# PART 1

FLORIDA POLY

## PROJECT NARRATIVE AND TECHNICAL APPROACH

# COVER PAGE

| SUMMARY TABLE | |
|---|---|
| Project Name/Title | Florida Poly Verification Methodology |
| Eligible Entity Applying to Receive Federal Funding | Florida Polytechnic University<br>4700 Research Way<br>Lakeland, FL 33805 |
| Point of Contact | Rahul Razdan Phd.<br>Director Advanced Mobility Institute<br>rrazdan@floridapoly.edu<br>(386) 943-5360 |
| Proposed Location (State(s) and Municipalities) for the Demonstration | Seaside FL (30A) |
| Proposed Technologies for the Demonstration | Level 3+ Autonomous Shuttle<br>Abstract Simulation<br>Abstract Test Generation<br>Aerospace/DoD/FAA High-fidelity simulation |
| Proposed Duration of the Demonstration | 10/1/2019 - 7/1/2022 |
| Federal Funding Amount Requested | $8,929,562.81 |
| Non-Federal Cost Share Amount Proposed | $0.00 |
| Total Project Cost | $8,929,562.81 |

# Project Narrative and Technical Approach

## 1. EXECUTIVE SUMMARY

Autonomous Vehicle (AV) technology has the potential to fundamentally transform the automotive industry, reorient transportation infrastructure, and significantly impact the energy sector. Rapid progress is being made in the core artificial intelligence engines which form the basis of AV technology. However, without a quantum leap in the test, verification and simulation approaches and systems the autonomous vehicle industry is using, the full capabilities of AV technology will not be realized, and regulators will not have the tools to install the safety apparatus needed for broad-based AV proliferation.

Critical features of the test, verification, simulation and regulatory regime must be a clear model upon which one can reason about the operation, testing, and validation of AV systems. Further, this model must be able to connect to the physical world directly through a feedback process from real accidents and the state space must be understood sufficiently to get to a notion of completeness. All of this must be done in a manner which maximizes safety and builds confidence in the public.

We observe that the current state has none of these characteristics. The current commercial solutions are using ad-hoc methods such as miles driven to provide some indication of validation, but no fundamental structure has been offered to demonstrate the robustness of the solutions. In fact, there is nothing which assures that all the various software updates are actually adding to safety. Finally, the use of "real world" testing through "shadow and safety" driving is highly suspect both from the point-of-view of verification convergence and public safety.

Figure 1: Seaside

At Florida Poly, we have developed a Florida Poly AV Verification Framework (FLPolyVF) as a methodology to enable rule-making for AVs. A center point of this framework is a model for a “scenario” which we call FLPolySA (Scenario Abstraction). FLPolySA allows for a mathematical language to specify AV test environment, so allowing regulators a language to clearly communicate with OEMs. Using FLPolySA, FLPolyVF can build an environment which can build a model for completeness, a flow for constant update from physical feedback, and drive the test generation process for “edge” test cases.

To prove the utility of this framework, we propose a demonstration project consisting of using this methodology on a very specific physical demonstration project at Seaside, FL. Seaside, Florida, described by *Time* magazine as “the most astonishing design achievement of its era and, one might hope, the most influential.” As the birthplace of a movement in land planning known as the new urbanism, Seaside’s influence has spread widely and is helping to revolutionize town planning in America. Seaside has won numerous awards for its architecture and town planning and has been the subject of three books and countless articles. However, Seaside’s success has led to large congestion issues on 30A the highway, and would like build the next level of a “new urbanism,” by **providing AV shuttle capabilities for the visitors, residents and workforce along 30A**. Because of the slower speeds and relatively low complexity of the 30A environment, Seaside offers an ideal starting point to hash out and prove a verification methodology.

Our proposed demonstration project would have three stages. First, the 30A shuttle route and environment would be modeled in the FLPoly Verification Flow. The intention at this stage would be to build a framework for coverage and generate representative “edge” test cases. Second, critical aspects of the 30A environment will be modeled in a full-motion real-time simulation modeled in the Dactle system. The intention of this system is to safely test as realistically as possible the ego vehicle. FPoly will generate the test cases to be run in the Dactle environment. The final stage will be a physical demonstration on 30A. Beyond verifying functionality, key aspects of this demonstration will be to instrument the shuttle in a manner to detect test conditions which were not examined in simulation. This very critical feedback process creates a situation to debug the test and validation system.

The most important result of this demonstration will be to prove an AV test and verification methodology which can then be subsequently scaled to other environments. With success, the current ad-hoc and ineffective methods of verification using “shadow and safety” driving can be retired.

Florida Polytechnic University, a state public university, will be the primary agency entering into an agreement with the USDOT. Key partners will be Dactle and Nova technologies who will provide the detailed physical simulation capability, and ITIC (International Transportation Innovation Center) who will support the physical testing and deployment of AV shuttle.

## 2. AV Verification Methodology Structure

In the recent years, AVs have attracted great attention from academia, industries, and governments. Perception, decision making, and action are the three major processes required for driving a vehicle. Currently, decision-making and perception are controlled by human drivers, and the action process is performed by the vehicles.

According to a report of the National Highway Traffic Safety Administration (NHTSA), 94 percent of the 37,461 traffic fatalities in 2016 were due to human error. AVs are designed to conduct the decision making and perception aspects of driving, and it is hoped that will reduce accidents related to human error. In addition, studies also show that autonomous driving technologies can positively impact economy, safety, and traffic congestion. Despite all these advantages, the major barrier for wide-scale adoption of AVs is the test and verification regime to assure safety. To address this barrier, a process, which builds an engineering argument for assuring safety, must be developed. Typically, this argument is built based on the following principles:

- Conceptual model: A conceptual understanding of the problem is built and supported through virtual models.
- Test Regime: Using the conceptual model, a test regime is built to test the model and build an argument for correctness.
- Completeness: The state space of tests is examined within the modeling environment to develop metrics for completeness.
- Accumulative Learning: A structure is constructed where field testing feeds back into this flow such that safety is always rising.

Intertwined with the above methodology is the classic V paradigm model which is used as a mechanism to enable concurrent design and test. In this paradigm, mathematical models, which have been correlated with a bottom-up component level characterization stage, are used early in the design stage. As the design is refined, physical components can be substituted to a point when system level tests can be performed on the whole physical design. Modeling issues are often corrected with a virtual to physical diagnostics flow. The combination of the conceptual safety regime and the V design process have been effectively used to build robust safe systems in many domains.

In fact, the above flow has been used very successfully by the automotive industry to verify conventional cars for many years. However, the addition of the perception and decision making has added an order-of-magnitude level of complexity to solving the safety problem. Critical open issues can be listed as follows:

- Conceptual model: What are the conceptual models, which are appropriate for the perception and decision making stages of AV operation?
- Test Regime: What is the test regime, which can build confidence as to the operation of the AV?
- Completeness: How do we understand the state space sufficiently to understand the risks surrounding completeness? How do we address the issues of tester bias?
- Accumulative Learning: How do we know the next version of the vehicle or even software update is increasing safety?

We observe that today none of the above questions is answered for AVs. The current commercial solutions are using ad-hoc methods such as miles driven [kalra2016driving] to provide some indication of safety, but no fundamental structure has been offered to demonstrate the robustness of AV products. Further, without the answers to above mentioned questions, regulators do not have the means to address safety issues or even to communicate clearly safety issues to operators or the public.

## 2.1. Characteristics of the Solution

### 2.1.1. Hardware Design

We observe that the realm of AV verification has quite a few similarities as compared to complex hardware verification. In the realm of hardware, millions of extremely complex components (transistors) are assembled together to form a higher-level function which is embodied in a semiconductor chip. The cost of the development of these semiconductor chips is very high and simulating the whole chip at the physics level is impossible. Even at the highest level of abstraction, most semi-conductor chips can only be simulated in the low kilo-hertz range while the chips themselves run in the giga-hertz range. Thus, in any design project, a very limited simulation budget exists (a day of real-world operation) to run all the tests to assure safe operation for the lifetime of the part.

What are the analogies to AV verification? First, much like hardware, AVs consist of complex components (sensors, object recognition systems, radar, etc). Simulating everything at the most detailed level is similarly impossible. Second, much like hardware, AVs need to compress 18 years of human traffic learning into a reasonable development cycle. Third, AVs have the same need for robustness relative to environmental conditions, and finally, they have the same need for completeness and accumulated learning.

World class hardware verification teams solve these problems with a variety of approaches. The first and most important of these is the use of abstraction as a powerful tool to decompose the problem. Within hardware, various abstraction levels have been developed (transistor, gate, RTL, micro-architecture, architecture, network stack) which separate concerns and build an inductive proof for verifying the whole semiconductor chip.

As an example, the transistor design team focus entirely on the task of making sure the semi-conductor physics process produces a behavior consistent with a transistor under all environmental conditions. A cell designer relies on this behavior to build larger components and only verifies that the combination works as expected. Thus, via this recursive process, a chip is built and verified. These separation of concerns and abstractions are so powerful that whole multi-billion-dollar markets (fabless, ASIC) have been built based on these concepts.

This inductive process is very effective in demonstrating equivalence between abstraction levels, but this still leaves the verification of the highest level of abstraction. In this area, hardware verification has used a variety of techniques such as formal verification, constrained-random test generation, and real-world test injection to model and test the overall function [kropf2013] introduction, melham1988abstraction}. Finally, a deep concept of coverage analysis exists in order to model completeness. The combination of all of the above has created an environment where most semi-conductor chips are typically functional on the first pass of manufacturing despite the enormous complexity and size.

How can AV verification use the powerful methods developed for hardware verification? The key is the development of an enabling abstraction level.

### 2.1.2. AV Framework

In this section, we introduce a key enabling abstraction approach to aid in the task of AV verification, which we will term FLPOLY Scenario Abstraction (FLPolySA). FLPolySA has the following high-level characteristics:

- Wireframe/Building Block: Components are very simple recto-linear objects which model the physical characteristics of the environment.
- Dynamic and Static: There are two types of objects. Static objects which do not move and dynamic objects which move with a pre-determined vector. These components are not responsive.
- Assertions: Both the dynamic and static components contain meta-data which assert expected behavior. Example: A static component such as a Stop-sign might assert that all cars approaching it must stop.
- Newtonian Physics: The behavior of the components follows the rules of simple Newtonian physics. A critical idea which is modeled is the notion that mass with velocity/acceleration/gravity will lead to the expected behavior.
- Unit Under Test (UUT): An ego car can enter this test and has the potential to achieve success or failure if none of the error assertions fire.

The precise details of the model will be explained in succeeding sections. However, at this point, it is important to motivate the design of FLPolySA. There are many powerful consequences for choosing this higher-level simple abstraction structure:

- Mathematical language: The abstraction allows for a clear capture of the modeling environment and provides a basis to reason about this environment.
- Extended Coverage: A single abstract model contains within it many underlying combinations. As an example, the black-box car could be used to model vehicles from any brand.
- System Coverage: A signature process for the whole model can be used to drive a test coverage process. A signature process is required to understand what scenarios have been examined earlier and when a scenario is indeed new.
- Separation of Concerns: The abstraction allows the sensor/object recognition problem to be addressed independently from the decision-making problem.

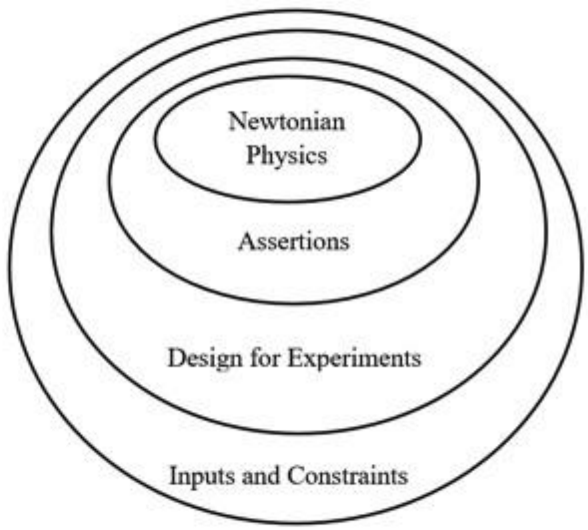



Figure 2: Conceptual Abstraction Layer

Overall, FLPolySA helps address many of the critical issues mentioned in the introduction as missing aspects for AV verification. Figure 2 shows the conceptual layering of this abstraction approach.

- Newtonian Physics: The function of this layer is to model movement, momentum, and detect collisions. This is the physical world.
- Assertions: Layered over the physical world is the idea of good or bad.

- Design for Experiments: This environment is setup such that the test has no memory and thus is repeatable.
- Inputs and Constraints: Layered in the outermost layer is the machinery for inputs and constraints which drive a singular test. This machinery is setup to be able to run constrained pseudo-random tests and collect ongoing coverage.

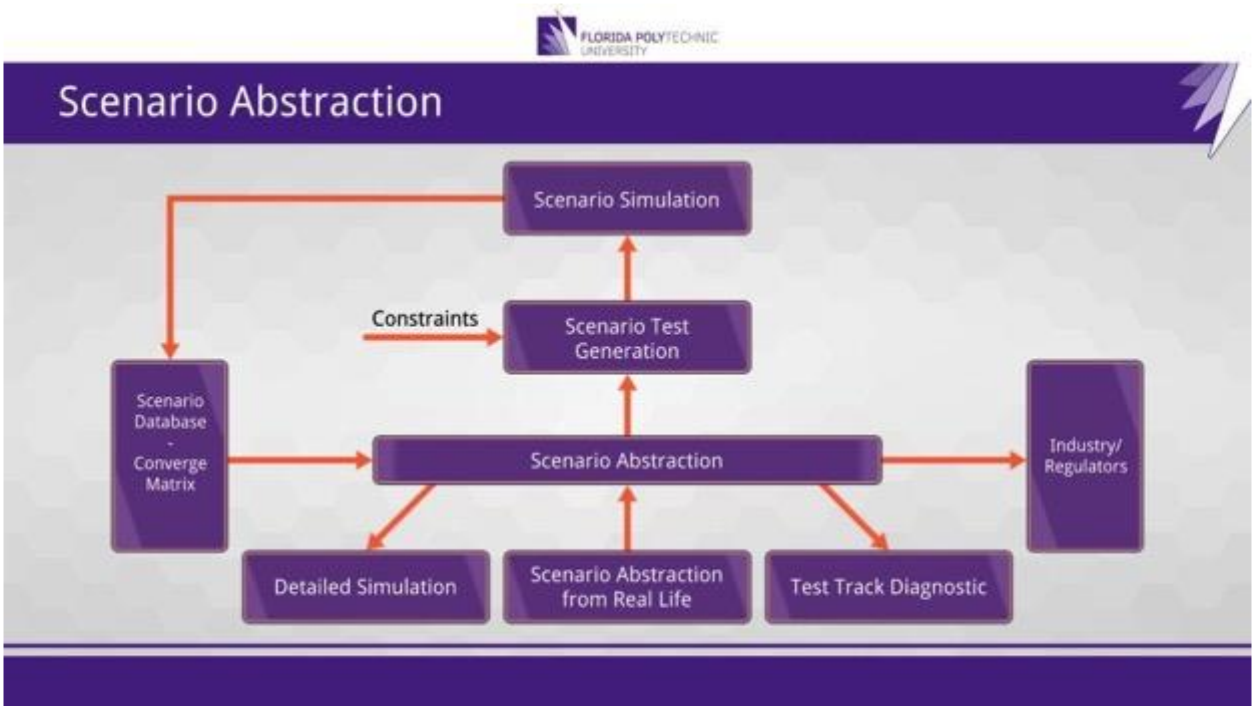



Figure 3: FLPolyVF

With the introduction of FLPolySA, an overall architecture can be developed for FLPolyVF. Figure 3 shows the blocks and flows among the blocks in this environment. The critical pieces include:

- Scenario Test Generation: Using a seed and a constraints matrix, pseudo random experiments can be generated and tested automatically against the unit under test (UUT). The objective of this part of the flow is to simulate millions of configurations in simulation and try to find test cases which cause ego-car to fail. This part of the flow requires the mathematics to understand input constraints and enable a signature flow for deciding whether a test has been generated earlier.
- Scenario Database/Coverage: The notion of scenario coverage is critical in FLPolyVF. This is the method which can be used as a basis for regulatory approval, limit redundant simulation, and form the basis of an ongoing test and verification database. The scenario abstraction provides an excellent method to capture this information and further the defined mathematics has a strong concept of equivalence classes which allows for a deeper optimization of the coverage database.

- Test Track Diagnostic: Once a test fails, there is a need to diagnose the test in a deeper level of abstraction. This can be more detailed simulation models all the way to a physical test track environment, and FLPolyVF allows for this path through a synthesis process from the scenario test to a physical instantiation.
- Sensor Test and Verification: A critical part of an AV is the object recognition system and sensors. The scenario abstraction can provide test cases with built-in criteria of success such that the sensor/object recognition problem can be verified separately.
- Scenario Abstraction: Just as there is a need for synthesis from the scenario level, there is a need to abstract the scenario construct from the physical environment. A classic example is a ``real-world" accident which must be analyzed much more completely in the simulation framework. In addition, it is common to test for ``cousin" bugs with this flow.
- Industry/Regulators Communication: There is a need to communicate in an unambiguous manner among the industry participants and the regulators. The FLPolyVF scenario abstraction provides an excellent means for this communication.

In the context of this project, the 30A test environment will be modeled in the FLPolyVF with the objective of building edge-case edge conditions, building a coverage framework for 30A test routes, and a framework for accumulative learning. As the test cases are produced by the FLPolyVF system, they will need to be tested, and the preferred safe method of doing so is a much more detailed PHYSICS based simulation environment. Finally, after confidence has been built within the simulation environments, controlled testing will be done on 30A corridor. A very important part of this controlled testing will be a feedback process from the physical shuttle to detect unanalyzed scenarios in the simulation environment.

## 3. DACTLE PHYSICS Simulation:

Dactle provides the following capabilities leveraging aerospace/DoD/FAA Simulation Technology the current Autonomous Vehicle industry does not utilize:

1) High fidelity, with accurate Newtonian physics, vehicle models which are easily adaptable
2) High fidelity, with accurate performance and anomalies included, sensor simulations
3) High fidelity digital representation and processing of the area to verify the vehicle performance

To support these properties Dactle has developed a construct which allows for a modular simulation architecture which allows for functionality to be adjusted with either the interchange of software modules or data configuration files. This allows for the adjustment within a scenario to be completed by either exchanging data files or selecting which capability model is executed. This allows for a single baseline to be utilized for passenger, light utility, truck and larger vehicles by the user. The use of aerospace/DoD/FAA simulation technology resolves the significant capability gaps the simulation systems in the Av industry utilize. Technology gaps that will cause false confidence, improperly trained ML systems and eventual real-world problems. These gaps involve real-time and model precision issues. Those models including vehicles, tires, roads, sensors and the environment. This approach will permit the replacement of 99.9% of the public shadow and safety driving with simulation. This solution will significantly reduce cost and time to market and avoid any more tragedies such as those Tesla and Uber have experienced.

The core capability is supported by a high-fidelity representation in software of the vehicle under test. This would include the equations of motion of the body of the vehicle (taking in the mass momentum, center of gravity and control reaction capabilities, etc.) that presents the AV management stack the accurate reactive capability of the vehicle. The vehicle is defined in two forms. The standardized equations of motion and the data which represents the response curves the equations of motion need to respond to in real time. The generalized form of the vehicle data flow is as shown in Figure 4.

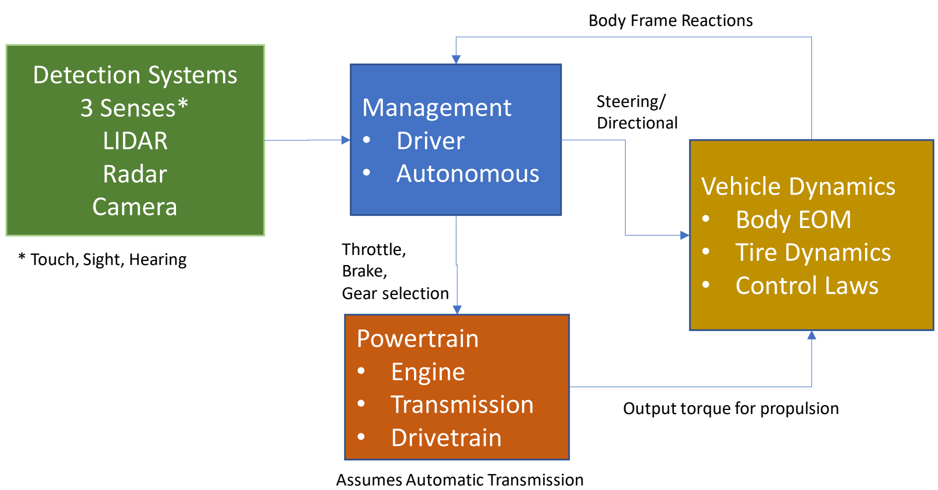



*Figure 4 - General Vehicle Model*

The data which is required to be defined to support this concept must be obtained from a representative vehicle of type. The vehicle under test must be instrumented for data gathering of critical aspects such as weight, center of gravity, acceleration and deceleration rates (both brake on and brake off), control response (input to body axis reaction times and magnitudes), and tire characterization. The generalization of the simulation of the powertrain is as shown in Figure 5, with the annotation of the data types developed for each segment.

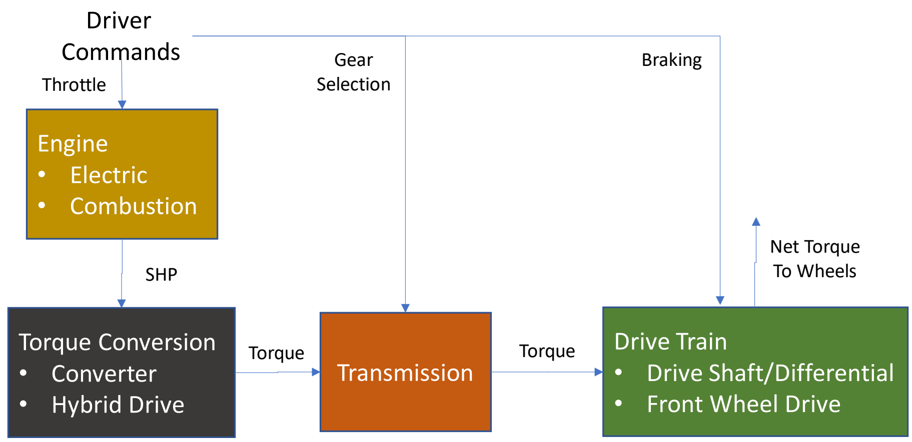
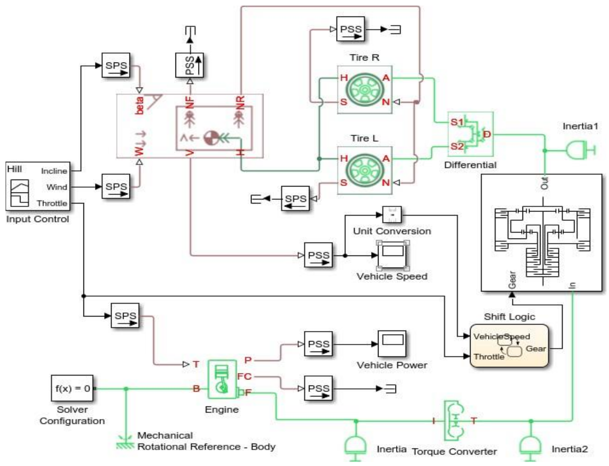



*Figure 5 - Simplified Propulsion Concept*

To support this process, Dactle is proposing the inclusion of the instrumenting and data gathering process for a vehicle of the type to be used on the road testing (as noted in the prior sections) to gather the data, providing the baseline of data which will support the simulation. The data set required will be limited in nature, due to both time and funding restrictions, however limiting the area and vehicle under test to 30A removes the need for gathering data which would be considered "winter" test cases.

This process is proposed to be completed in a static environment (closed course) and is expected to be completed within 60 days of activation. This process will involve the data gathering under specific test matrix (coming from FLPolyVF), data reduction for post gathering to confirm a reasonable appearance of the data, and execution of additional testing to gather additional data for either data collection issue correction (outliers or inconsistent data cases). Once verified, the data will be integrated with the vehicular model to confirm proper operation with the Automatic Testing Guide (ATG) which is an element of the Dactle Simulation.

The ATG functionality is a management process of executing sequential test cases with the simulation that presents a data initialization to the simulation, manages the execution of the simulation (start/stop) and the data gathering mechanization for later review/comparison to the expected results. The synchronization and data management process is the key to successful execution of the testing process. Working in concert with the scenario management data, the ATG is the autonomous component of the simulation testing process, with recursive testing capability and the ability to manage a rapid validation of the AV component as it is developed and revalidated. The structural difference between the Dactle approach and the standard laboratory approach is the use of discrete software algorithms instead of collaborated, packaged math models in a standardized form. The representation of a standard laboratory approach using SimuLink or Matlab will look as the depiction in Figure 6.




*The abstraction of this is that a component (engine, transmission, etc.) is represented in the model as a widget (as shown in*

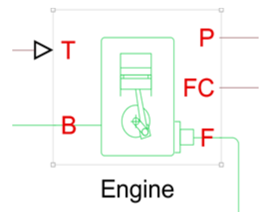



*Figure 6 - Representative Vehicle Model in Simulink*

*Figure 7 - Graphical Widget For Modeling*

The same component in the Dactle solution would be a discrete level of software capability which is data driven for engine specifications and speed/torque curves which would be obtained from the component manufacturer. The data tables which the Dactle model has contained within it form the seed data from which all other performance is derived. The effects of the environment (temperature, humidity, etc.) along with the interaction with the surface the vehicle is travelling upon is all factored into the performance model coded into the system. The method provides a much better management structure to account for anomalies in the simulation as well as variations between models of vehicles (reducing the cost of development for a new vehicle type).

The second leg of the Dactle approach is a high-fidelity representation of sensor data processing and the presentation to either the AV control system or the driver for decision making. The key element of the ability of the Autonomous Vehicle to make safe control determination is that it has solid situational awareness of the events and elements within distance of the vehicle that potentially affects the vehicle. These items include things such as other vehicles, pedestrians, animals, moving objects in the path of the vehicle (ball crossing the path, objects protruding from other vehicles, etc.), and traffic management elements (lights, signs, etc.). The Dactle approach is to provide a very high-fidelity digital representation of the environment in which the vehicle will be traversing and an accurate performance-based simulation of the sensors which will be interrogating that environment. Figure 8 contains a screen capture of the level of data which we believe this represents.

Given there are no “aids” incorporated in the National Transportation System as of yet for electronic identification of any of these items (no signals from lights, stop signs, other vehicles, etc.) the entire recognition system is a combination of active and passive systems. Passive systems are typically video enabled, as well as G sensing systems to obtain body axis reactions due to either commanded movement or road roughness intake. Active systems include those of

Radio Detection and Ranging (RADAR) or Light (Imaging) Detection and Ranging (LIDAR) systems. These systems project a signal on specific frequency, with a specific period and then interprets the return the signal has when it reflects off of a surface. This gives primarily a distance and a direction of an object around the vehicle. These sensors are then considered by the AV electronics (similarly to the determination made by a human driver) to define the control inputs which are necessary to the vehicle. This can be an avoidance maneuver, acceleration or deceleration, or a directional change to a new path.

The ability of this process to be successful is the combination of the high-fidelity environmental representation of the area under test along with a simulation of sensors which resolves to the same level of discrimination as the real sensors. The sensors have limitations, restrictions and performance boundaries which must be taken into consideration when the simulation is employed. The placement of the sensors is a critical issue with the ability of the sensor to properly detect and discriminate the environment completely to aid in the decision processing. The quantity of sensor inputs as well as the fidelity of the data being presented to the AV management electronics needs to be an exact replication to give the same events and decision processes to be executed as would be done in the "Real world". Figure 9 contains a representation of what a LIDAR sensor would reflect to the sensor intake for the vehicle.

*Figure 8 - High Fidelity Environment for Simulation*

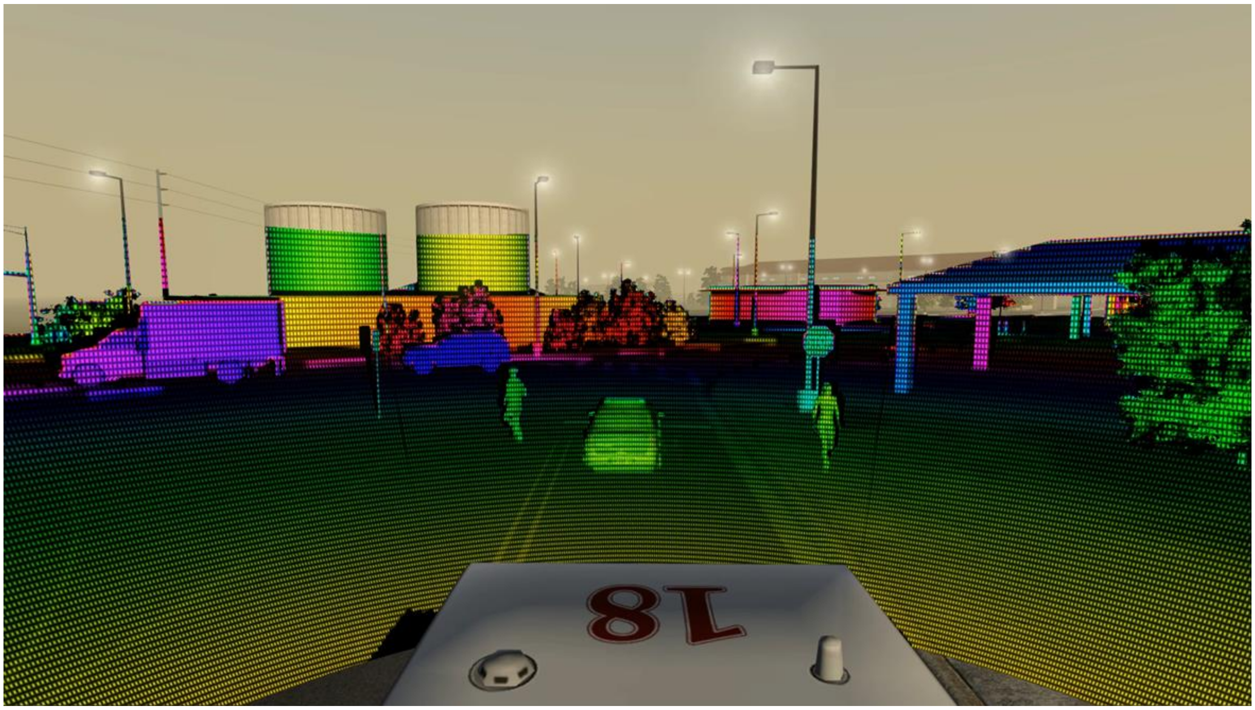


*Figure 9 - Sensor Representation through Simulation*

The collaboration of the sensor simulation and the high-fidelity environment provide a basis for evaluating the AVs ability to discern the decisions necessary to navigate the environment it is working in.  The sensor simulation is an essential component, combined with the vehicle simulation to achieve correct measure of the decision loop for latency, accuracy and completeness assessments.  Without both of those elements being exact representations of the system under test the results are at best an approximation of performance, not giving that level of confidence which is critical to determination of completeness in capability.

To coordinate all of this technology to a common goal of a structured testing process is the Scenario management system which supports a structured execution system for verification and the processing of the ATG system in the simulation.  The design of the scenario management system is driven by the ability to support a variety of vehicles, under a variety of conditions with a range of testing dynamics within each parametric. The system must also be capable of supporting a specific singularity of a scenario to support accident investigation or a specific data case which must be examined.  The general user concept of the types of data which must be defined in the scenario is shown in Figure 10.

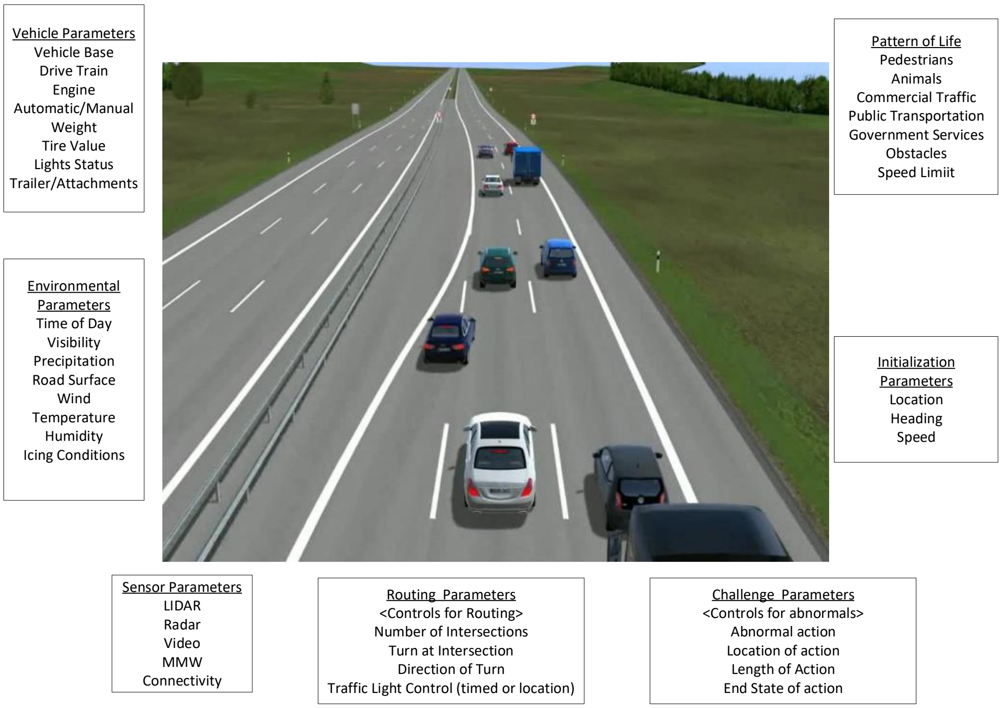



*Figure 10 - Scenario Data Typing and Controls*

The scenario management understructure (driven by FLPolyVF) is to provide a basis for the establishment of the flow of testing parametrics to determine specific case and elements of the validation matrix. That development of the validation matrix is the key to providing a standardized method of assessment for the Autonomous Vehicle component set. Using this standardized process also allows to reflect extreme cases under test as well as the establishment of the persistent standard of success for an AV management set. The flow process for the establishment of the scenario is as show in Figure 11. Using these three components collaborated into a real time matrix of validation will allow for the confirmation that the AV components will be able to manage the vehicle in a competent and safe manner under a set of varying environments.

There are number of levels of fidelity which are available to the simulated event process. The fidelity of the simulation models, the sensor models and the level of validation of the scenario process all affect the outcomes. One of the major components of this process is the physical device utilized in the process as well. Figure 12 reflects the possible options for physical fidelity of the simulation process.

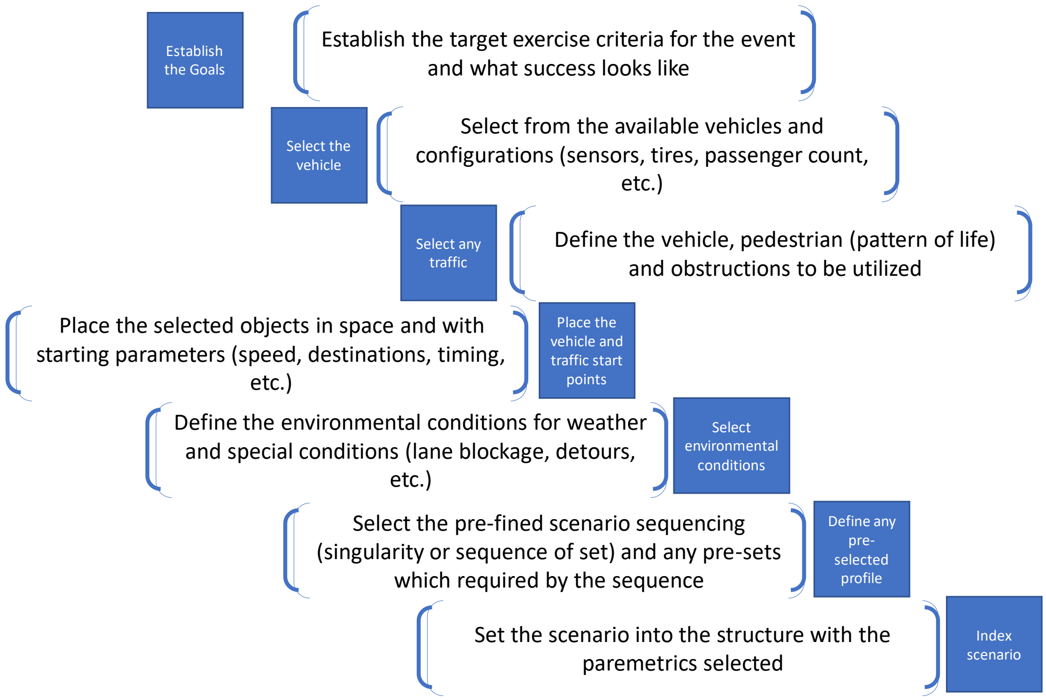



*Figure 11 - Scenario Sequencing Method*

*Figure 12 - Examples of physical simulation platforms*

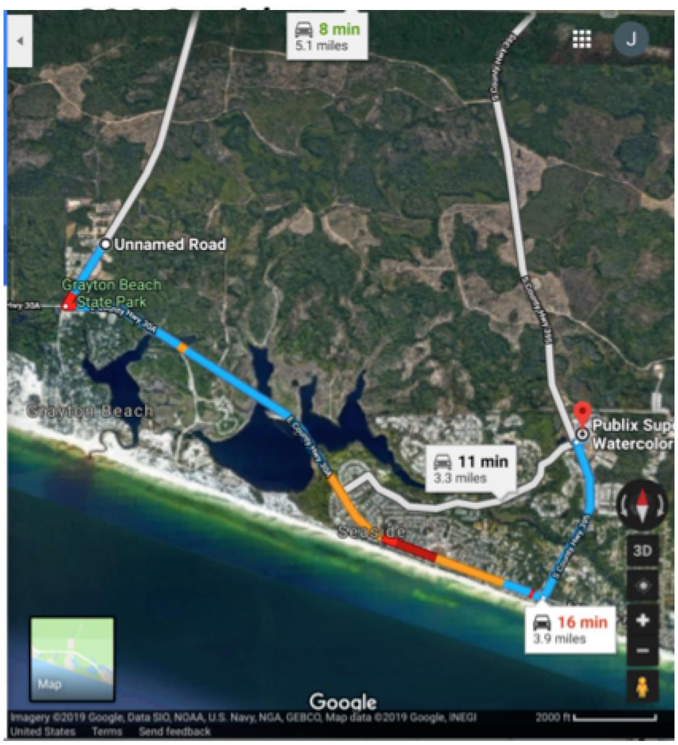
The left panel is the simplest, with a gaming level control system and a simple panel display. This configuration is good for task-based learning but not for reflecting any motion cues (such as abrupt starts, stops or turns). That level of fidelity begins to get reflected in the center panel where a more immersive experience is delivered with a larger visualization as well as a more complete user cab. This cab includes a small cabin and a much higher fidelity control system with a motion cuing system for seat movement. The most complex system (as well as the most expensive) is the one on the right (this is NADS-1 at the University of Iowa) which involves a 360-degree visualization, a large motion base and the most advanced physical environment in the country.

These have varying levels of cost associated with them. For the purposes of this grant Dactle has chosen a system consistent with the middle level of fidelity, which should be enough to reach the level of feedback intended for the effort. This level of fidelity is more than sufficient to validate the management and control of an autonomous vehicle as well as the affects of the vehicles maneuvers on the occupants. The lowest level of fidelity provides no feedback, and the highest level of fidelity has an exorbitant purchase and support cost, nowhere near the value which would be returned to any study or experiment for autonomous vehicles. The simulation hardware will be a residual asset for use in later evaluations and experiments if desired.

## 4. 30A Physical Experiment:

**Motivation:**

Seaside and the 30A corridor are plagued with massive traffic issues. This leads to safety, economic, and lifestyle consequences which are not positive. The goal at 30A is to develop an on-demand AV shuttle system which can co-exist with other traffic participants via fixed routes considering the use of dedicated AV lanes. The biggest concern to operate an AV shuttle system along 30A is linked to safety. At peak season, many tourists are populating the beaches which leads to high traffic density on the roads as well as spontaneous road crossings of pedestrians over unmarked sections.




Although the 30A corridor stretches over an area of 20 miles, we will implement an AV shuttle pilot route over 3.9 miles between Grayton Beach and Seaside to study in detail AV vehicle interaction scenarios from a safety perspective as well as the traffic impact of using dedicates AV lanes along CR30A.The suggested test route starts at a County park lot at Grayton Beach at CR283 and ends at the parking lot of the Publix Super Market in Seagrove Beach at CR395. CR283 is a main feeder road into CR30A. The 30A section between the end points passes Watercolor and Seaside which are considered the most attractive beach communities along 30A.

Figure 14: Google Maps of 30A

A key strategic goal of the Seaside founder Robert Davis for 30A is to incentivize tourists, residents and service workers not to use their private vehicle but to use on-demand mobility services such as AV shuttles. This means to motivate people to leave their vehicles at remote parking lots and to switch the transportation mode as they come closer to the beaches.

Here are the key locations along the suggested pilot route:


Grayton Beach parking lot

Water Color Inn

Seaside Bud and Aley’s

Publix at Seagrove Beach

The envisioned implementation is planned to be a dedicated lane either in the middle of the street as shown in the figure below or extended lanes on the side which reuse existing bike paths. As highly automated vehicles are fully controlled at all times, they do not require the same lane width as traditional manually driven vehicles. Large SUV’s and Pick-Up trucks can be up to 7ft wide whereas smaller AV shuttles such as the one show below is just about 5ft wide.

Given the specific circumstances of this test pilot, the simulation systems will fully model the relevant pieces. These include:

1) Definition of relevant driving scenarios in the interaction between the AV shuttle and its environment along the route.
2) Analysis which sections along the pilot route are safety critical.
3) Analysis the dependency of the identified safety risk level from road infrastructure layout, third party sensor input and communication infrastructure layout.
4) Simulation of critical safety risk scenarios and identification of improvement options (vehicle side, infrastructure side) with simulated data
5) Use of collected data through a physical AV test vehicle and installed sensors along the route to validate safety simulations and to validate suggested improvement measures.

Based on the findings of the simulation and the validation of the simulation model through physical experiments it should be possible to build an enhanced model to address the following topics:

> Operation of a fleet of AV vehicles both on single AV lane and double AV lane layout (up to 4 vehicles to be operated in parallel)
> Operation of a fleet of AV vehicles on dedicated AV lane only versus operation in mixed traffic (partially or fully) and its impact on traffic flow as well as operational safety risk

Dedicated uni-directional AV lane (Source: Seaside Institute)

What is important to note that along the suggested 30A AV shuttle pilot route a single low speed AV lane integrated in a shared space with bicycles and pedestrians separated from the main road is in principle feasible. A dedicated double AV lane would require a change of the existing road layout.

AV Shuttle (source: Local Motors)

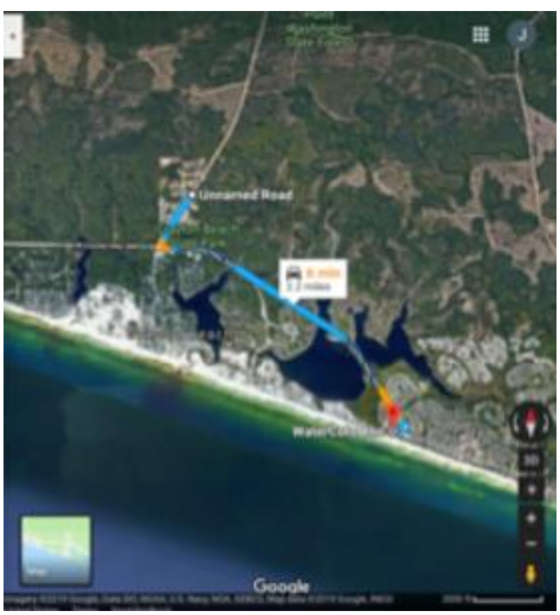
The following use cases will be considered for both the simulated as well as the actual experiments:

A) People movement:
- Use case 1: drop-off and pick-up of people at dedicated AV shuttle stations
- Use case 2: impact of different vehicle dynamics scenarios (e.g. hard braking) on people inside the shuttle while AV shuttle is moving at different speed levels (assuming cruising speed around 12-15 mph and maximum speed at 25 mph) – both considering normal operation as well as emergency maneuvers.
- Use case 3: impact of different AV vehicle driving scenarios on other traffic participants (reaction pattern analysis, e.g. interaction with bicycles drivers or golf karts)

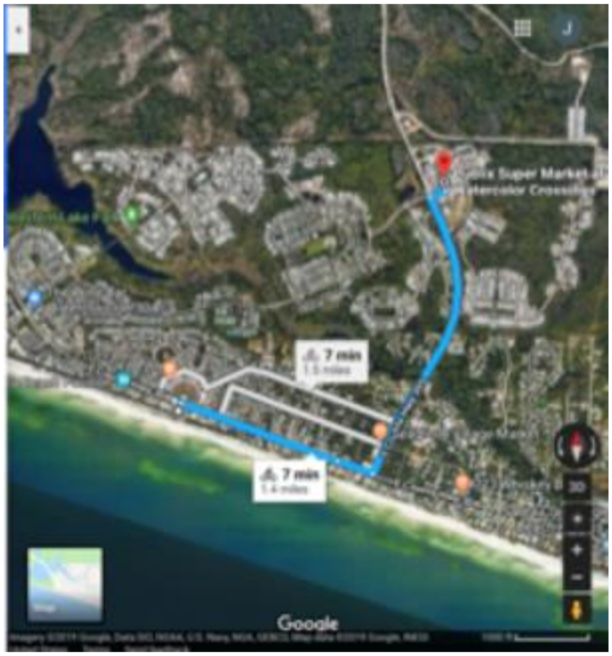
B) goods movement
- Use case 1: loading and unloading of goods into the vehicle at dedicated AV shuttle stations
- Use case 2: point to point movement of goods along fixed route without intermediate stops

The 30A AV shuttle pilot route can be sub segmented into three sections as follows:

Section one (2.2 miles long) from Grayton Beach County Parking Lot to Water Color Inn (on left). Within this section are several bridges which do not allow separate lanes. Section two (0.5 miles long) is from Water Color Inn to Seaside Bud & Aley's (on right). This section has a very high risk of heavy traffic during peak season and has parking lanes on both sides of the road.




Section three (1.4 miles long) is from Bud & Aley's to Publix at Seagrove Beach. Within this section is a complicated T-bone traffic intersection (CR30A/CR395). Prior to this intersection are complicated non-symmetrical road conditions in terms of parking and sidewalks, after the intersection is a dedicated dedicated bike lane leading to the Publix super market.




From a safety perspective all three sections need to be carefully explored in terms of potential road layout configurations and vehicle/bicycle/pedestrian configurations. It needs to be determined how a low-speed AV shuttle operation can be implemented both in the existing road configuration and lane layouts as well as how future road modifications and lane configurations should look like. Ultimately a solution needs to be found with the lowest safety risk level with reasonable improvement in travel time in highly congested traffic situations. Long term the goal is to shift capacity from private car use to shared AV shuttle use and to keep private cars parked outside the beach areas. Furthermore goods transportation such as luggage delivery and food delivery should also be shifted to AV shuttles as far as possible.

# 5. FOCUS AREAS

**Significant Public Benefit:** The significant public benefit from this project is that it provides a methodology (FLPolyVF) for the regulation of autonomous vehicles. This framework combines simulation of various forms with a physical demonstration. In addition, it builds a feedback process for evaluation of simulation correctness. Finally, this methodology provides a method for regulators to take real-world crashes and move them into the virtual domain for further analysis. The pilot combines simulation methodologies, test generation methodologies, and a well understood physical implementation (at 30A) to demonstrate the viability of the FLPolyVF. With success, this methodology can be generalized by DOT for use throughout the country.

**Addressing Market Failures:** To build an effective verification structure upon which one can have confidence, one must have a conceptual model upon which to reason, a test regime around this conceptual model, some sense of completeness, and some method for accumulative learning. We observe that today none of the above is true. The current commercial solutions use ad-hoc methods such as miles driven to provide some indication of safety, but no fundamental structure to show robustness of their solutions.

**Economic Vitality:** Without a regulatory structure and technical verification structure, it is not possible for autonomous vehicle technology to reach its promise of access, cost, and convenience.

**Complexity of Technology:** The verification structure presented focuses on AV solutions which are SAE level 3 or higher. In addition, the techniques presented can be generalized to a broader set of AI/mobility solutions (marine, drone, etc).

**Diversity of Projects:** This project focuses on rural Walton county which can have urban style traffic patterns during peak season. Walking, biking, (privately organized) transit, and driving are all modes of transportation represented in these areas.

**Transportation-Challenged Populations:** As a rural area without public transportation, this project will enable the transportation for visitors, residents as well as workers who must travel to or along 30A.

**Prototypes:** The key prototype produced from this work will be a working physical demonstration and associated virtual infrastructure used for regulatory purposes. This consists of:

1. FLPolySA: High level simulation models of the 30A environment with associated test generation engines.
2. Dactile: Detailed high-fidelity simulation environments to recreate problematic scenarios.
3. 30A Transit: A physical transit demonstration with built-in feedback systems to #1 and #2.

# 6. REQUIREMENTS

**Focus on ADS:** The focus areas of verification infrastructure, testing methodology and physical demonstration are all focused ADS.

**Demonstration:** There will be a physical demonstration connected with two levels of simulation models.

**Data Exchange:** All the data for the project will be hosted on AWS cloud storage with access given to DOT.

**User Interface:** The physical demonstration will be an automated shuttle with a safety driver.

**Scalability and Agency Outreach:** The results of this demonstration will be open-sourced and the FloridaPoly will publish several papers on the lessons learned from this demonstration.

# 7. APPROACH

**Implement and Evaluate the Demonstration**: Our approach takes the real problem of Autonomous Vehicle safety verification of a real situation at 30A (seaside) and builds a safety verification flow for that situation. In this process, the implementation will be the generation and execution of test cases. The demonstration of success will be the ability of the shuttle to operate error-free or the set of tests which expose failure.

**Legal, regulatory, environmental, and/or other**: To the best of our knowledge, the demonstrations do not require any exception to Florida Statue, Federal Motor Vehicle Safety Standards (FMVSS), Federal Motor Carrier Safety Regulations (FMCSR), or any other regulation. In terms of the buy America provisions, all the components will be American made except for the Full-motion simulator [xxxx] where we would seek an exception.

**Commitment to provide data and participate in safety evaluations**: As stated in *Section 4.3 Data Exchange*, this project is committed to sharing all the available data. We intend to open source the methodology and publish materials in academic journals.

**Risk Identification, Mitigation, and Management**: Before demonstrations occur, a risk management plan will be developed.

**Contribution and Management of non-Federal resources (cost share)**: There is no cost share.

## 8. Schedule and Major Deliverables:

| | | | |
|---|---|---|---|
| | Year 1 | Year 2 | Year 3 |
| AMI Goal | Abstract Simulation Model of 30A Route | Maintain and Update AMI model based on changes to route/shuttle | Maintain and Update AMI model based on changes to route/shuttle |
| AMI Goal | Coverage Matrix For 30A Route | Update Coverage Matrix For 30A Route/shuttle combination | Update Coverage Matrix For 30A Route/shuttle combination |
| AMI Goal | Prioritized List of Test Scenarios for Edge Conditions | Generate Prioritized List of Test Scenarios for Edge Conditions | Generate Prioritized List of Test Scenarios for Edge Conditions |
| | | | |
| DACTLE Goal | Receive/Install Ansible Simulator | Maintain Ansible Simulator | Maintain Ansible Simulator |
| DACTLE Goal | Characterize shuttle and 2-3 likely scenarios | Grow Scenarios based on physical and AMI input. | Grow Scenarios based on physical and AMI input. |
| DACTLE Goal | Build software bridge with AMI team for future scenario generation. | | |
| | | | |
| ITIC Goal | Receive and "install" first Local Motors shuttle | Receive and "install" 2 additional Local Motors shuttles | Work with government to build transit authority |
| ITIC Goal | Support Characterization activity for simulation environments. | Build safety plan for a restricted pilot | Enable the capture of data for analysis by AMI and Dactle teams. |
| ITIC Goal | | Enable the capture of data for analysis by AMI and Dactle teams. | |
| | | | |
| | | | |