



DEVCOM DAC-TR-2023-048  
June 2023

---

# **Adapting U.S. Army Acquisition to Ensure the Reliability and Safety of Autonomous Vehicles**

by Patiana Theragene, Martin Wayne, and Nathan Herbert

### **DISCLAIMER**

The findings in this report are not to be construed as an official Department of the Army position unless so specified by other official documentation.

### **WARNING**

Information and data contained in this document are based on the input available at the time of preparation.

### **TRADE NAMES**

The use of trade names in this report does not constitute an official endorsement or approval of the use of such commercial hardware or software. The report may not be cited for purposes of advertisement.



DEVCOM DAC-TR-2023-048  
June 2023

---

# Adapting U.S. Army Acquisition to Ensure the Reliability and Safety of Autonomous Vehicles

by Patiana Theragene, Martin Wayne, and Nathan Herbert  
*DEVCOM Analysis Center*

## REPORT DOCUMENTATION PAGE

<b>1. REPORT DATE</b>		<b>2. REPORT TYPE</b>		<b>3. DATES COVERED</b>	
June 2023		Technical Report		<b>START DATE</b> 10/27/2022	<b>END DATE</b> 03/13/2023
<b>4. TITLE AND SUBTITLE</b> Adapting U.S. Army Acquisition to Ensure the Reliability and Safety of Autonomous Vehicles					
<b>5a. CONTRACT NUMBER</b>		<b>5b. GRANT NUMBER</b>		<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>5d. PROJECT NUMBER</b>		<b>5e. TASK NUMBER</b>		<b>5f. WORK UNIT NUMBER</b>	
<b>6. AUTHOR(S)</b> Patiana Theragene, Martin Wayne, and Nathan Herbert					
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Director DEVCOM Analysis Center ATTN: FCDD-DAS-L 6896 Mauchly Street Aberdeen Proving Ground, MD 21005				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>  DEVCOM DAC-TR-2023-048	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>			<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b> DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b> This report presents several challenges that the U.S. Army will face in the transition to autonomous vehicles—challenges that are only magnified in the current acquisition environment with limited testing. Artificial intelligence algorithms introduce additional complexity, resulting in systems with a complex combination of human, machine, and autonomous controllers. The acquisition process for these autonomous vehicles will require a different approach to ensure that these systems are safe and reliable. Employing a “systems view” that appropriately considers the complexity of the interactions, along with a combination of currently available tools and techniques early in the development of these systems, is vital to the success of the programs. This will enable the Army to efficiently acquire these capabilities while also providing assurance that they will function as intended.					
<b>15. SUBJECT TERMS</b> autonomous vehicles, AI, artificial intelligence, reliability, testing, risk assessment					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>		<b>18. NUMBER OF PAGES</b>
<b>a. REPORT</b> UNCLASSIFIED	<b>b. ABSTRACT</b> UNCLASSIFIED	<b>c. THIS PAGE</b> UNCLASSIFIED	UU		19
<b>19a. NAME OF RESPONSIBLE PERSON</b> Patiana Theragene				<b>19b. PHONE NUMBER (Include area code)</b> (410) 278-9261	

**STANDARD FORM 298 (REV. 5/2020)**

*Prescribed by ANSI Std. Z39.18*

---

---

## (U) Table of Contents

List of Figures .....	iv
1. INTRODUCTION .....	1
2. CURRENT STATE OF TESTING IN THE ARMY .....	3
3. IMPACTS OF AUTONOMY .....	4
4. ACCOMODATING THE COMPLEXITY OF AI.....	6
5. CONCLUSION AND FUTURE WORK.....	9
6. REFERENCES .....	10
List of Acronyms .....	12
Distribution List .....	13

---

---

## List of Figures

Figure 1.	Sample Control Structure of Autonomous System.....	8
-----------	--	---

---

---

# 1. INTRODUCTION

Artificial intelligence (AI) has become prevalent in many fields in the modern world, ranging from vacuum cleaners to lawn mowers and commercial automobiles. These capabilities are continuing to evolve and become a part of more products and systems every day, with numerous potential benefits to humans. AI is of particular interest in autonomous vehicles (AVs), where the benefits include reduced cognitive workload, increased efficiency, and improved safety for human operators.

Numerous investments from academia and industry have been made recently with the intent of improving the enabling technologies for AVs. Google and Tesla are two of the more well-known examples in industry, with Google developing a self-driving car and Tesla providing its Full Self-Driving (FSD) autopilot system. Ford and BMW are also working on their own AVs. These systems continue to remain a challenge despite these investments.

Recent National Highway Traffic Safety Administration (NHTSA) investigations involving Tesla highlight the importance and difficulty of these technologies. NHTSA Campaign Number 23V085000 (NHSTA, 2023a) involves the recall of over 300,000 vehicles equipped with the FSD Beta system. The recall summary states that the FSD Beta system may result in unsafe actions with the vehicle. These unsafe actions include traveling straight through an intersection while in a turn-only lane, entering a STOP-sign-controlled intersection without coming to a complete stop, and proceeding without caution into an intersection during a steady yellow traffic signal. The recall also states that the system may fail to respond to changes in posted speed limits and not account for the driver's adjustment of the vehicle's speed beyond these limits.

NHTSA Campaign Number 23V037000 (NHSTA, 2023b) was an earlier recall involving approximately 53,000 vehicles that also discusses issues with the FSD Beta system. The system included a "rolling stop" functionality in the software that could allow the vehicle to travel through an all-way stop intersection without stopping. Both recalls were able to be addressed through remote Over-the-Air software updates, which is actually an important benefit that AVs will likely have.

A recent survey of deep learning methods for AVs also recognized the difficulties that remain for widespread use of AVs. The work identified 10 challenges that must be resolved: 1) Complexity of Autonomous Driving (AD) Systems; 2) Dynamicity of Road Environment; 3) Big Data and Real-Time Processing; 4) Intelligent Data Prioritization; 5) Robustness and Adaptability; 6) Integration/Fusion of Sensory Data for Dynamic Decision Making; 7) Fairness, Accountability, and Transparency in deep learning for

---

---

AD; 8) Online Learning Capabilities in AD; 9) Robustness Against Adversarial Attacks; and 10) Variability of Traffic Sign Boards (Muhammad et al., 2021).

According to NHTSA guidance, the overall safety of AVs is left to the companies that build them. Each company must comply with Federal Motor Vehicle Safety Standards while also certifying that their vehicles do not have inherent safety risks. While many companies are currently testing vehicles with higher levels of automation to ensure that they operate as intended, the NHSTA indicates that numerous experts state that more work remains to be done by vehicle developers to ensure their safe operation before they are available commercially (NHSTA, n.d.). Detailed studies on the algorithmic methods for safe autonomous driving have not yet been completed. These studies should be considered the backbone of the safety of AVs, including those developed for use by the Army.

Considering the issues raised here, the Army must be prepared to adapt its current strategies for development and acquisition of these systems. They must ensure that the systems are beneficial to the units that receive them and do not lead to operational vulnerabilities or unforeseen safety hazards. Before adapting any acquisition strategies though, it is useful to consider the fundamental limitations of AI and how they relate to systems of interest to the Army. AI systems are often designed to function in complex environments with varying levels of uncertainty, but the algorithms are not able to achieve flawless performance. Therefore, systems that are technically reliable may still be expected to occasionally produce abnormal results. This point may be counter to the desired performance of AVs, but it helps to emphasize the importance of approaching the development and acquisition process carefully. These systems are no longer just machines or human-machine interactions. They are instead complex combinations of the human, machine, and autonomous controller. Any approach to the development and acquisition of these systems must consider this complex combination as its foundation.

The rest of this report is organized as follows. Section 2 presents an overview of the current state of testing within the U.S. Army. Section 3 discusses the various impacts of autonomy and the requirements that it would bring under the current acquisition approach. Section 4 provides alternatives to the current acquisition approach that should be considered for AVs, and Section 5 contains conclusions and recommendations for future work.

---

---

## 2. CURRENT STATE OF TESTING IN THE ARMY

Currently fielded Army vehicles do not have high levels of autonomy. Much of the developmental testing of these vehicles focuses on the reliability and durability of their mechanical and electromechanical subsystems. The testing is typically conducted on Army-approved test tracks that cover the terrains expected in-theater.

In the past, Army programs had substantial amounts of developmental testing, culminating with an operational test scoped to verify compliance with reliability requirements. This entire body of evidence would support production and fielding decisions. However, the draft update of the test and evaluation policy Army Regulation 73-1 states that operational tests are no longer scoped for reliability data collection. It says that developmental tests will be used to collect the data needed to show compliance with reliability requirements (U.S. Department of the Army, 2022).

At the same time, the Army is moving toward more agile forms of acquisition. The goal is to field systems in a timely manner while keeping pace with the speed of technological innovation. DoD Instruction 5000.02 outlines several acquisition pathways that support programs moving faster (Under Secretary of Defense for Acquisition, Technology, and Logistics, 2022). By compressing the acquisition timeline, it likely decreases the opportunities for developmental testing. This presents a challenge to programs as they are forced to say more with less.

As future Army vehicles incorporate more autonomy and elements of AI, the scope of reliability testing will need to expand accordingly, not decrease. In addition to stressing the mechanical elements, testing will also need to help ensure autonomy and AI algorithms in the vehicles are correctly tuned to support safe operations in expected and even unexpected environments and scenarios. Prior to fielding these AVs, the Army will need to preserve adequate amounts of developmental testing, limiting how fast these systems can be acquired.

---

---

### 3. IMPACTS OF AUTONOMY

It is not unwarranted that AVs could lessen, or even eliminate, the issues that occur with human driving. However, the complexities of autonomy bring about a unique set of problems that require a rigorous amount of testing. We begin by studying the standards of automation. According to SAE J3016 (Society of Automotive Engineers, 2019), there are six levels of automation: levels 0–2 have driver support features with a human in control, ranging from little to no automation, and levels 3–5 have automated driving features with the system in control, the highest level being fully autonomous. Note that different parts of the driving system can be at different levels of automation. For example, a vehicle may have lane centering (level 2 feature) but may not have a steering wheel (level 4 feature). In addition, the more-advanced-level features in the vehicle, the more the driver will have to maintain situational awareness, as there is an expectation, not a requirement, for the human driver to be ready for manual takeover in levels 3–5. Thus, the changes in the role of the human driver with each automation level will affect the creation of AV safety standards. Current safety standards are limiting: They build off standards created for non-autonomous vehicles and require a full understanding of the system (Joint Technical Committee, 2020) whereas AI systems, not just AVs, are dynamic and perpetually updating with new information.

When we consider the actual physical testing of AVs, accumulating high mileage is one of the most crucial factors. The California Department of Motor Vehicles (CA DMV) recently released their 2020-2021 disengagement and mileage reports for 1175 vehicles tested in autonomous mode (CA DMV, 2022). Disengagement is defined as “a deactivation of the autonomous mode when a failure of the autonomous technology is detected or when the safe operation of the vehicle requires that the AV test driver disengage the autonomous mode and take immediate manual control of the vehicle.” On average, each vehicle tested for about 3500 miles and had 2–3 disengagements. In total, there were over 4 million miles driven and 2,676 disengagements reported. Consider those numbers compared to usual Army testing miles for non-autonomous vehicles, which is typically significantly less.

Furthermore, recent work from the RAND Corporation found that fully AVs would need to drive millions of miles to ensure reliability in terms of fatalities and injuries. They also stated that for the autonomous failure rate to be statistically significantly lower than the human driver failure rate, it would require several billion miles of driving. For example, it would take roughly 5 billion miles to demonstrate that the autonomous failure rate was 20% lower than the assumed human failure rate of 1.09 reported fatalities per 100 million miles (Kalra & Paddock, 2016).

---

---

When looking at injuries, a test for comparing two Poisson means can be used to design a test with no failures assuming the human injury rate of 77 reported injuries per 100 million miles (Nelson, 1982). The test indicates that even in the unlikely event of having no failures, we would still need to test these vehicles for close to 5 million miles. This becomes a tougher issue when considering the influence of dynamic environments and the need to test across the spectrum of intended operation.

For autonomous systems, sustained testing and interactions with a dynamic environment are critical to the reliability of the system. Designing systems that are developed and tested in static and structured environments, and then having them perform as needed in changing and unstructured environments, however, is no small feat. The intricacies of the physical environment render the space of system inputs/outputs and environmental variables combinatorically intractable. Existing test and evaluation practices do not have the needed fidelity to deal with this emergent behavior or the built-in capacity to learn without human supervision. Testing for the unpredictable is an arduous, imperfect process as can be evidenced by the reasons for disengagement in the CA DMV report. For example, some disengagements were issued by the test driver due to the system moving too slow or too fast for given traffic conditions despite going the “correct” speed according to road signs. Other disengagements were issued because the vehicle was out of the operational design domain (CA DMV, 2022). These two reasons portray the subtleties of AVs that may not always be accounted for in testing. If we wish to incorporate autonomous systems into everyday operations, it is vital to include these nuances of autonomy into the reliability testing of these vehicles.

---

---

## 4. ACCOMODATING THE COMPLEXITY OF AI

The AVs will introduce new complex failure modes involving interactions between multiple components within the broader vehicle system. Certain failure modes will be difficult to detect, even with substantial amounts of testing. This highlights the need for additional approaches to manage the risk. As discussed in *Reliability Assurance for AI Systems* (Blood et al., 2023), many of the failure modes and risks associated with autonomous systems can be identified and mitigated using rigorous Design-for-Reliability (DfR) techniques. These systems have significantly higher complexity than traditional systems that lack autonomous capabilities, but they are still likely to experience many commonly occurring failure modes that can be mitigated with traditional methods such as Fault Tree Analysis, Failure Mode Effects and Criticality Analysis (FMECA), and Physics of Failure modeling.

Simulation, where it is used to supplement physical testing, is an additional approach worth considering. A combination of modeling and simulation, lab testing, and full-system-level testing was recommended in *Testing in a Complex World* (LaRose et al., 2019) as part of a crawl–walk–run approach for Army ground vehicles. The crawl stage involved traditional DfR techniques, while the walk stage was a combination of digital simulation and physical simulation in the laboratory. The run stage was built on the previous work and consisted of full-system-level testing in operationally relevant environments. Carruth et al. (2022) conducted simulated and real-life experiments of an autonomous ground vehicle in off-road and less-than-ideal environments to find the impacts this had on the perception, control, and planning subsystems. Different algorithms were tested, and they compared the effects of hard surface and soft soil on the vehicle. They found that the simulation matched the real tests when it came to trajectories and speeds for both types of soil, though it overpredicted the speed for soft soil. Nonetheless, the failure rates matched within a 90% confidence interval.

Investigators at Kostas Research Institute for Homeland Security from Northeastern University used a modified “Grand Theft Auto V” video game to help simulate off-road environments and a simulation platform to test and evaluate model performance on real versus simulated data sets (Singh et al., 2023). In addition, they used a style transfer to help close the domain gap between simulated and real environments. These approaches helped determine which algorithms and lidar methods were more accurate. Of course, simulation does have its limitations. But simulation and physical testing complement each other, as they can further help identify problems that may have been missed when looking at each part separately.

Another potentially powerful approach for understanding and mitigating risks associated with AVs uses a systems-view of the AV. The “systems-view” considers the complexity

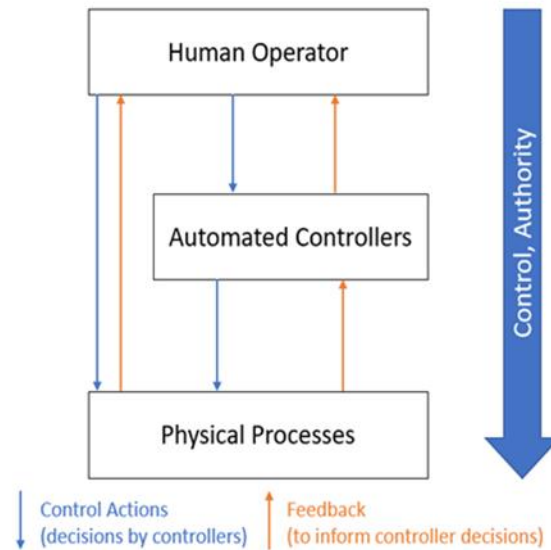
---

---

of the entire AV, including the human operator, physical hardware, software, and autonomous controller. System-Theoretic Process Analysis (STPA) is an approach that employs such a systems-view to further understand complex systems.

STPA is a hazard analysis technique based on an accident causation model. It assumes that accidents, or undesirable outcomes, are caused not only by component failures, but also through unsafe interactions between fully functioning components (Levenson & Thomas, 2018). It also presents numerous advantages over standard probability risk assessment and other techniques. The analysis can be started very early in system development and easily integrated with more traditional DfR activities. It can also document functionality in complex systems that is often difficult to understand through traditional approaches. This provides numerous benefits throughout early design evolution, as more complex “unknown unknowns” can be identified and mitigated earlier in development (Levenson & Thomas, 2018). Comparisons of STPA to traditional methods such as fault trees and FMECAs show that STPA finds the same causal scenarios as those approaches, along with other scenarios that are often software-based, where no actual failure has occurred. STPA also includes both software and human operators. This helps to ensure that the analysis includes all relevant causal factors, while also providing a direct connection to failure modes associated with the human–computer interface.

STPA examines potential losses and hazards by analyzing a system’s control structure. A sample control structure for an autonomous system is shown in Figure 1. In this basic control structure, the downward arrows illustrate control actions and the upward arrows illustrate feedback paths. In general, control actions are based on the mission and feedback received.



**Figure 1. Sample Control Structure of Autonomous System**

Systems with autonomy are ideal candidates for STPA, as the autonomy introduces a new level of complexity. For example, in the sample control structure, both the human operator and automated controller are controlling the physical processes. What if there is a conflict in control actions between the two? What if different feedback is presented to each? How does the system resolve this conflict? STPA provides a methodical approach to work through undesirable control actions and the conditions or scenarios that can cause them.

In the end, STPA will identify system requirements and constraints that will help mitigate potential failure modes and undesirable interactions between system components. It will also inform efficient test and evaluation strategies to be used during development.

---

---

## **5. CONCLUSION AND FUTURE WORK**

This report presents several challenges that the Army will face in the transition to AVs and which are only magnified in the current acquisition environment with limited testing. The complexity introduced by the combination of human, machine, and autonomous controller in AVs will require a different approach to development and acquisition to ensure that these systems are safe and reliable. Employing a “systems-view” and a combination of currently available tools and techniques early in the development of these systems will enable the Army to efficiently acquire these capabilities while also providing assurance that they will function as intended.

Additional work is necessary to ensure the safe and reliable use of AVs, and the many current challenges must be addressed by a combined partnership of the Army, industry, and partners in academia. Detailed studies on the underlying AI algorithms enabling autonomy are necessary, along with a systematic approach to analyzing and/or mitigating the risks of these algorithms within the broader systems in which they exist.

---

---

## 6. REFERENCES

- Blood, J., Herbert, N., & Wayne, M. (2023) *Reliability assurance for AI systems*. Proceedings of the Reliability and Maintainability Symposium (RAMS).
- [CA DMV] California Department of Motor Vehicles. (2022). *2021 Autonomous vehicle disengagement & mileage reports*. [Online]. <https://www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/disengagement-reports/>
- Carruth, D. W., Goodin, C., Dabbiru, L., Scherrer, N., Moore, M. N., Hudson, C. H., Cagle, L. D., & Jayakumar, P. (2022, August 4). *Measuring the influence of environmental factors on obstacle detection and avoidance with an autonomous ground vehicle*. Center for Advanced Vehicle Systems, Mississippi State University [Online]. <https://apps.dtic.mil/sti/citations/AD1177696>
- Joint Technical Committee ISO/IEC JTC 1, Information Technology, Subcommittee SC 7, Software and Systems Engineering. (2020, November). *Software and systems engineering – software testing – Part 11: guidelines on the testing of AI-based systems*. ISO/IEC. [Online]. <https://www.iso.org/obp/ui/#iso:std:iso-iec:tr:29119:-11:ed-1:v1:en>
- Kalra, N., & Paddock, S. (2016, September 23) *Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?* (RR-1478-RC). RAND Corporation.
- LaRose, B., Morgan, M., Stark, K., Kosinski, D., Fischer, K., Brudnak, M., Schultz, G., Christino, A., Wayne, M., & Baseski, I. (2019, August). *Testing in a complex world*. Proceedings of the Ground Vehicle Systems Engineering and Technology Symposium (GVSETS), NDIA.
- Levenson, N., & Thomas, J. (2018, March). *STPA handbook*. Boston, Massachusetts.
- Muhammad, K., Ullah, A., Lloret, J., Del Ser, J., de Albuquerque, V. (2021, July). *Deep learning for safe autonomous driving: current challenges and future directions*. IEEE Transactions on Intelligent Transportation Systems, 22, 7.
- Nelson, W. (1982) *Applied life data analysis*. Wiley & Sons.
- [NHTSA] National Highway Traffic Safety Administration. (n.d.). *Automated vehicles for safety*. <https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety>
- [NHTSA] National Highway Traffic Safety Administration. (2023a, February 15). *NHTSA Campaign Number 23V085000*. <https://www.nhtsa.gov/vehicle/2022/TESLA/MODEL%2525203/4%252520DR/RWD#recalls>

---

---

[NHTSA] National Highway Traffic Safety Administration. (2023b, January 27). *NHTSA Campaign Number 23V037000*. <https://www.nhtsa.gov/vehicle/2022/TESLA/MODEL%2525203/4%252520DR/RWD#recalls>

Singh, H., Ostadabbas, S., Mortin, E., Kaveti, P., Bai, X., Luo, Y., Gupta, A., & Jiang, L. (2023, February 2). *MYO 6: Assessing autonomous maneuver & mobility, project 6.2: Scoping requirements for autonomy simulation*. Kostas Research Institute for Homeland Security at Northeastern University, Proceedings of DEVCOM Analysis Center AI/AA Analytical Framework Cooperative Agreement Year 1 Technical Reviews.

Society of Automotive Engineers. (2019, January 7). *SAE J3016: Levels of driving automation*. <https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic>

Under Secretary of Defense for Acquisition, Technology, and Logistics. (2022). *DoD Instruction 5000.02, operation of the adaptive acquisition framework*.

U.S. Department of the Army. (2022). *US Army Regulation 73-1, test and evaluation policy, 2022 draft*.

---

---

## LIST OF ACRONYMS

AD	autonomous driving
AI	artificial intelligence
AV	autonomous vehicle
CA DMV	California Department of Motor Vehicles
DfR	Design for Reliability
FMECA	Failure Mode Effects and Criticality Analysis
FSD	Full Self-Driving
FTA	Fault Tree Analysis
NHSTA	National Highway Traffic Safety Administration
STPA	System-Theoretic Process Analysis

---

---

## DISTRIBUTION LIST

DEVCOM Analysis Center  
FCDD-DAS-L/N. Herbert  
FCDD-DAS-L/P. Theragene  
FCDD-DAS-L/M. Wayne  
6896 Mauchly St.  
Aberdeen Proving Ground, MD 21005-5071

DEVCOM Army Research Laboratory  
FCDD-RLB-CI/Tech Library  
2800 Powder Mill Rd.  
Adelphi, MD 20783

Defense Technical Information Center  
ATTN: DTIC-O  
8725 John J. Kingman Rd.  
Fort Belvoir, VA 22060-6218