Georgia Tech Research Institute

Problem. Solved.

Autonomy Validation, Introspection, and Assessment (AVIA)

System Description and Lessons Learned

By Miles Thompson

Research Engineer II

Aerospace, Transportation, and Advanced Systems Lab

DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

Intro to AVIA

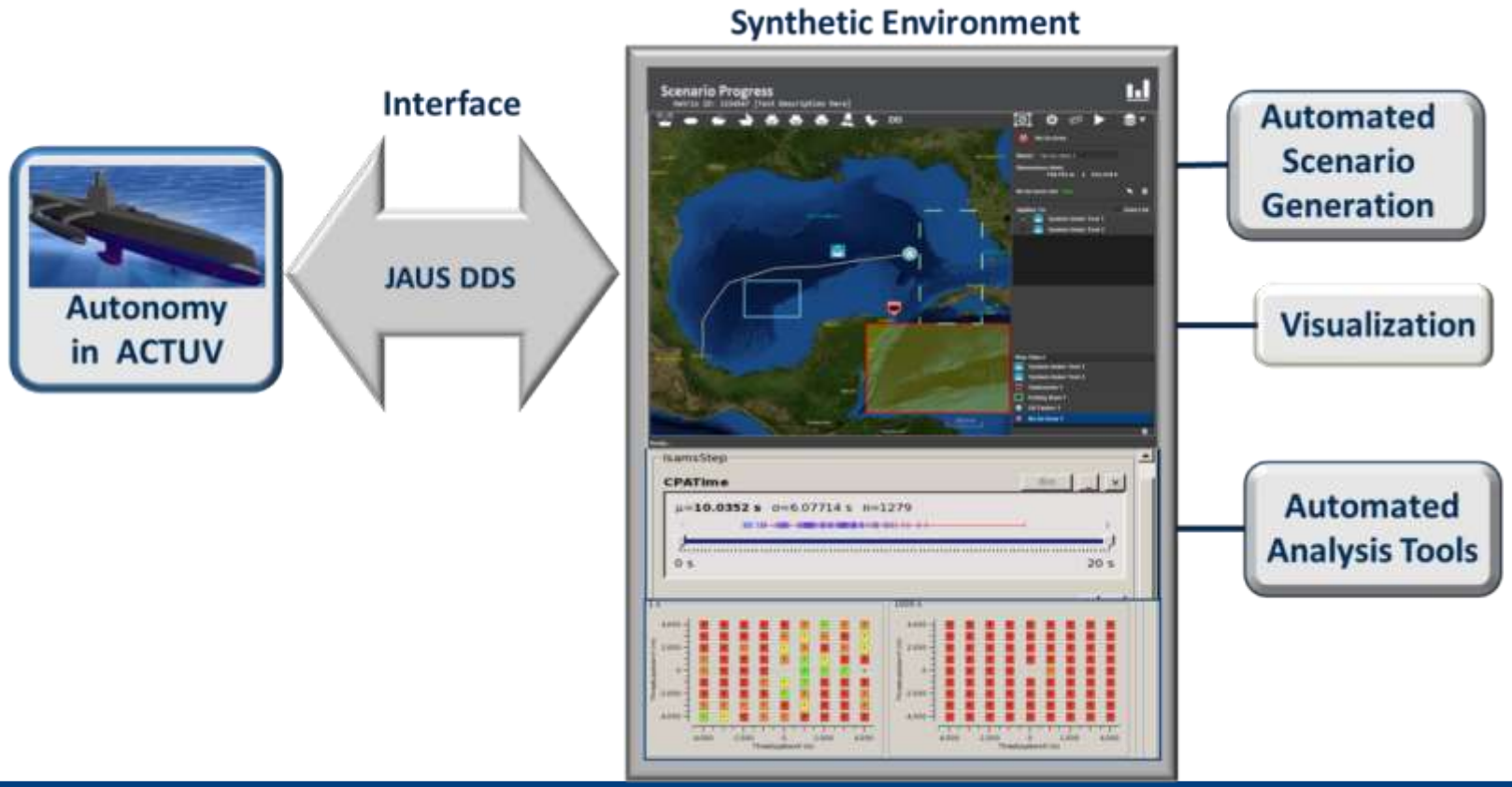
- AVIA is a DARPA project to develop assessment tools for the Anti-Submarine Warfare Continuous Trail Unmanned Vessel (ACTUV) program
- Specifically designed to validate, introspect, and assess autonomous systems (autonomy)
- Developed unique capabilities:
  - Metrics for autonomous behaviors (tracking, collision regulations)
  - Iterative test-selection based on uncertainty
- Designed with modular infrastructure to facilitate future capabilities
  - Systems Under Test (SUTs)
  - Test-selection methods
  - Metrics

DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

Autonomy Validation, Introspection, and Assessment (AVIA)

Georgia Tech Research Institute

Rapidly conduct massive numbers of analytic assessments of ACTUV autonomy in complex, dynamic scenarios



Impact

- Assess capabilities beyond operational envelope: 10x – 100x more scenarios
- Address ACTUV operational concerns
- Provide >10x reduction in time to run & generate results from each simulation, from 42 days to < 1 day

DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

How does AVIA accomplish this?

Georgia Tech Research Institute

- **Sample entire test domain intelligently and without bias**
  - Currently using Latin hypercube sampling
  - Allow perturbations and emergent conditions to arise
  - Upgradeable: can implement future sampling techniques
- **Parallel scenario execution**
  - Demonstrated 1000 1-hour scenarios in < 24 hours (16 hours)
  - Automatic metrics evaluation
  - Metrics stored in MySQL database, easy to visualize with analysis graphical user interface (GUI)
- **Metrics are requirement dependent and include introspection to increase understanding of autonomy logic**

DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

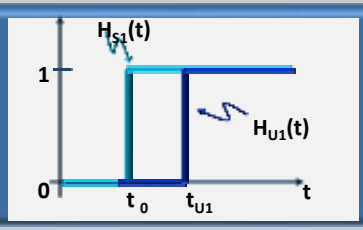
AVIA Handles all Autonomy Processes

Georgia Tech Research Institute

Assessment Tools

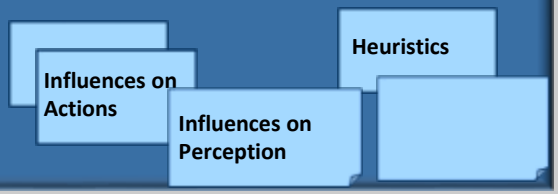
Configurable environment
(Automated scenario generator)

Unexpected event generator



Critics (evaluation modules)

Library of evaluation modules



Exploration tools

Assessment tools

Presentation tools
(including uncertainty)

System Under Test (SUT)

Reflective Processes

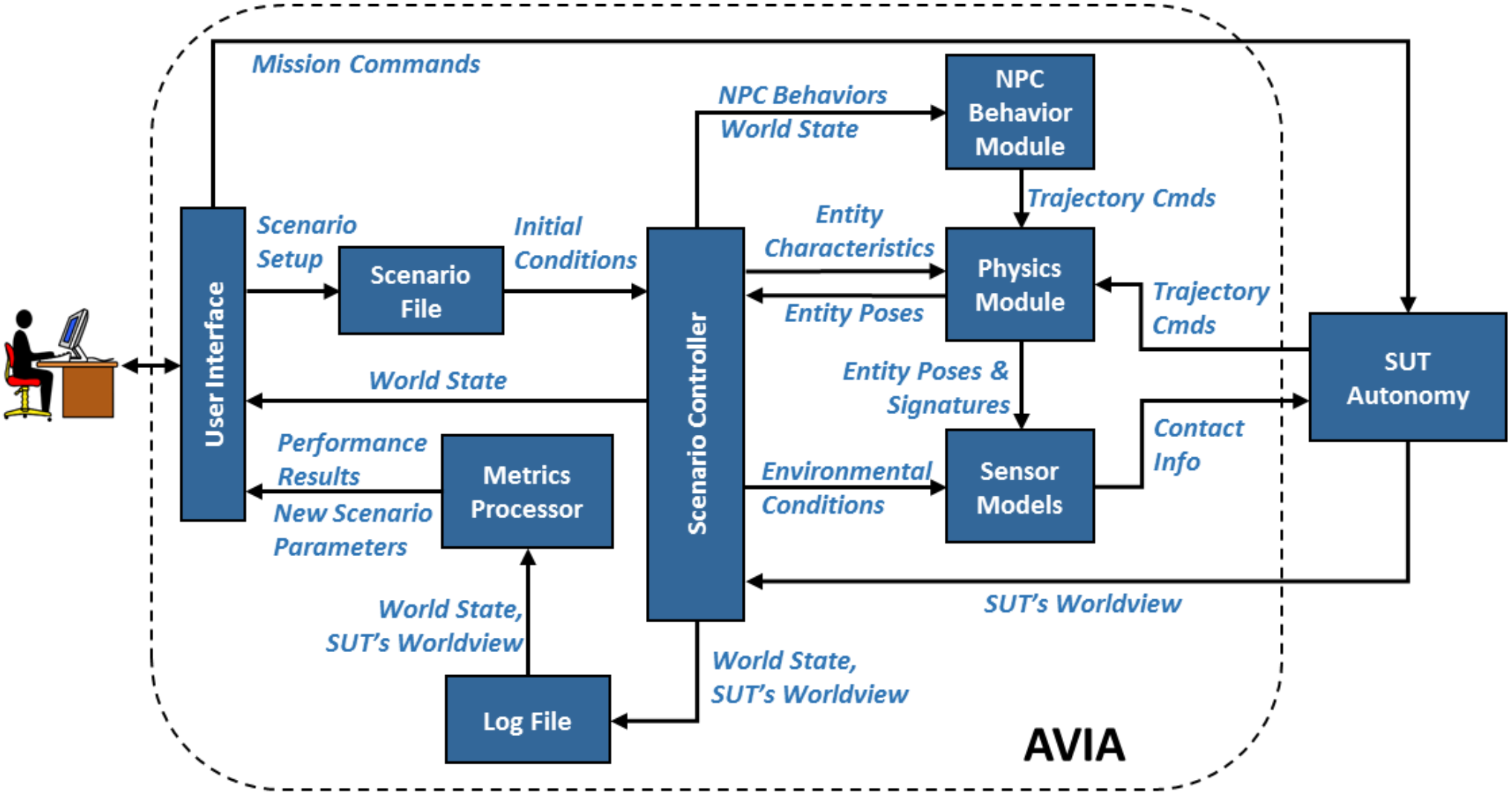
Deliberative Processes

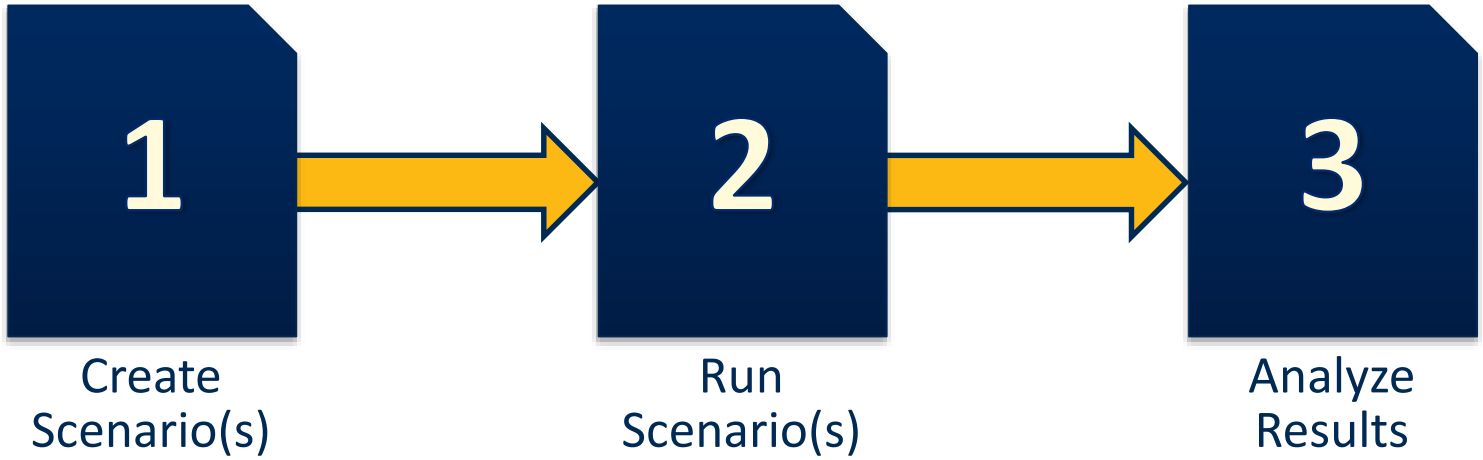
Reactive Processes

Autonomy Processes

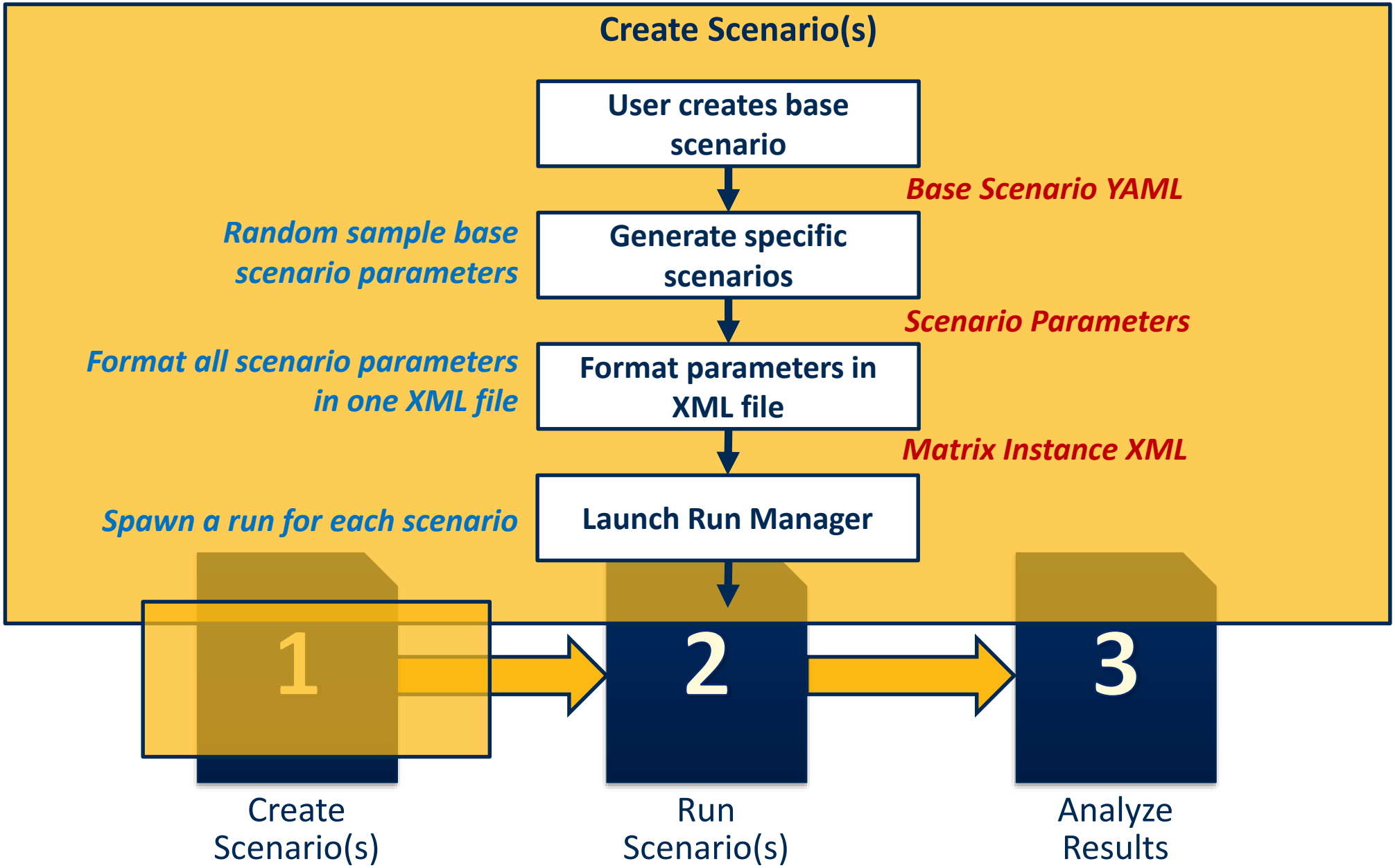
DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

AVIA Architecture

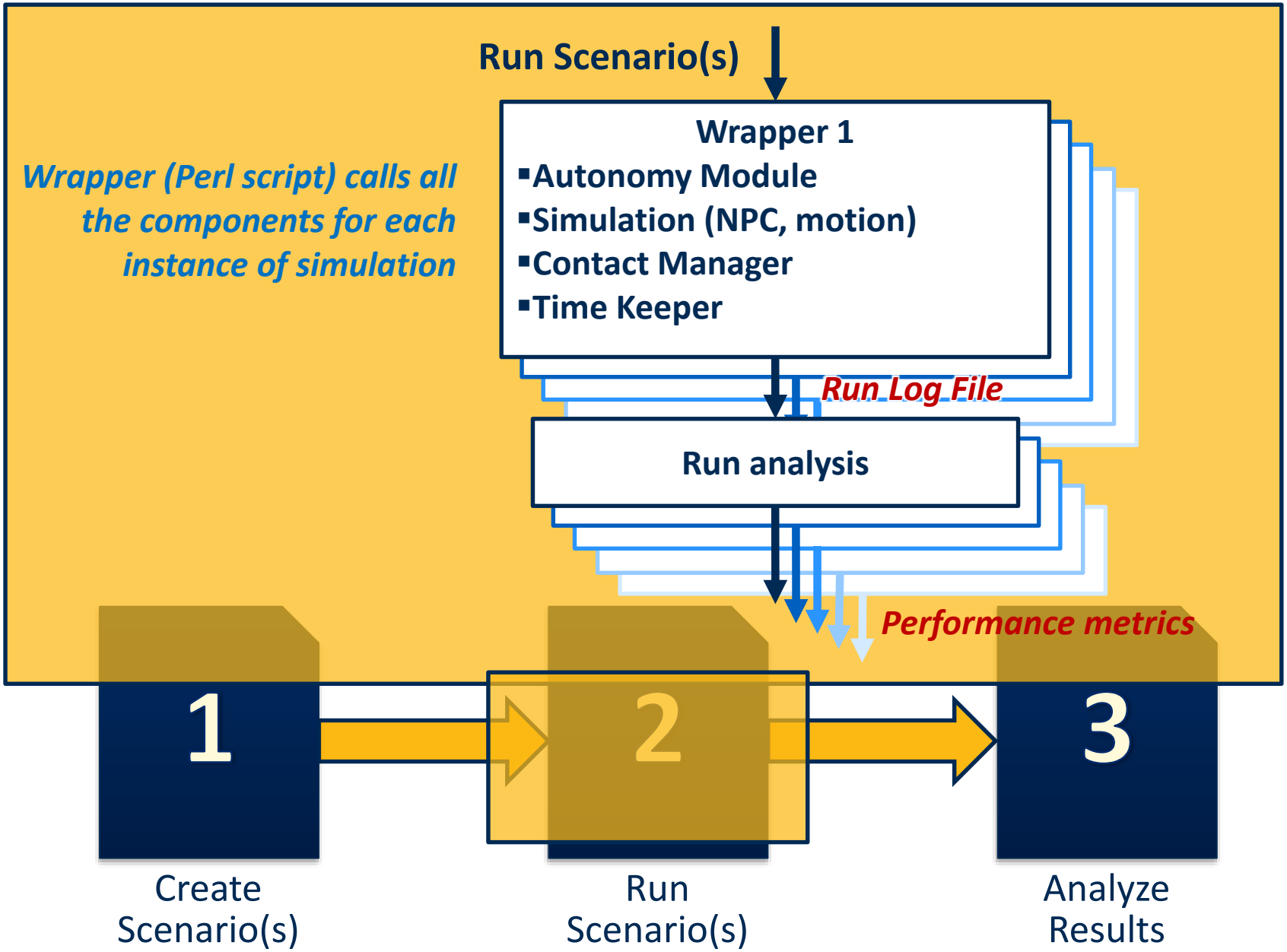




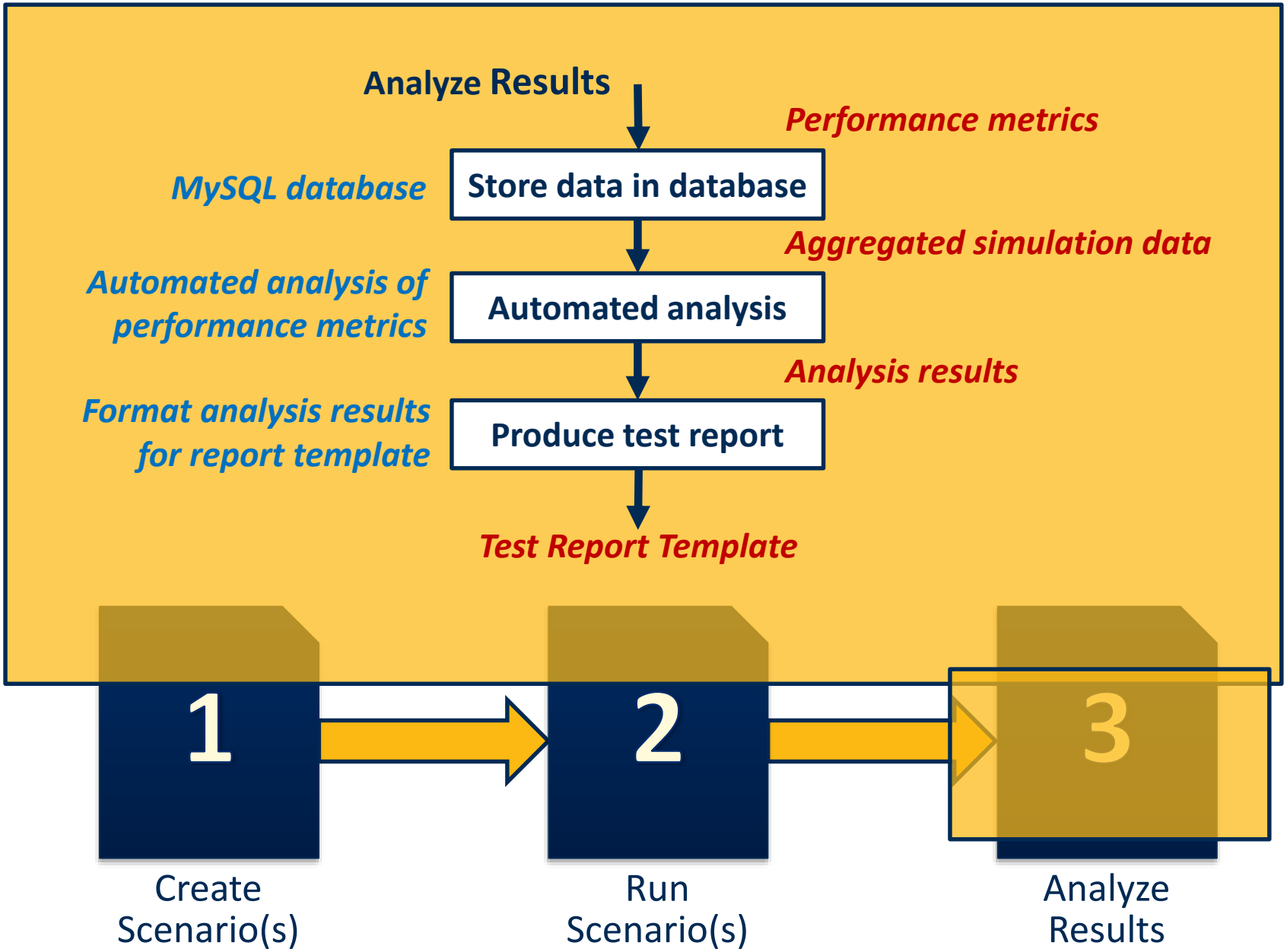
DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited



DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

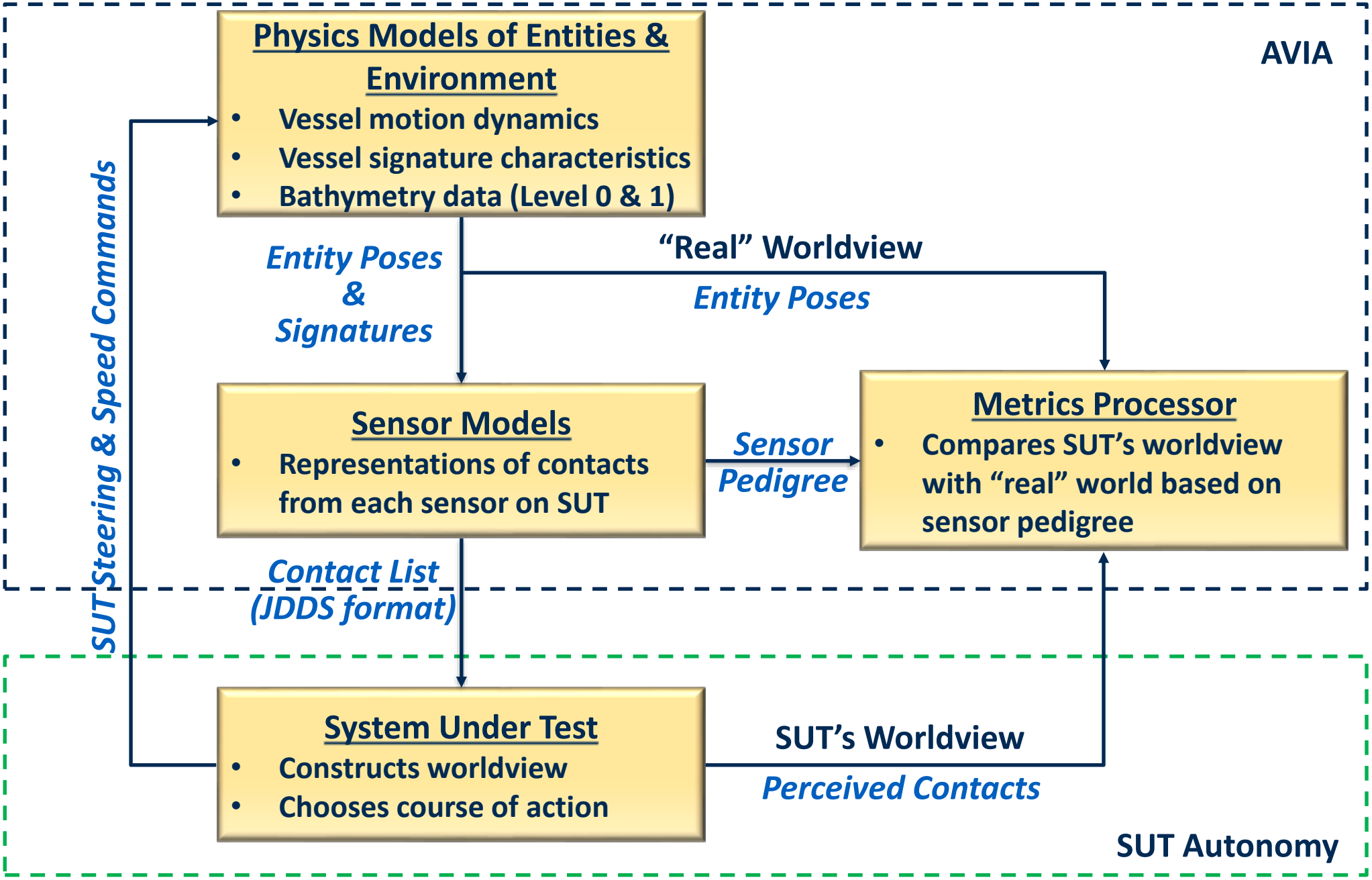


DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

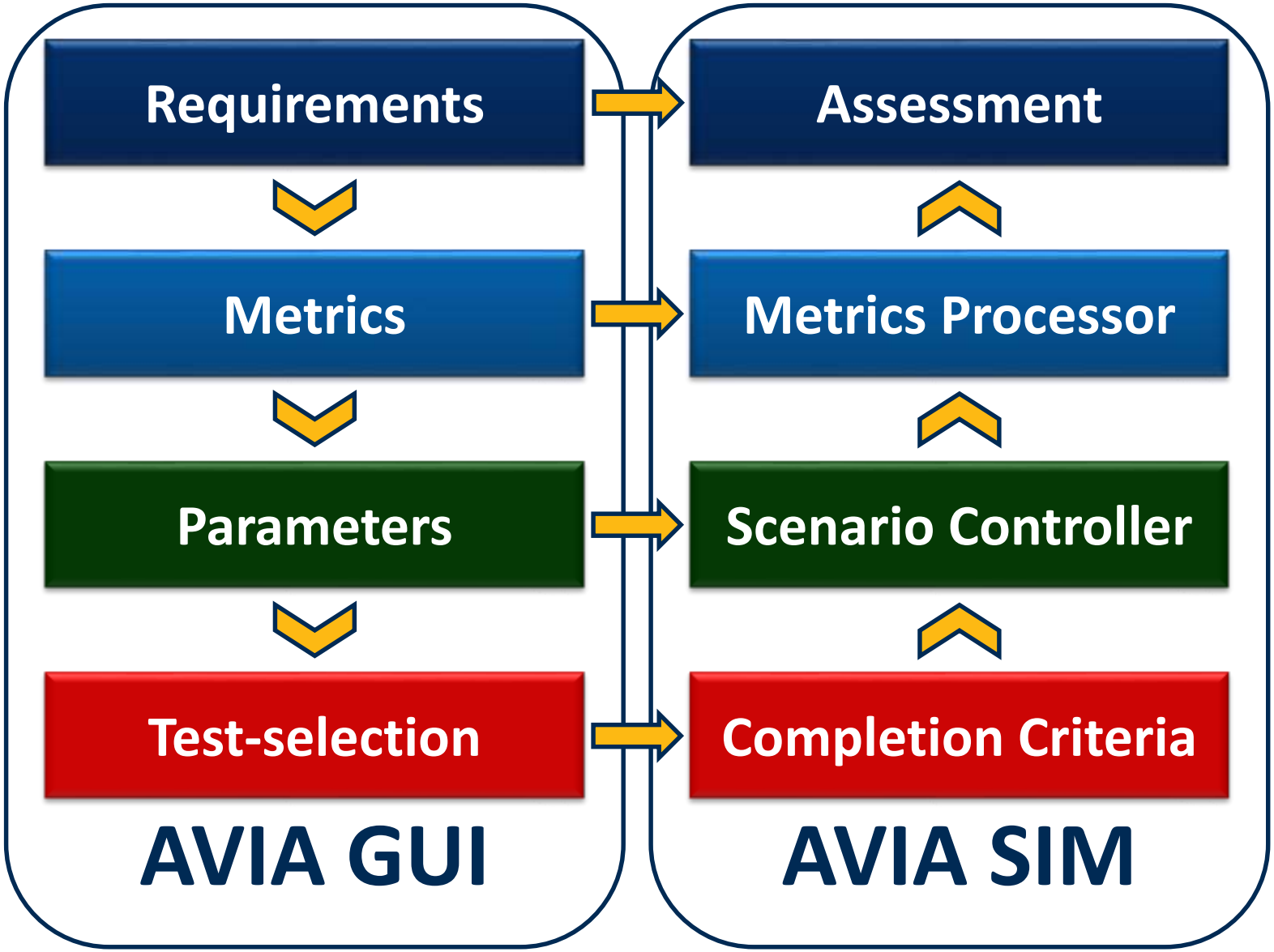


DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

How Data Flows Thru AVIA

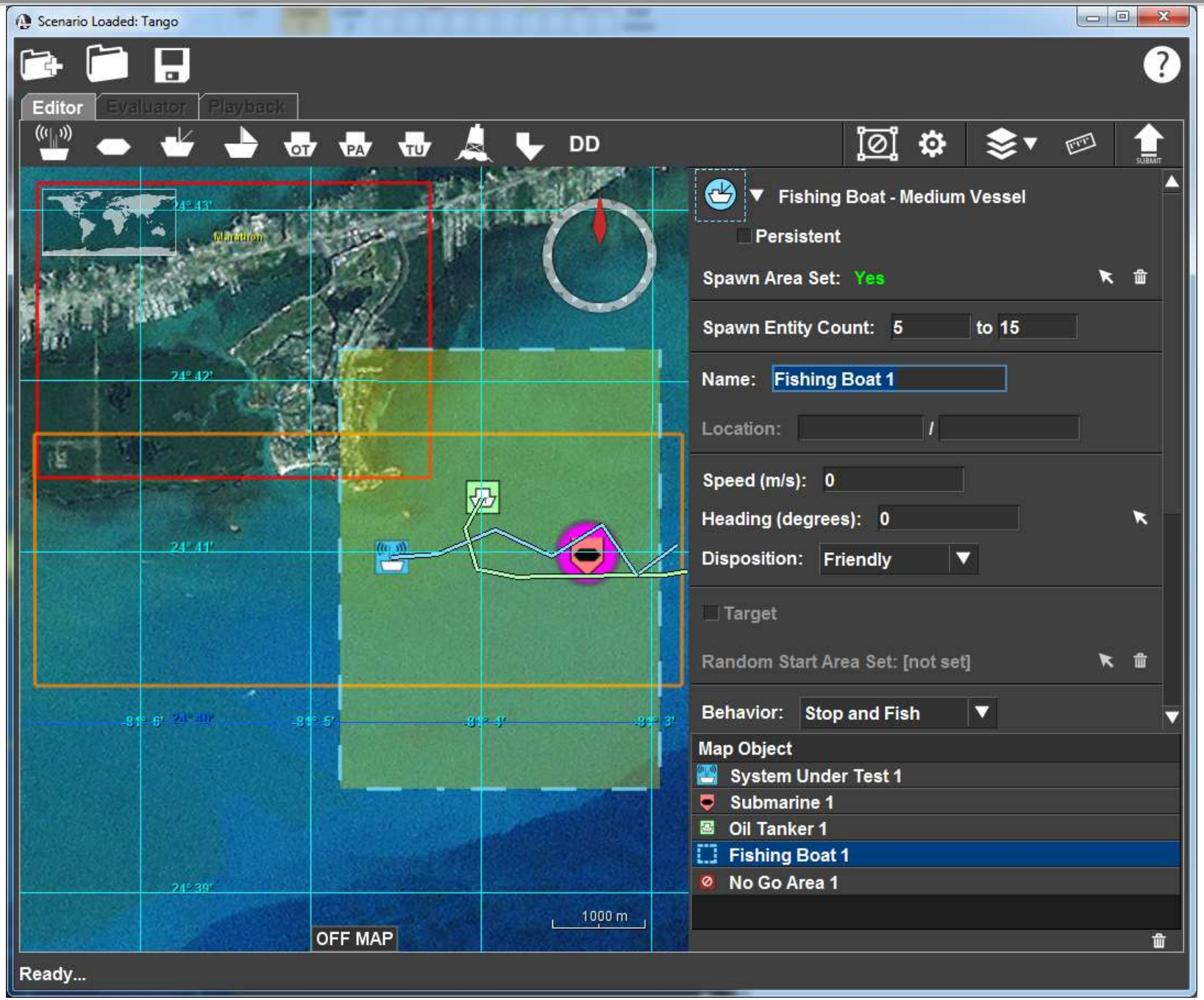


AVIA is Designed for the Tester



DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

AVIA GUI


Scenario Loaded: Tango
Editor
Evaluator
Playback
OT
PA
TU
DD
Fishing Boat - Medium Vessel
Persistent
Spawn Area Set: Yes
Spawn Entity Count: 5 to 15
Name: Fishing Boat 1
Location: /
Speed (m/s): 0
Heading (degrees): 0
Disposition: Friendly
Target
Random Start Area Set: [not set]
Behavior: Stop and Fish
Map Object
System Under Test 1
Submarine 1
Oil Tanker 1
Fishing Boat 1
No Go Area 1
1000 m
OFF MAP
Ready...

AVIA GUI Evaluator Tab

Scenario Loaded: Dolomite

Editor | Evaluator | Playback

Scenario Progress

Matrix ID: 1012

Metric Variance

Steady State: 56.00%
Converging: 21.00%
Diverging: 23.00%

A description of what we are looking for in this area might go here. Here's...

Sample Status: Complete

Percent Complete: 100%
Total Runs: 5
Completed: 5
Errors: 0

Scenario Details

Name: Dolomite
Last Edited: 08/19/2015
By: tm132
Submitted: 08/19/2015

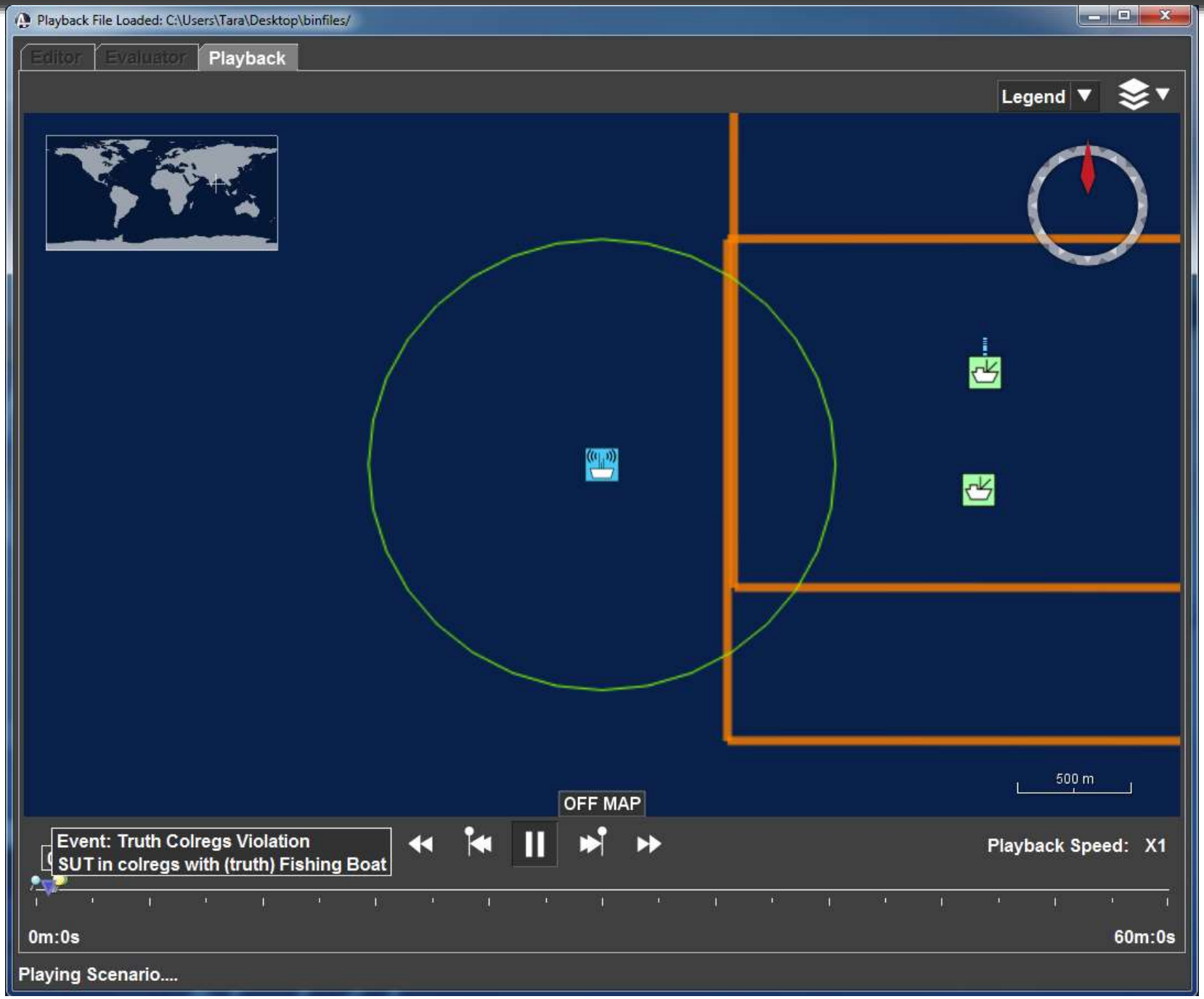
Completed Tasks | Metric Details | Raw TMT Log

☑ Show Errors ☑ Show Completed Without Errors ☑ Show Uneventful

Test Case 434205	COMPLETE ✓
Test Case 434206	COMPLETE ✓
Test Case 434207	COMPLETE ✓
Test Case 434208	COMPLETE ✓
Test Case 434209	COMPLETE ✓

Ready...

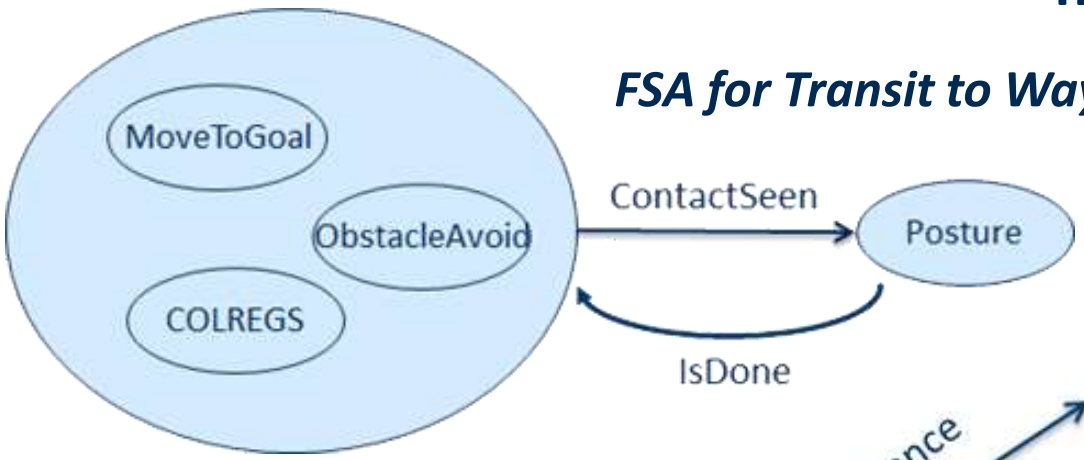
AVIA GUI Playback Tab

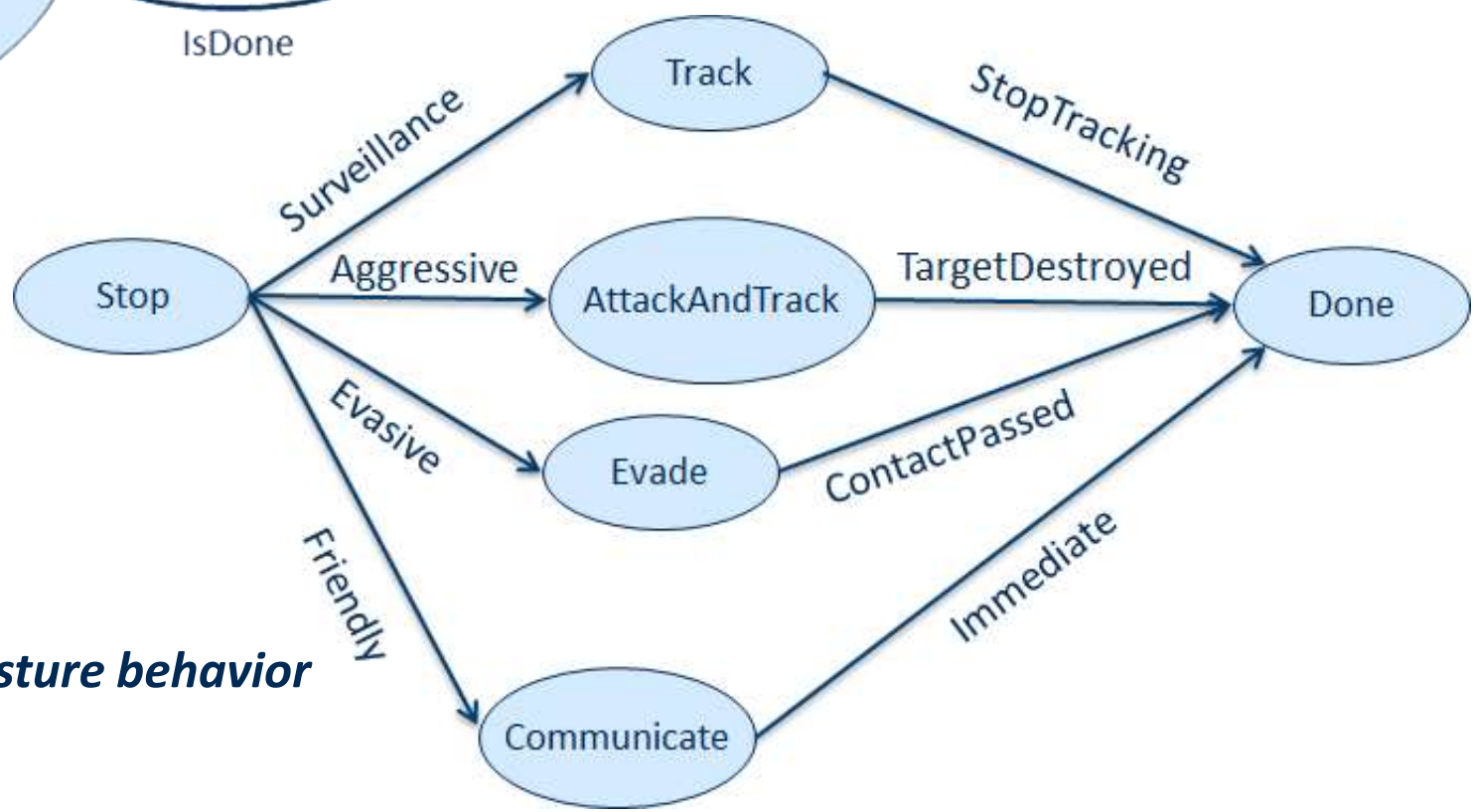


Scenario Controller

Other Vessel Behavior Module – uses Finite State Automata (FSA) for non-player characters (NPCs)

FSA for Transit to Waypoint



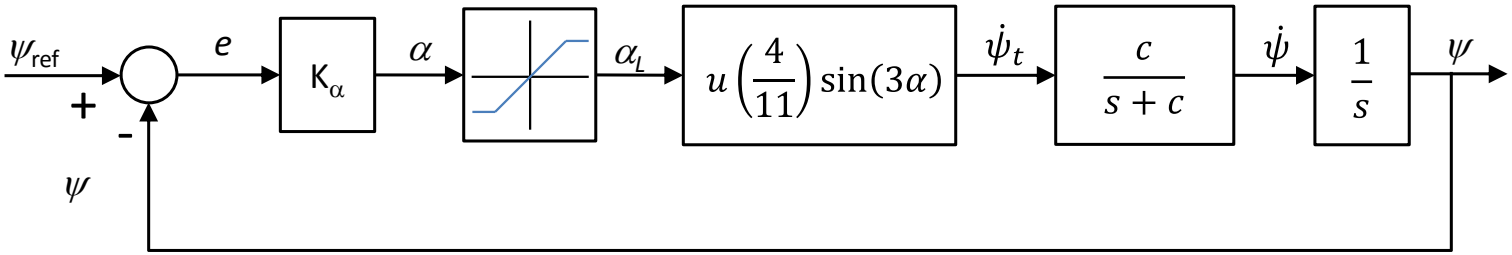


FSA for Posture behavior

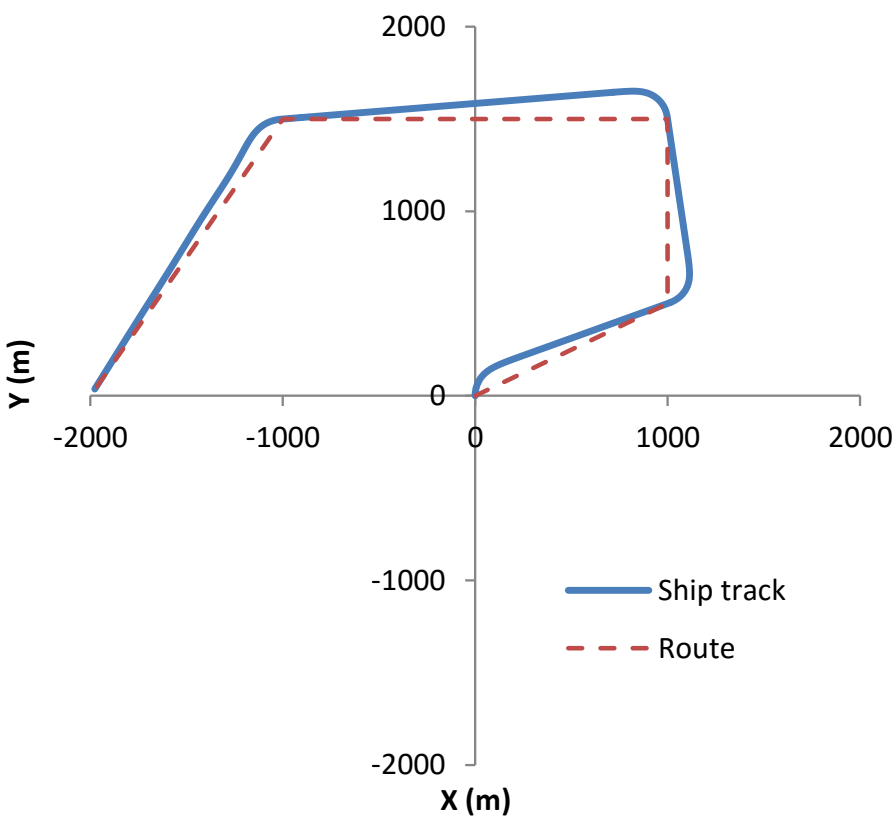
DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

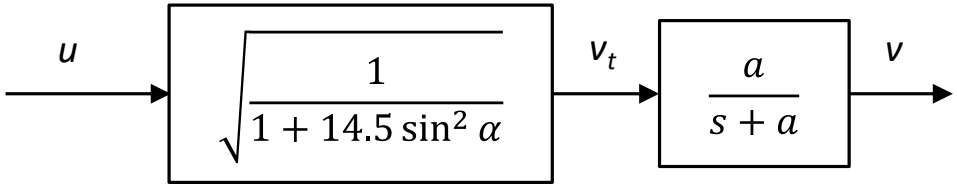
Scenario Controller

Physics Module – simplified models for speed and heading dynamics



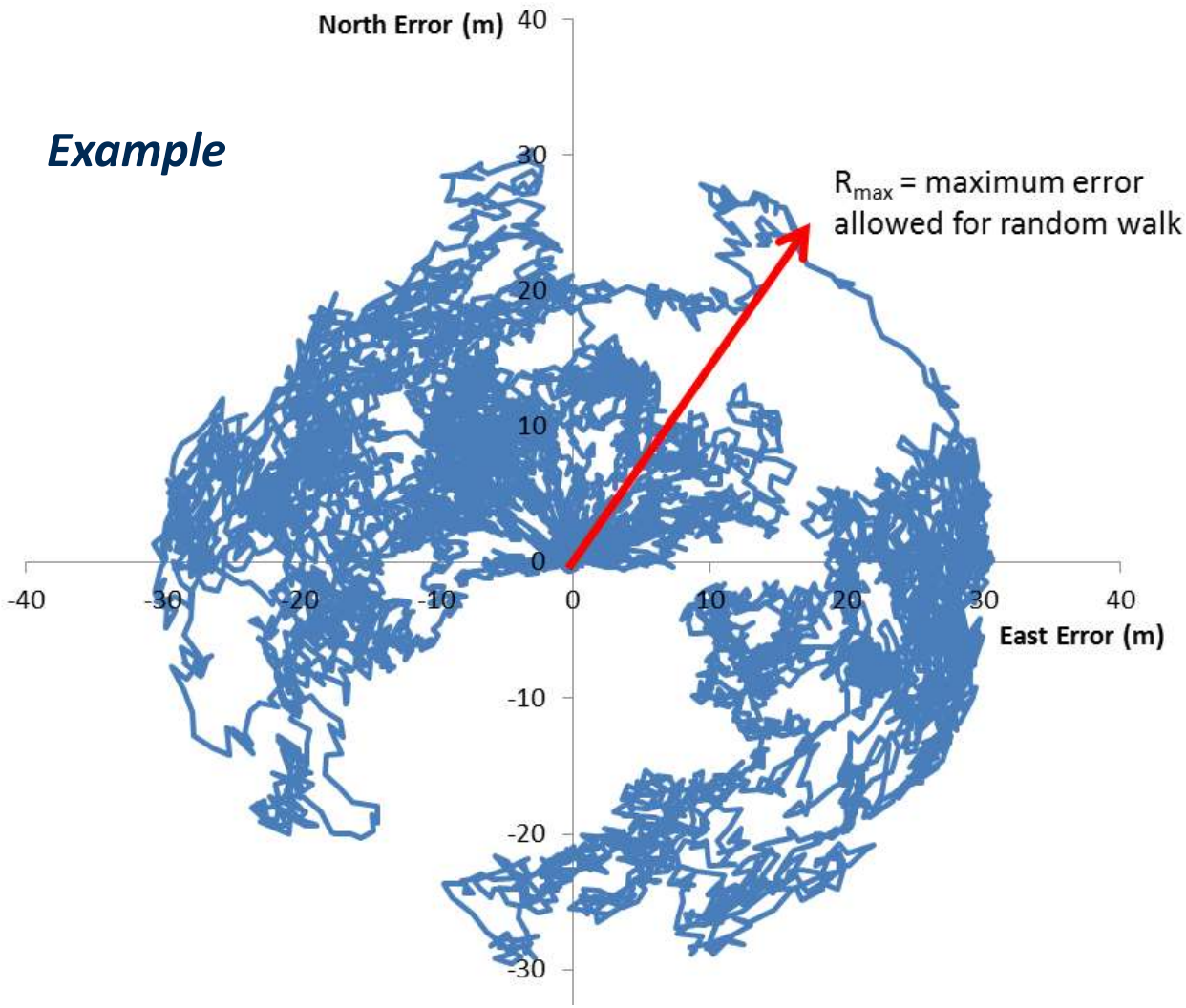
Heading model





Speed model

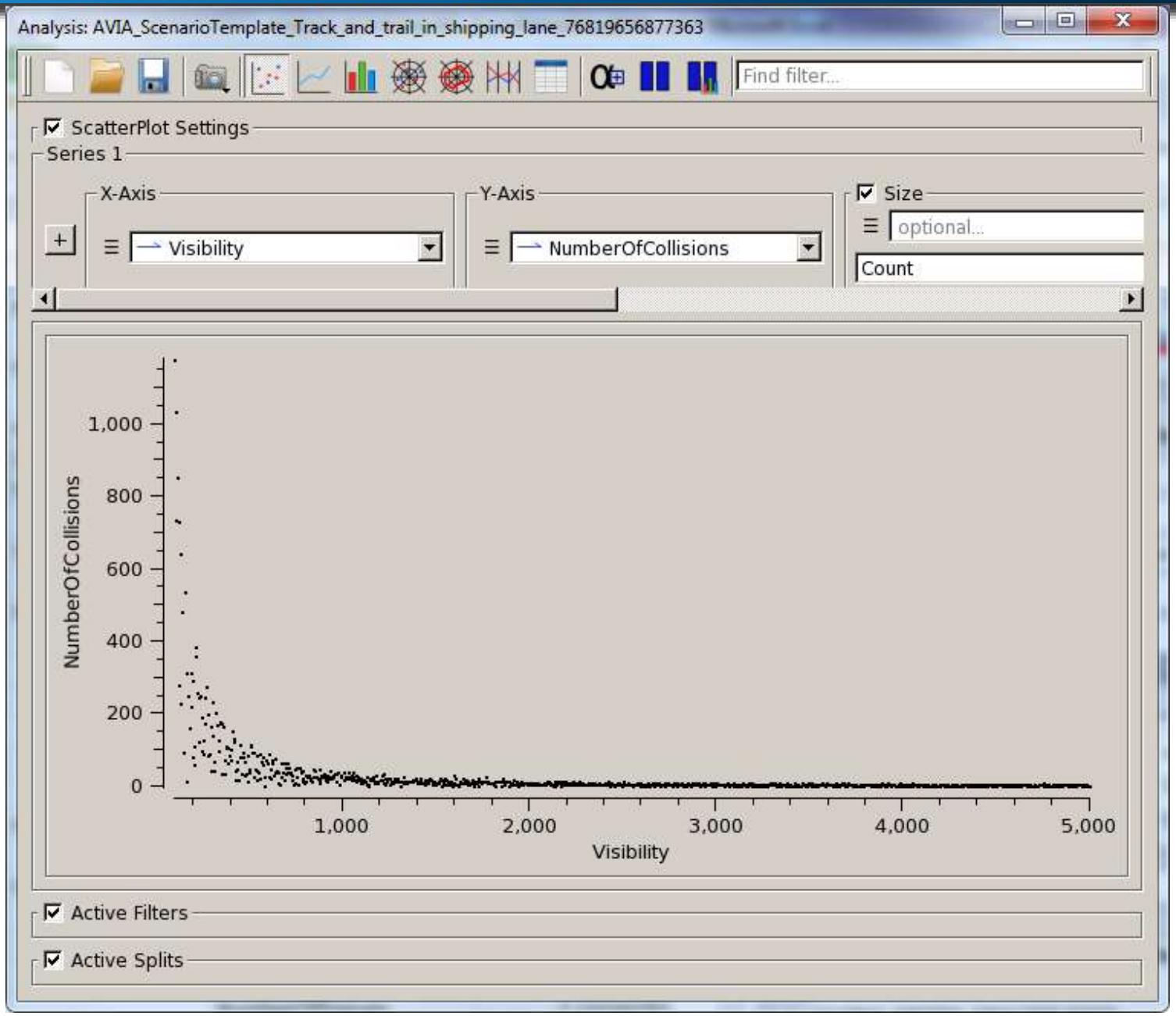
Sensor Models – random walk noise added to true data



Random walk error added to position data

DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

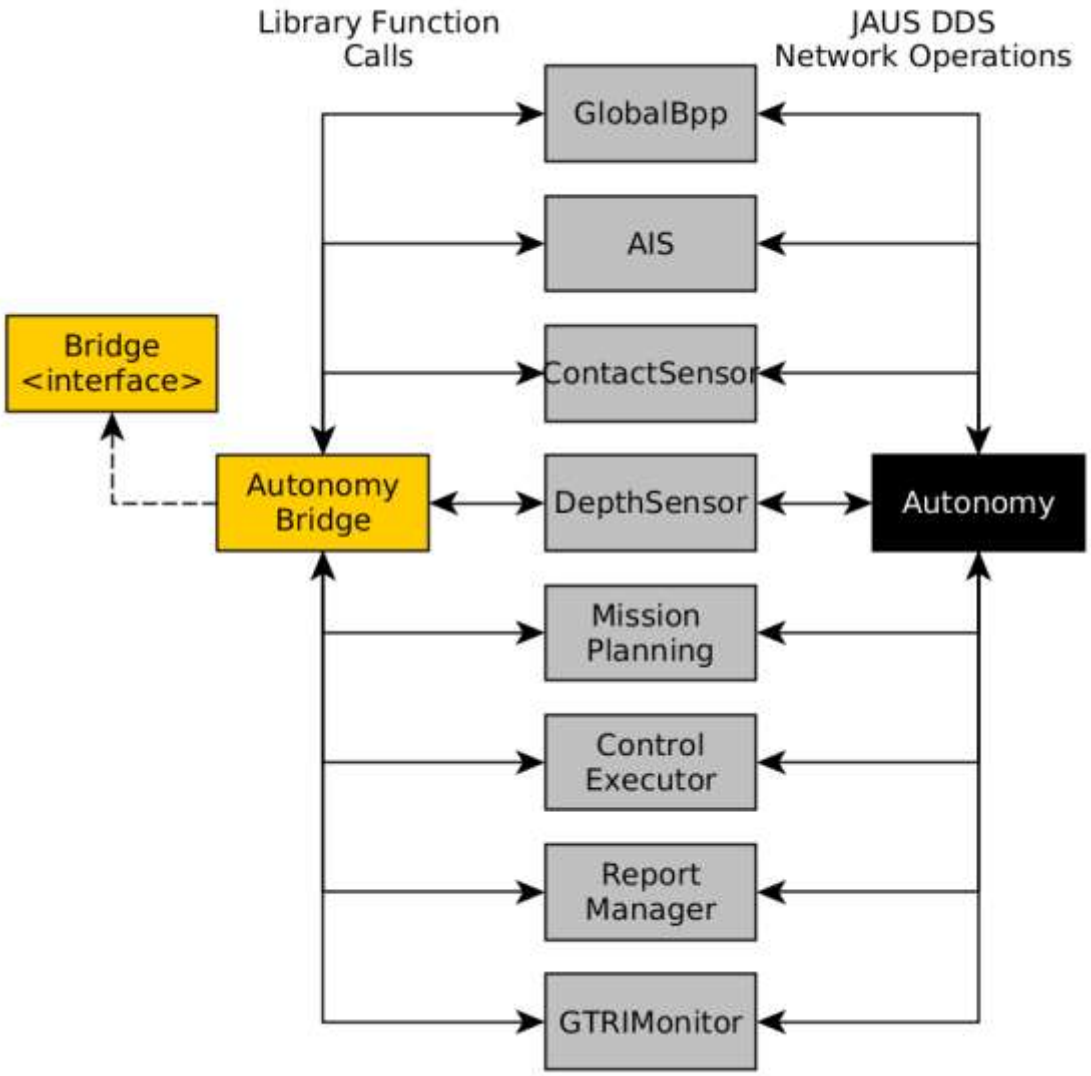
Metrics Processor & Automated Analysis



Test Matrix Tool analysis GUI: Scatterplot visualization

DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

SUT Autonomy



Autonomy communication bridge

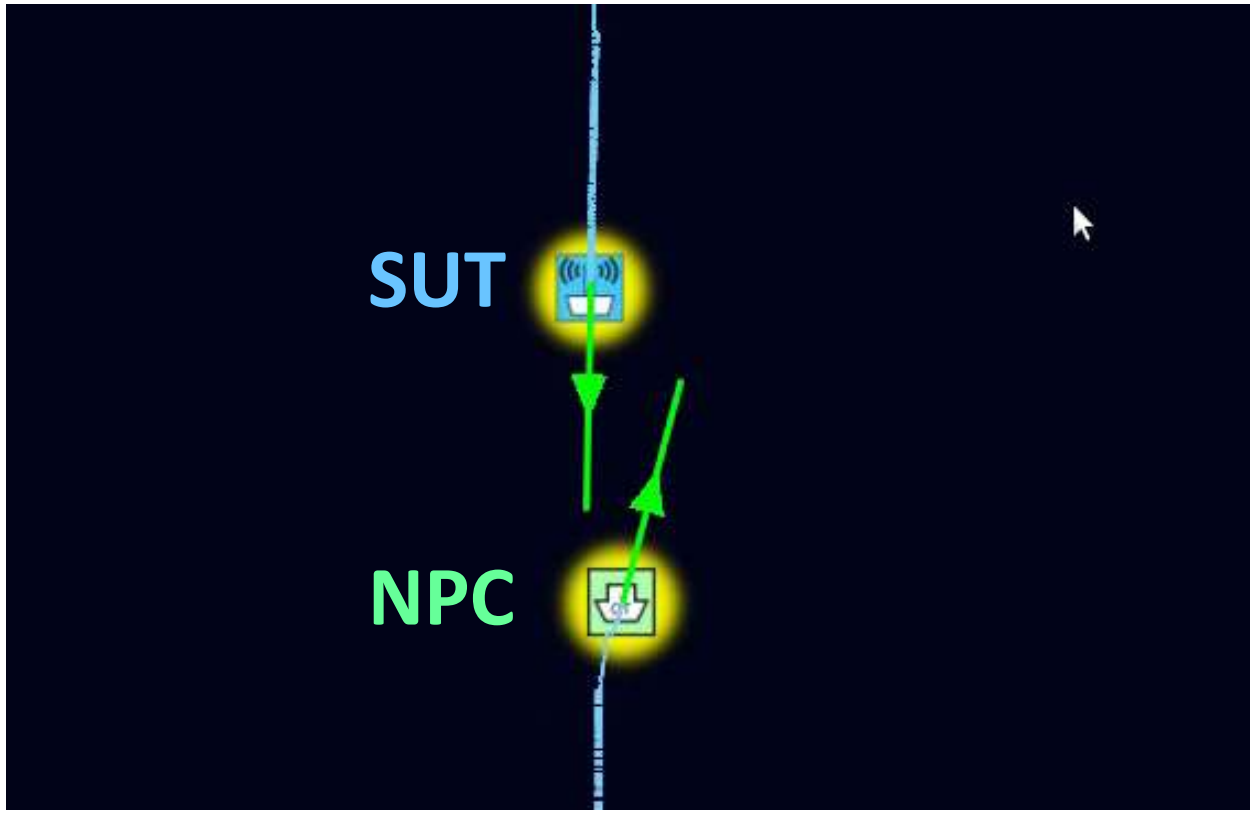
Iterative Test-selection

- Added iterative test-selection based on performance of previous samples
  - Searches the test domain for the highest estimated variances, the highest uncertainty
  - Continues to iterate until completion criteria are met
    - » Completion criteria is programmable, currently using steady-state variance of performance across the test domain
  - Estimates calculated using spatial interpolation
  - Does not bias to best/worst performance

DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

Lessons Learned

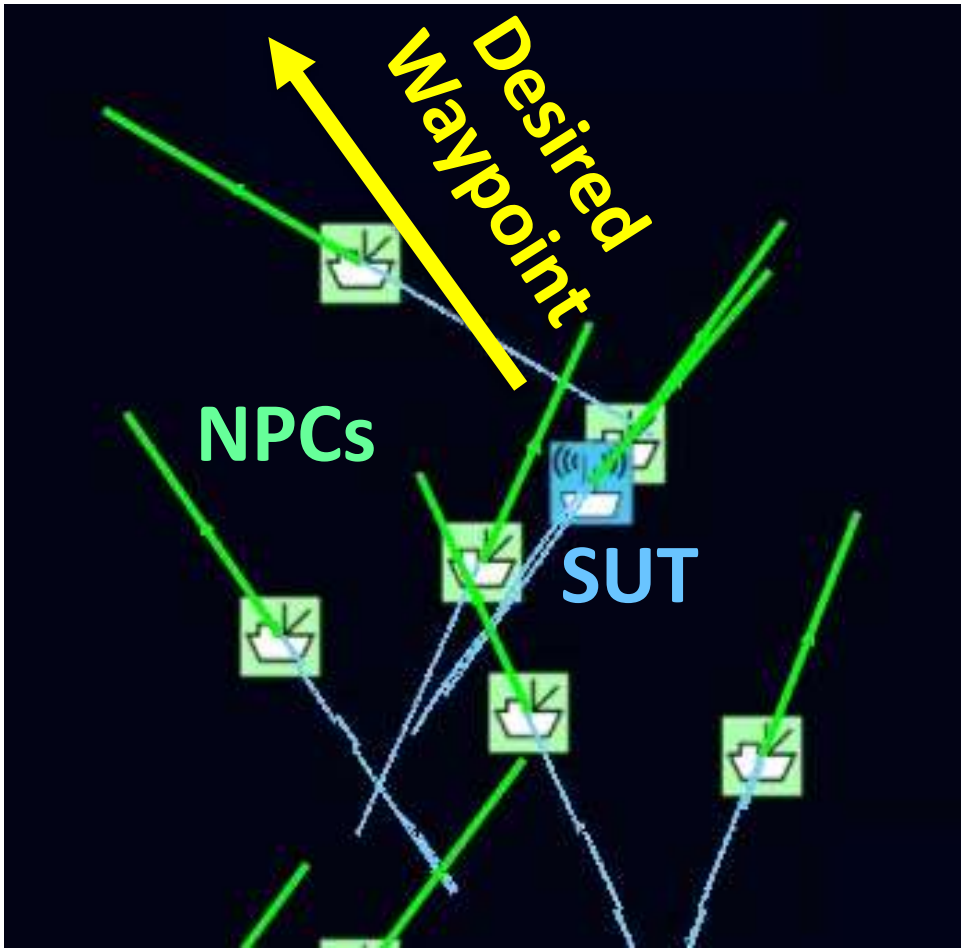
- **COLREG classification can fluctuate over time**



  — **Would have been head-on (both give-way), but SUT all-stop so now the non-player character (NPC) is give-way**

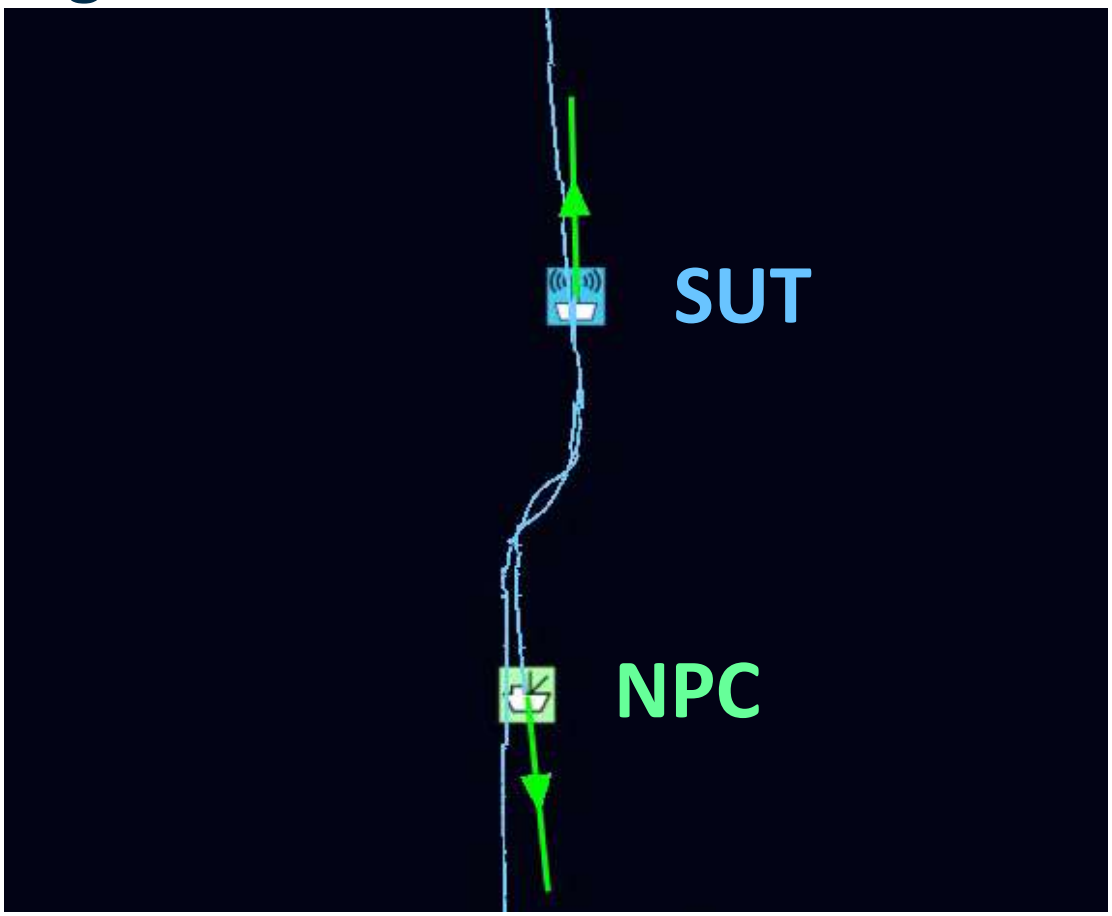
Lessons Learned

- **This SUT will try to pass on the right, even if it is the Stand-On vessel**



Lessons Learned

- Once again, SUT will try to pass on the right, even if it is the Stand-On vessel
- NPC using similar (flawed) COLREG algorithm so also passes on right



DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited

AVIA Conclusion

Georgia Tech Research Institute

- **Achievements for the AVIA program include:**
  - Bounded performance for desired scenarios
  - Performed 1000 1-hour scenarios within 16 hours
  - Identified errant action and attributed to either errant perception or errant behavior (or errant simulation)
  - Confirmed stochastic nature of autonomy performance
- **Applicable to other programs**
  - Scalable to massive parallelization on high performance computers (HPCs)
  - Architecture supports SUTs, simulations, metrics plugins
  - AVIA scenario controller -> sensor, physics, scenario plugins

ANY DOMAIN - HETEROGENEOUS SUTs - REGRESSION TESTING

DISTRIBUTION STATEMENT A – Approved for public release, distribution unlimited