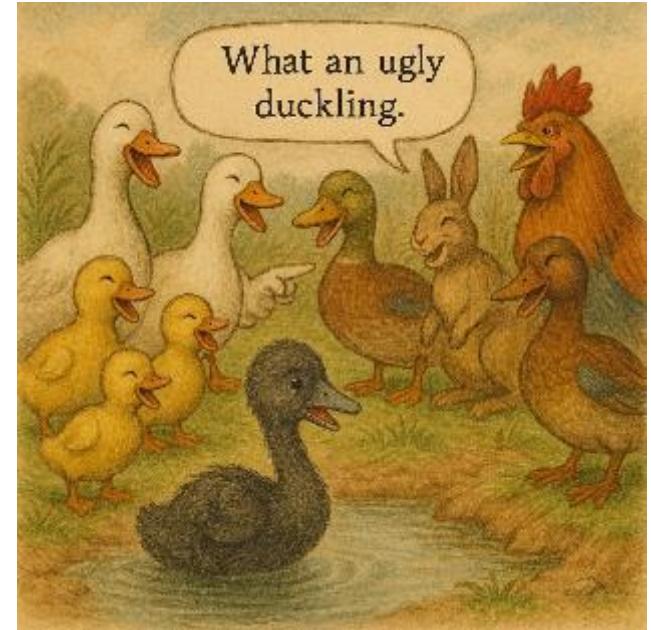
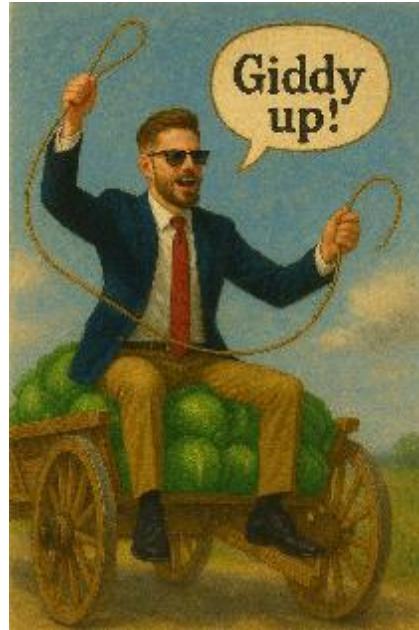
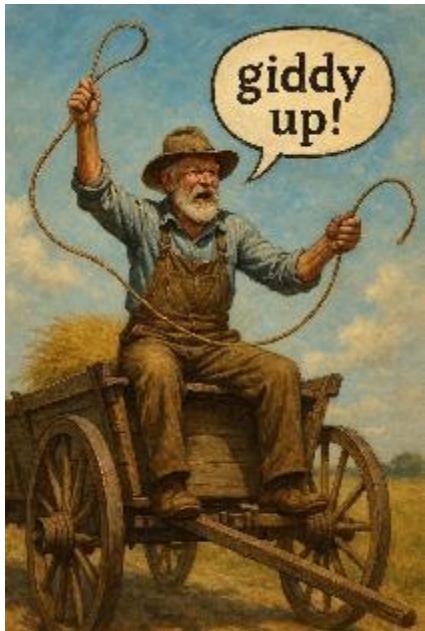
Dr. Taylan Topcu, Mr. Geoff Kerr, Virginia Tech

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

**THINK** Systems Thinking

The Air Force admits the F-35 fighter jet costs too much. So it wants to spend even more.

Developing and procuring a brand new nonstealth plane to save money makes sense only if the Pentagon can defy its entire history of defense spending.

The Air Force admits the F-35 fighter jet costs too much. So it wants to spend even more.

Developing and procuring a brand new nonstealth plane to save money makes sense only if the Pentagon can defy its entire history of defense spending.

The U.S. Air Force Lockheed Martin F-35 Lightning stealth fighter flies over the San Francisco Bay, on Oct. 12, 2008. (Human Eye / AP Photo via AP File)

March 1, 2024 4:46 AM EST

**GAO**  
U.S. Government Accountability Office

For Congress | Press Center | Careers | Blog

Search GAO.gov

REPORTS & TESTIMONIES | VIEW TOPICS | VIEW AGENCIES | BID PROTESTS & APPROPRIATIONS LAW | ABOUT

< WatchBlog: Following the Federal Dollar

**The F-35 Will Now Exceed \$2 Trillion As the Military Plans to Fly It Less**

Posted on May 16, 2024

[Facebook](#) [Twitter](#) [LinkedIn](#) [Email](#)

The F-35 Lightning II Joint Strike Fighter is the Department of Defense's most ambitious weapons program and plays a crucial role in DOD's ability to meet current and future U.S. national security goals. At the same time, the F-35 has been plagued by mounting costs and delays resulting in what some have called a staggering price tag of more than \$2 trillion over several decades.

In March, the fighter jet marked an important milestone – “full-rate production,” generally the point when development reaches an acceptable level of performance and reliability to start building more of them, faster. F-35s have already been in production at or near full rate for several years. At the same time, the military services that fly the F-35 (the Air Force, Marines, and Navy) plan to use it less.

What's going on with the F-35? Today's WatchBlog post looks at our two new reports about challenges with modernizing the F-35 and its rising costs and declining use.

**Related Posts**

**AO's Priority Open recommendations**

**Our 2024 Update**

Blog Post Priority Recommendation Round-Up—Actions that Can Bring the Best

NEWS CAREERS COMMENTARY JOURNALS

SCIENCE & SPACE

Science

SEARCH LOG IN BECOME A MEMBER

**NASA announces more delays for giant space telescope**

Launch of James Webb Space Telescope now delayed to May 2020

27 MAR 2018 BY DANIEL CLEARY

SHARE

Facebook Twitter LinkedIn Email

529 Sunny

Search

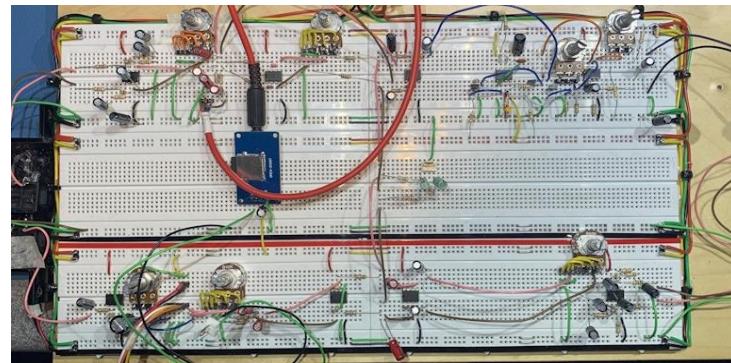
11:00 AM 3/20/2020

# 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 system 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)

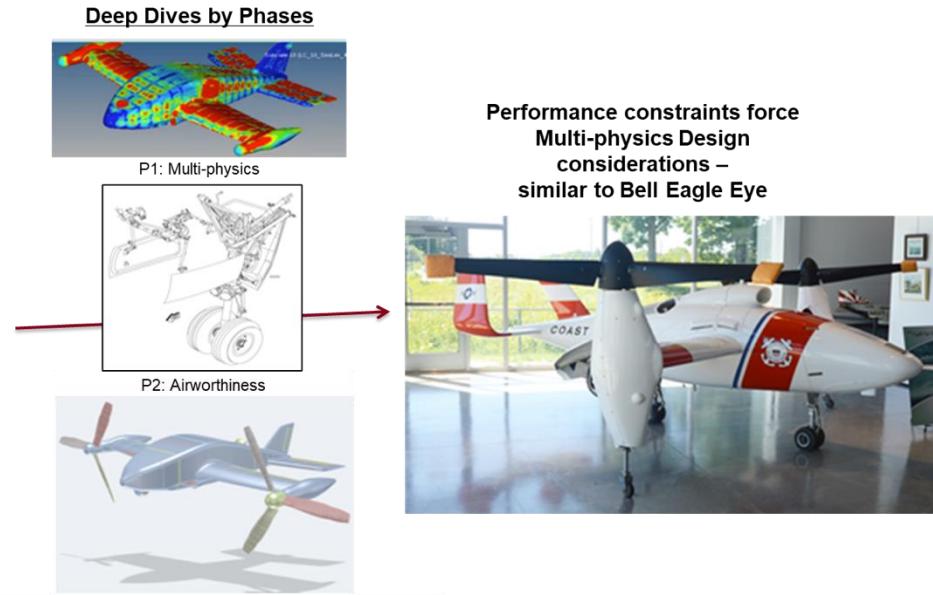
# Use of proxies/analogies is regular practice

- 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



**\*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)**



Dr. Alejandro  
Salado, UA

Dr. Joe  
Gregory, UA

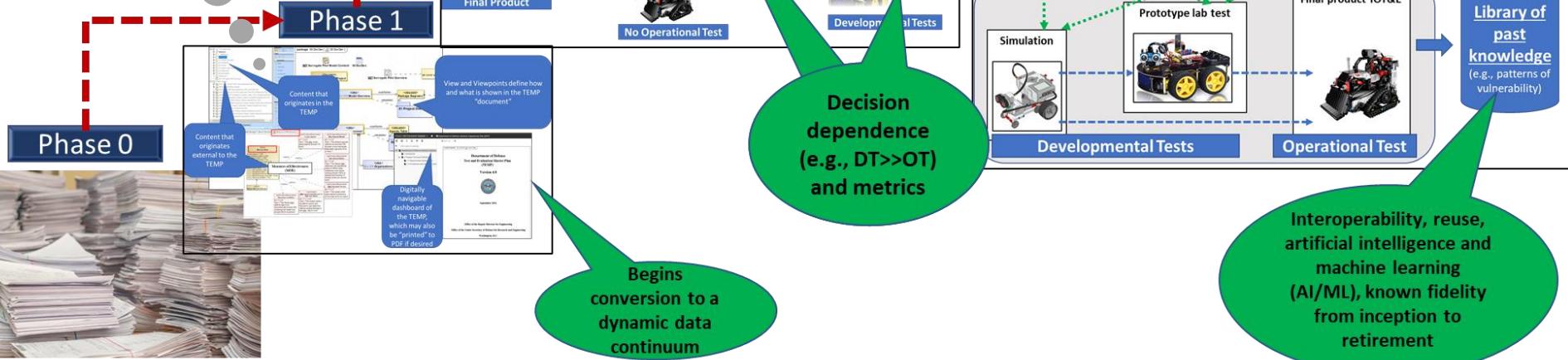
Dr. Paul Wach,  
VT

Geoffrey Kerr,  
VT

Brandt  
Sandman, VT

Taylan Topcu,  
VT

# Phases of T&E Transformation



# Historical Fiction?!

## Mission proxy

- Operation Safe Passage



- Operation Safe Lego



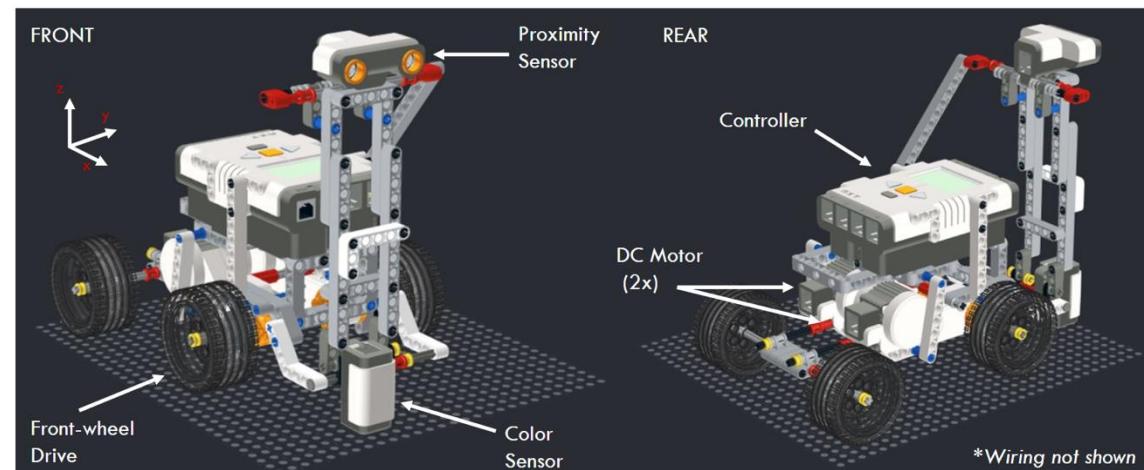
# Historical Fiction?!

## 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



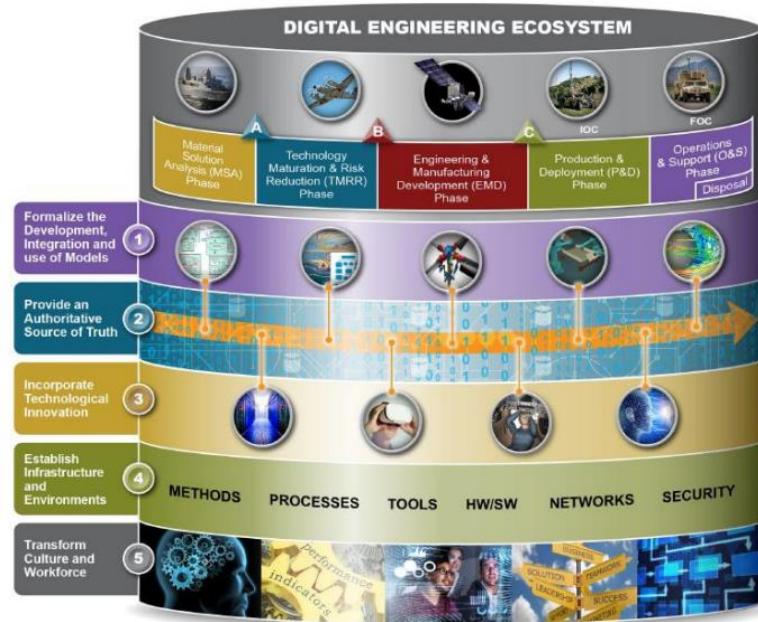
- Operation Safe Lego



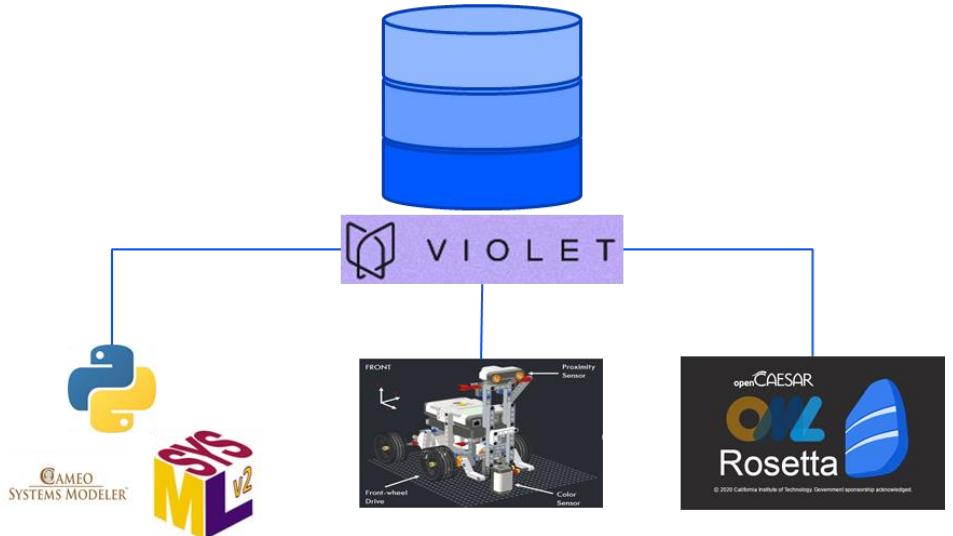
# Historical Fiction?!

## System proxy

- Operation Safe Passage



- Operation Safe Lego



# Key Performance Parameters (KPPs)

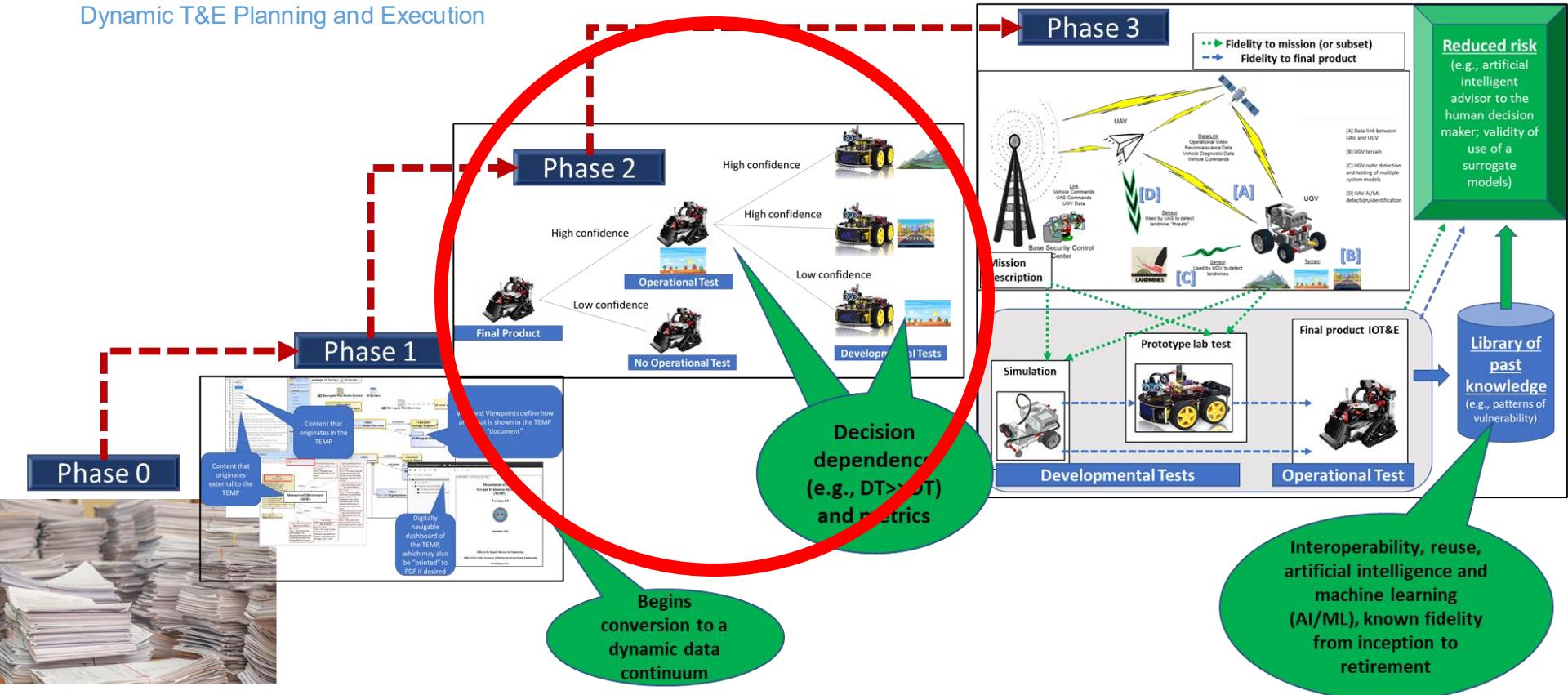
Measures	Definition
Detection Rate*	% mines detected correctly
False Alarm Rate, Detection*	% mine detection false alarms
Classification Rate*	% mines classified correctly
False Alarm Rate, Classification*	% mine classification false alarms
Pathway Creation Time*	Time taken to clear a pathway from start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

# Mock interim design review" (IDR)

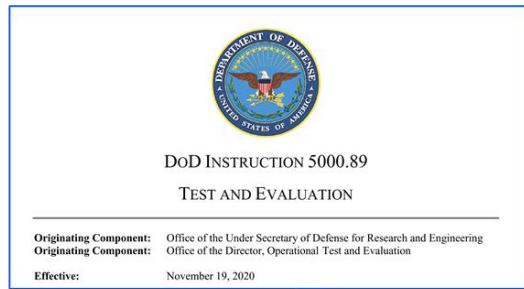
Category	Description
Entities involved and role	<ol style="list-style-type: none"><li>1. DOT&amp;E personnel with a role as the T&amp;E decision authority</li><li>2. AIRC researchers with a role to conduct the T&amp;E activities and provide artifacts to the decision authority</li></ol>
Supporting artifacts	<ol style="list-style-type: none"><li>1. PowerPoint slides as a read ahead</li><li>2. Lego mission model</li><li>3. Lego (SysML) system model</li><li>4. Lego CAD models</li><li>5. MBTEMP</li><li>6. Modeling and Simulation (M&amp;S) models (Python-based)</li><li>7. Analytical results</li><li>8. IDSK (decision dashboard)</li></ol>
Means of conducting the review	<ol style="list-style-type: none"><li>1. Virtual (MS Teams)</li><li>2. Use of PowerPoint to guide the review</li><li>3. Use of real-time presentation of models and data (e.g., MBTEMP and IDSK)</li></ol>
T&E use cases	<ol style="list-style-type: none"><li>1. Developmental tests resulting in design change after T&amp;E showed the inability to achieve traction at lower weights</li><li>2. Developmental tests resulting in design and performance trade after T&amp;E showed the tire diameter and terrain are co-dependent</li><li>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</li></ol>
Review phases	<ol style="list-style-type: none"><li>1. Initial IDR (June 2024)</li><li>2. Delta IDR (Aug 2024)</li></ol>

# Phase of T&E Transformation

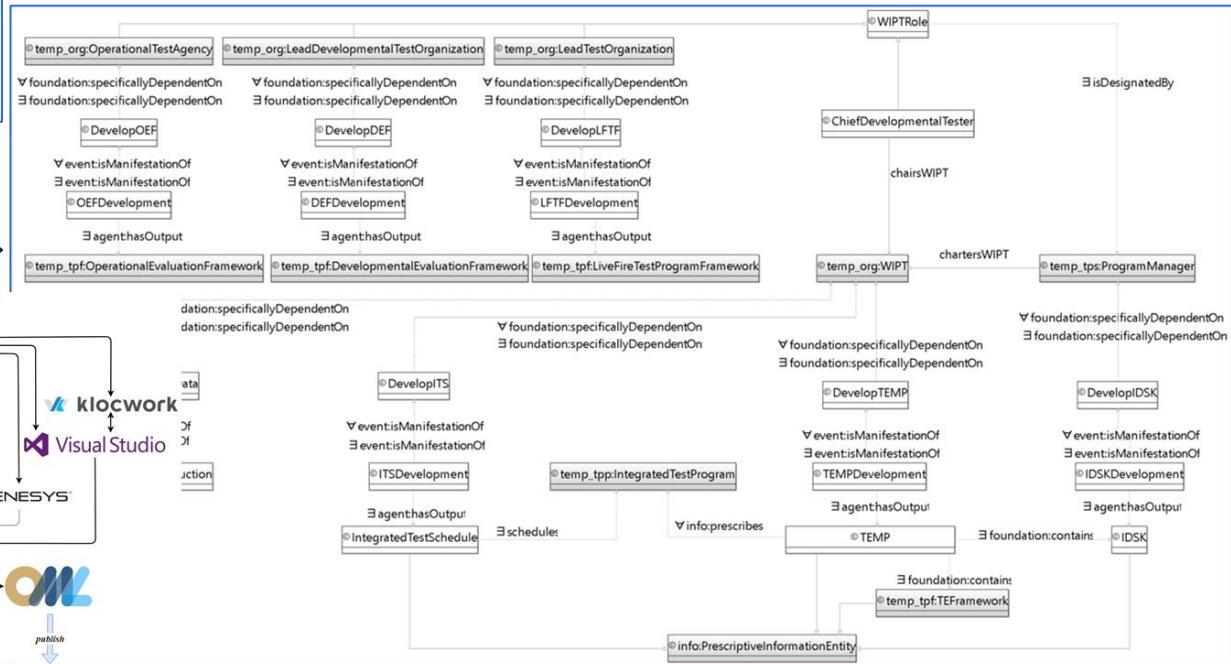
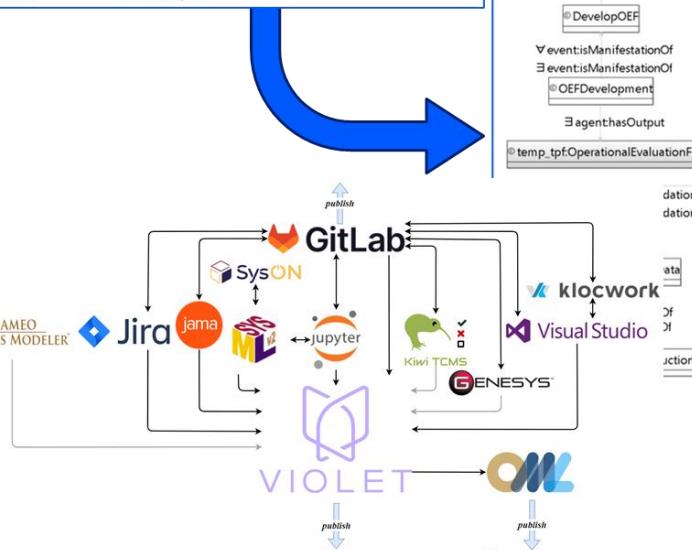
## Dynamic T&E Planning and Execution



# Digital Test and Evaluation Master Plan (dTEMP)



## Mapping the DoDI Into Foundational Tooling



# Digital Test and Evaluation Master Plan (dTEMP)

## Dashboard

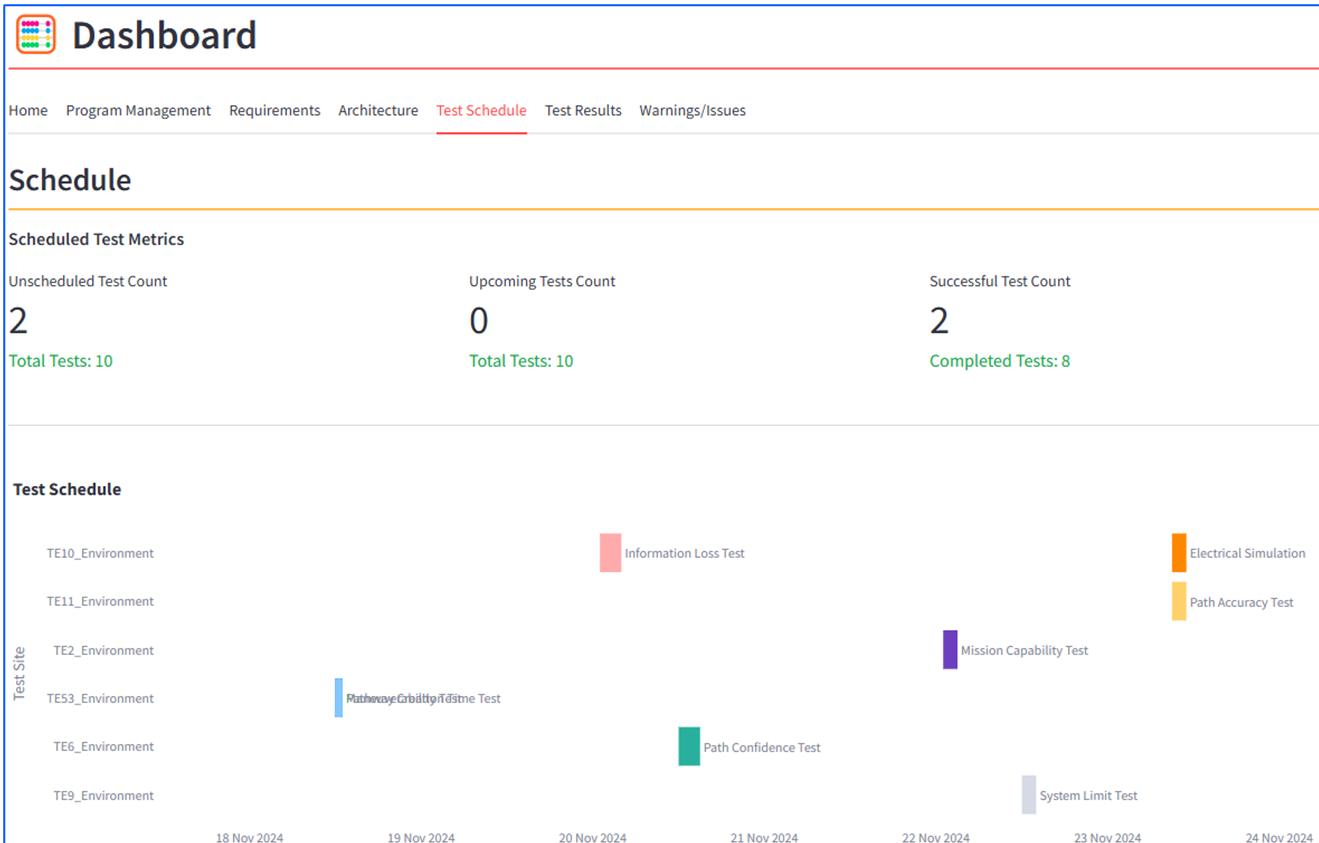
Home Program Management Requirements Architecture **Test Schedule** Test Results Warnings/Issues

### Schedule

**Scheduled Test Metrics**

Unscheduled Test Count	Upcoming Tests Count	Successful Test Count
2	0	2
Total Tests: 10	Total Tests: 10	Completed Tests: 8

**Test Schedule**



Test Site

Test Type Legend:

- Electrical Simulation
- Path Accuracy Test
- Pathway Creation Time Test
- Information Loss Test
- Path Confidence Test
- Mission Capability Test
- System Limit Test

Test Details:

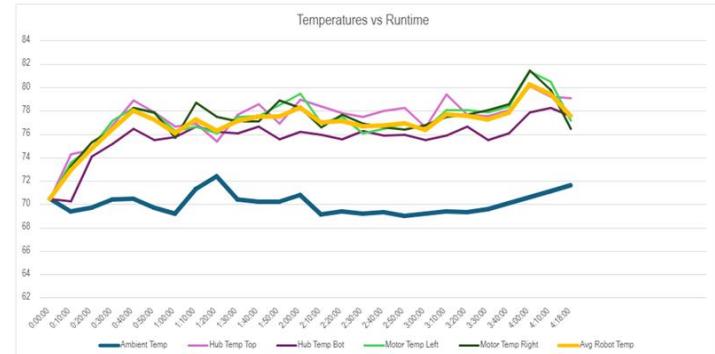
- TE10\_Environment: Electrical Simulation
- TE11\_Environment: Path Accuracy Test
- TE2\_Environment: Pathway Creation Time Test
- TE53\_Environment: Pathway Creation Time Test
- TE6\_Environment: Pathway Creation Time Test
- TE9\_Environment: Pathway Creation Time Test
- Pathway Creation Time Test: Pathway Creation Time Test
- Information Loss Test: Information Loss Test
- Path Confidence Test: Path Confidence Test
- Mission Capability Test: Mission Capability Test
- System Limit Test: System Limit Test

Timeline: 18 Nov 2024, 19 Nov 2024, 20 Nov 2024, 21 Nov 2024, 22 Nov 2024, 23 Nov 2024, 24 Nov 2024

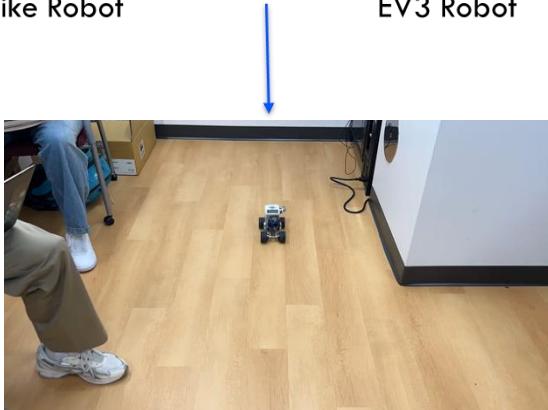
# A job fit for students: Battery testing



EV3  
Test 1

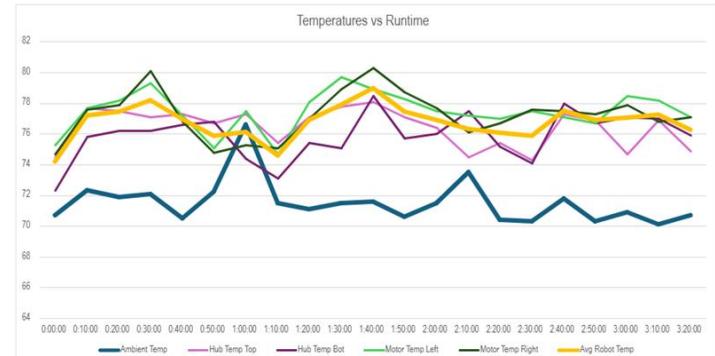


Spike Robot



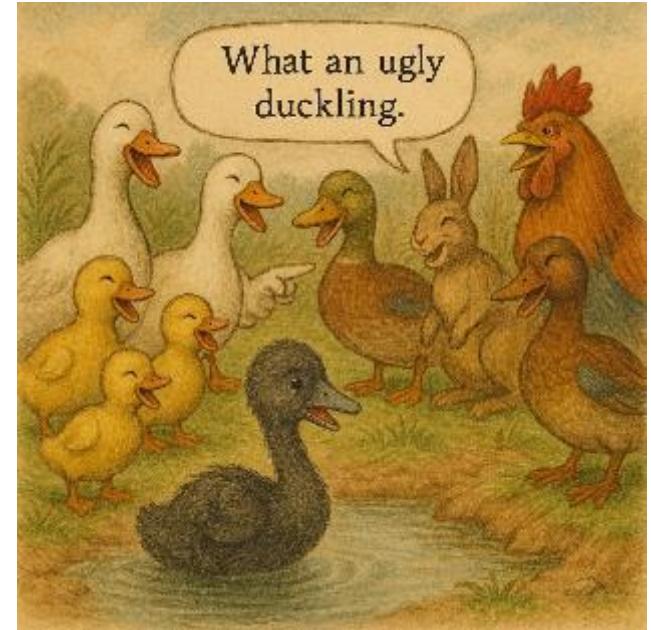
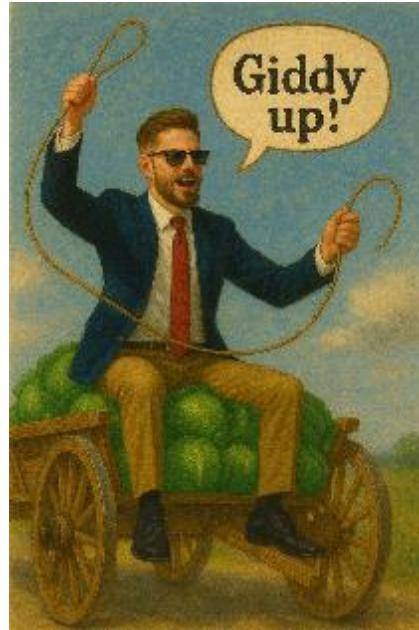
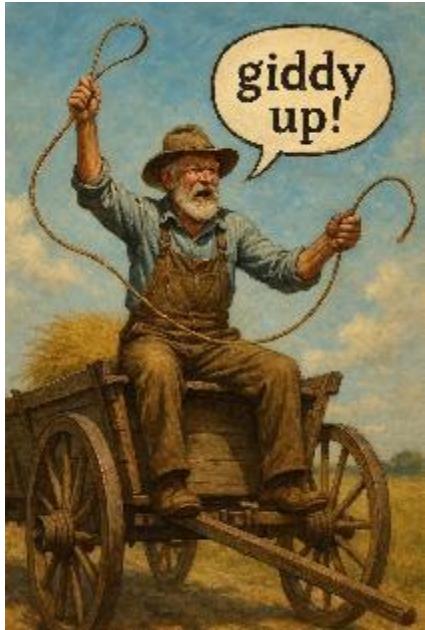
EV3 Robot

EV3  
Test 2



# Reactions to our work???

We hope not



# Questions?

Paul Wach, [paulw86@vt.edu](mailto:paulw86@vt.edu)

Alejandro Salado, [alejandrosalado@arizona.edu](mailto:alejandrosalado@arizona.edu)

Joe Gregory, [joegregory@arizona.edu](mailto: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



Low confidence



High confidence



Final Product

No Operational Test

# A job fit for students



Brandt Sandman



Trevor lerardi



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

Legacy System



New System



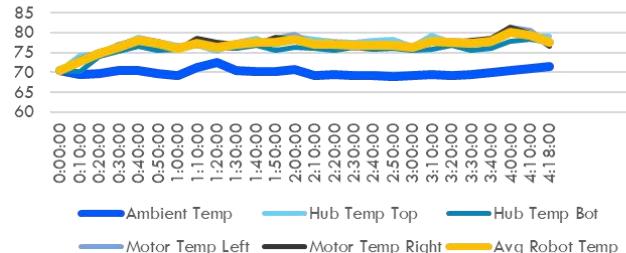
High confidence



High confidence



Temperatures vs Runtime



Battery Voltage vs Runtime

