

DOI: [https://doi.org/10.48009/3\\_iis\\_2024\\_136](https://doi.org/10.48009/3_iis_2024_136)

## Cross-disciplinary threats mitigation framework for AI autonomous vehicles

**Gaston Elongha**, *Marymount University, gle85144@marymount.edu*

**Mohammed Almuthaybiri**, *Marymount University, mma84343@marymount.edu*

**Innocent Mgembe**, *Marymount University, ipm87755@marymount.edu*

### Abstract

Despite the benefits such as road mobility and capacity enlargement, mortality lowering, and emissions drop, AI Autonomous Vehicles (AI-AVs) have introduced emerging threats that paired with cybersecurity threats, are changing the cybersecurity threat landscape. AI-AVs can be used as a moving intelligence collection for espionage and surveillance. They can be weaponized and operated for physical (logistical and transport) or furtive cyberattacks. AI-AVs have also introduced concerns about data privacy, governance, and ethical considerations. Therefore, it is crucial for both a domestic and foreign AI-AVs cross-disciplinary mitigation framework to address the threats introduced by AI-AVs. This study emphasizes solutions to AI-AV threats and addresses them through the cross-disciplinary framework mitigation approach that prepares society, nations, and individuals for ongoing transportation advancements, the rise of AI applications, and their threats despite the revolutionary benefits. The researchers proposed a cross-disciplinary approach that converges the strategic and resilient approach, including the transport department road safety administration, and all the entities involved in the coalition such as AI-AV manufacturers, transport stakeholders, lawmakers, advocates for safety, academia, local, and foreign entities.

**Keywords:** autonomous vehicles, artificial intelligence, cross-disciplinary, threats mitigation

### Introduction

In this era of outstanding and ongoing technological advancements, Artificial Intelligence (AI) applications have expanded to almost any sector, not just transport in the case of AVs, but also healthcare, education, aviation, and so on. The capability and limitations that could not be outsmarted or extended in the past are now improved and extended to the level that one day could outsmart humans. Any sector nowadays has some sort of AI in the background that runs either the process, predicts outcomes, or simply makes the driving decision. AI Autonomous vehicles (AI-AVs) are regarded as sustainable mobility technology as they are eco-friendly and help reduce collision and traffic obstruction, expand mobility, and road capacity, improving supply chain and logistic systems (Li et al., 2019). Despite those benefits, there are tremendous risks, security, and unintended ethical and privacy considerations (Taeihagh & Lim, 2018). The driving motions have changed from no-to-partial automation (level 0-2) to fully automated (0-5) (Society of Automotive Engineers International, 2014). Artificial Intelligence (AI) is now powering the full automation driving level for making judgments (Cunneen et al., 2019; Martinho et al., 2021). AI-AVs have also introduced concerns about data privacy and governance (Chen et al., 2020), moral and ethical concerns.

The US Department of Transport (USDOT) along with its annex, the National Highway Traffic Safety Administration (NHTSA) has made a tremendous effort to ensure the safety guidance plans for AVs, going

from the ADS 2.0, AV 3.0, and AV 4.0. The first plan was the Automated Driving Systems (ADS 2.0) published in September 2017 as guidance visionary plans for safety. The following one was Automated Vehicles 3.0 (AV 3.0), which pitched preparation for future AVs, and lastly, Automated Vehicles 4.0 (AV 4.0), like other guidance plans not only focused on safety but aimed to ensure self-sufficiency in AVs as well as the leadership in automated vehicle technologies (USDOT, n.d). The USDOT with the NHTSA also developed the Federal Autonomous Vehicles Policy published in September 2016 which has a framework for AVs' performance guidance and scope.

While reviewing the Federal AVs policy framework, we found that the scope includes all entities involved in AVs' sales, testing, deployments, and development. The guidance states the areas where the AV manufacturers must focus to ensure safety that satisfies the framework's scope (USDOT, n.d). The issue with the Federal AVs policy and USDOT safety guidelines is that they are **not mandatory**. They were developed for compliance with the Federal Motor Vehicles Safety Standards (FMVSS). The framework focuses more on safety on the roads and vehicle occupants. Autonomous vehicle manufacturers, nowadays, AI-AVS, only comply with the framework to show how AI-AVs test, deployment, and design include safety and are compliant with conventional FMVSS standards. It is unclear if the NHTSA will make certain framework areas mandatory. Meanwhile, domestic, and foreign AV manufacturers are deploying AI-AVs that only comply with conventional FMVSS standards, opening the door for new cybersecurity and emerging technology threats as the deployed AI-AVs include more functionalities that must be addressed through a mitigation framework.

The mitigation framework is important to help define, and address areas overlooked in the AI-AV's safety. Manufacturers may comply with the USDOT and NHTSA's Federal AVs policy framework for vehicle performance, but the framework is not mandatory, it is important to have a mitigation framework that takes on safety from AI and cybersecurity threats introduced by the testing, deployment, and development of AI-AVs. The mitigation framework will be beneficial to support the safety assessed in the framework for vehicle performance guidance by adding a threats mitigation framework from cybersecurity and AI perspectives. This study is critical for preparing society, nations, and individuals for this continuous transport and rise of AI applications and threats when embedded in AVs. The purpose of this study is to develop a cross-disciplinary approach that converges the strategic and resilient approach, including the transport department road safety administration, and all the entities involved in the coalition such as AI-AV manufacturers, transport stakeholders, lawmakers, advocates for safety, academia, local, and foreign entities.

### Literature Review

AI Autonomous Vehicles (AI-AVs) rely on different layers to navigate the roads. Priscila et al. (2022) highlighted how AI-AVs heavily depend on perception, network, and application layers to navigate the traffic. They showed how each layer also relies on different other devices, arguing that the perception layer relies on cameras, sensors, navigation & location data. They also highlighted how the network layer relies on broadband, satellite, and mobile communication devices. Finally, they explained how the application layer depends on the human interface, stores, processes, and handles all data. Different experiments have been conducted in past literature reviews that demonstrated successful cyber-attacks on these layers. In addition, the devices used in each layer are also susceptible to vulnerabilities that can be exploited and help threat actors take over the AI-AVs. We reviewed and grouped attacks, vulnerabilities, and experiments from current and past literature into *perception*, *network*, and *application* layers to understand AI AV technology and potential threats that will help develop the mitigation framework.

## Perception Layer

Many studies have proved the AI-AV's vulnerabilities and threats through the perception layers. The first was introduced and known by Petit et al. (2021) in experiments in sensors exploiting MobileEye camera vulnerabilities. They exploited particularly the MobileEye C2-270 vulnerabilities and LiDAR (the ibeo LUX 3). A similar experiment by Bhupathiraju et al. (2023) on LiDAR Sensors also proved how AVs' perception layer is vulnerable. They proved that targeting the LiDAR sensor's time of circuits (TOF) with adversarial interference due to an electromagnetic signal can make the AI-AVs misbehave affecting the perception layer to improperly categorize things, see inexistent ones and unreal obstacles (Bhupathiraju et al., 2023). MobileEye camera is the most used camera in AI-AVs worldwide, with 40 million pieces loaded in AI-AVs (Povolny et al., 2020). McAfee Advanced Threat Research (ATR) conducted eighteen months of research focusing on digital attacks at the perception layer. They exploited the MobileEye vulnerabilities in the black box known as adversarial Machine Learning (AML) and the white box. They were able to confuse the classifier with false negatives and positives when detecting objects and speed signs (Povolny et al., 2020). Eykholt et al., (2017) proved through a physical attack through stop signs with paints and stickers. They were able to confuse the Tesla X MobileEye camera to misclassify 35 mph to 85 mph and the stop sign to an added lane and 35 mph to 45mph (Okubo & The Hoffman Agency, 2020; Povolny & Fralick, 2020).

## Network & Application Layers

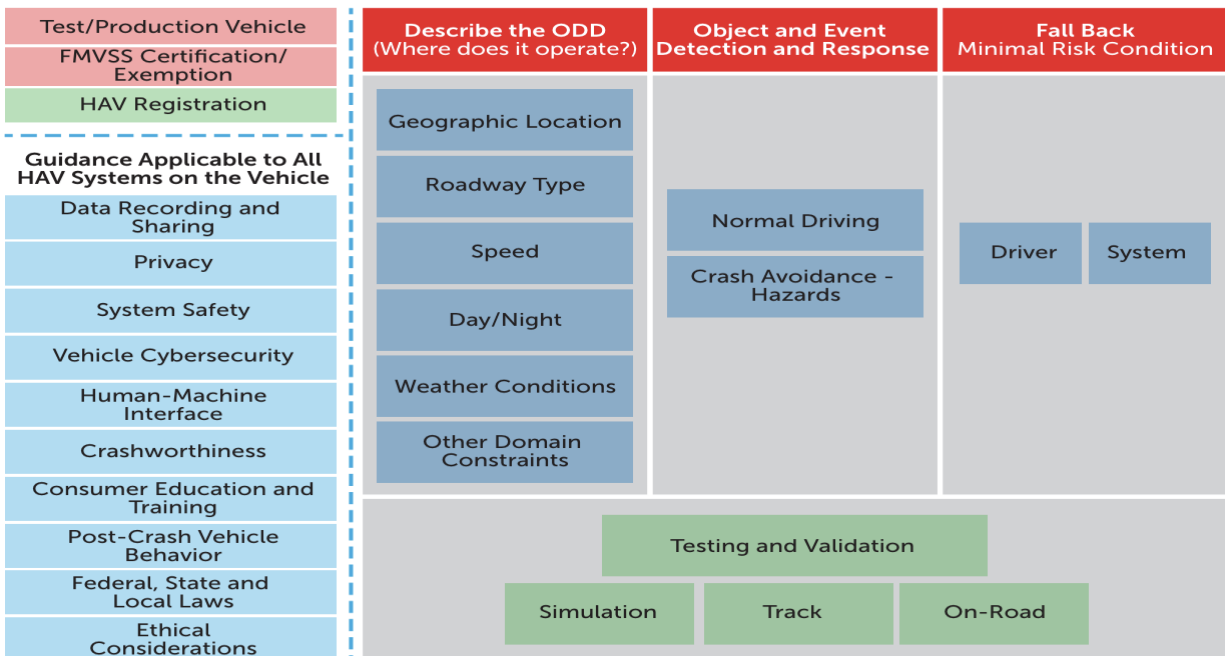
There have been proof of concepts and experiments that showed how AVs can be compromised through network and application layers. The first disabled AVs in 2015 using a mobile phone (Miller & Valasek, 2015; Seed & Kisow, 2023). The second was an application layer attack followed by a network layer, like the Miller & Valasek experiment. Still, they exploited the Tesla (Model S) WEB kit browser vulnerabilities, registering and disabling Tesla cars with a mobile phone (Nie et al., 2017; Seed & Kisow, 2023). Similar successful attacks were also experimented with another Tesla Model S and Model X where the network layer was compromised (Seed & Kisow, 2023). Recently, in 2024, it was proven by two security researchers named Talal Haj Bakry and Tommy Mysk. They proved that through Man-in-the-Middle (MiTM) phishing attacks, attackers could perform powerful actions, from registering new keys, hijacking accounts, and completely taking the Tesla car away (Toulas, 2024). Talal Haj Bakry and Tommy Mysk stated that the attack was successful for the Tesla application v.4.30.6 and 11.1 2024.2.7 for the software (Toulas, 2024). The two security researchers also reported using this version's vulnerabilities by creating a MiTM SSID network named the same as the one found in most Tesla charging stations to bait users (Toulas, 2024).

The cyberattacks, experiments, and threats that AI-AVs both on the network, application, and perception layers are valid for both domestic and foreign contexts. From a domestic context, bad actors can take advantage of domestic manufacturers' AI-AV vulnerabilities to carry out cyberattacks that can cause tremendous negative impacts on people, national critical infrastructures, and ecosystems. From the foreign AI-AVs deployed in each country, the foreign AI-AV threats are even worse as they implicate foreign entities, manufacturers, governments, ethics, data privacy, national security, and geopolitical conflicts. This can contribute to nation-state implications on leveraging AI-AVs to gain power over other countries.

## AI-AVs Frameworks, Gaps, and Insights

The AVs frameworks, guidance, and standards up to the time of authoring this research paper are the ones developed in the Federal AVs policy from the USDOT/NHTSA, NIST (DOT HS 812073) in 2014, and the ISO/SAE 21434. The framework developed in the Federal AVs policy even though it is **not mandatory** and focuses only on AV safety, the framework has scope and process guidance that High-performance Autonomous Vehicle (HAV) manufacturers are highly recommended to use to ensure their vehicles are

safe for roads, drivers, occupants, and surrounding infrastructure. The HAV guidance in the existing non-mandatory framework are also ones, the researcher considered the NHTSA recommended by default for AI-AVs for manufacturers extending AI applications in AVs. The framework, however, recommended more useful guidance even though non-mandatory but a good start for AI-AVs/HAVs manufacturers to address some areas that not just involve safety, but others as shown in Figure 1 (U.S. Department of Transportation, 2016; National Highway Traffic Safety Administration, 2016, 2022) such as cybersecurity in AVs, data in transit, and at rest, ethics, drivers training (U.S. Department of Transportation, 2016; National Highway Traffic Safety Administration, 2022), and many others areas. Many efforts and partnerships have been made from domestic (NIST) and international (ISO) perspectives to propose standards to address road vehicle risks.



**Figure 1: NHTSA Federal AVs Policy: Framework for AVs Implementation**

The National Institute of Standards and Technology partnered already with the US Department of Transport (USDOT) and NHTSA establishing the first National Institute of Standards and Technology Cybersecurity Risk Management Framework (NIST-C/RMF) that is applied to modern automobiles in October 2014 (Seed & Kisow, 2023; McCarthy & Harnett, 2014). The standards were developed from scanning different standards such as NIST SP 800-series, the Federal FIPS 199 & 200, initial road vehicles safety (ISO 26262), EVITA (E-Safety Vehicle Intrusion Protected Applications), RTCA DO-178C, the FAA Cybersecurity Certification and Accreditation and the SAE Vehicle Electrical System Security Committee’s task force, the Automotive Security Guidelines and Risk Development TF 2 (McCarthy & Harnett, 2014). The McCarthy & Harnett (2014, p.13) proposed framework (see Figure 2) does not specify the technology and tool for managing risk in road vehicles. However, the framework only lines up baselines and standards from SAE, NIST, and others throughout each NIST Risk Management Framework activity such as *categorizing, selecting, implementing, assessing, authorizing, and monitoring*.

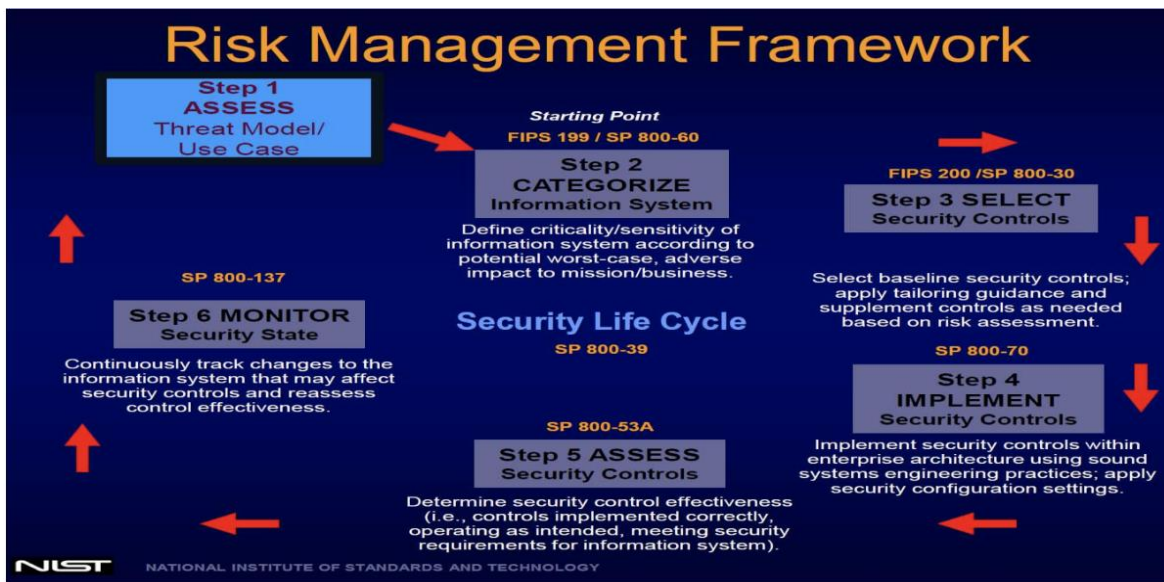


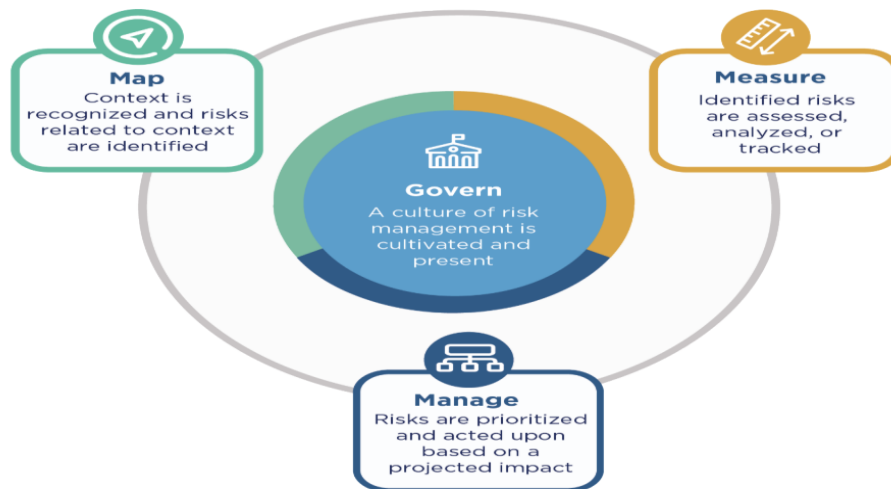
Figure 2: Risk Management Framework for Road Vehicles

Internationally, the International Organization for Standardization (ISO) recently developed the ISO/SAE 21434 standards for road automotive. The ISO/SAE 21434:2021 standards establish engineering “must” follow to effectively manage cybersecurity risks through various stages of the electrical and electronic systems (E/E) used in Autonomous Vehicles (AVs). The stages include conceptual strategy, production, operation, maintenance, and decommissioning of the E/E used in AVs. Furthermore, the standard focuses on E/E and its features, interfaces, and procedures for security against cyber threats (ISO, 2021). Finally, the ISO/SAE 21434 encourages threat analysis, organized cybersecurity management, risk assessment, and continued cybersecurity activities (Seed & Kisow, 2023; ISO, 2021). Like the USDOT/NHTSA & NIST (DOT HS 812073), the ISO was developed for compliance purposes only, therefore, it is non-mandatory (Seed & Kisow, 2023) and does not specify any type of explicit cybersecurity explanations, measures, and technology (ISO, 2021).

The rapid adoption of AI and its application to different sectors led Habbal et al. (2024) to publish a comprehensive framework called “AI TRiSM” which stands for Artificial Intelligence Trust, Risk and Security Management. AI TRiSM focuses on assessing the issues with systems and sectors employing AI. AI TRiSM's key considerations are establishing that AI application in AVs, for example, is reliable, unprejudiced, adequate, and guarantees privacy, governance, and trustworthiness (Habbal et al., 2024). This AI-focused framework was developed to increase AI trust and reliability and emerging threats monitoring and mitigation (Habbal et al., 2024). They continue arguing the importance of AI-focused frameworks like AI TRiSM in establishing AI trust that existing frameworks and standards are either failing to assess AI threats, and risks or improperly conducting a risk evaluation associated with AI application in each system. They also highlighted how the AI TRiSM framework will help manufacturers, entities, and any settings where AI is applied, to manage security and AI-related threats.

Additionally, the National Institute of Standards and Technology (NIST) recently developed the first version of the NIST AI Risk Management Framework (AI RMF 1.0) (Holistic AI Team, 2023; NIST, 2024). The framework in its first version establishes (see Figure 3) three cores: *map, measure, and manage* (Holistic AI Team, 2023; Raimondo et al., 2023). The framework was built based on how AI systems can harm people, organizations, and the ecosystem (Raimondo et al., 2023). AI RFM 1.0 is a baseline

requirement for AI manufacturers to guarantee the AI system's reliability, security, transparency, resiliency, safety, accountability, and privacy-enhanced, ethics (Raimondo et al., 2023, p.17).



**Figure 3: Cores of NIST AI RMF 1.0**

The mapping core in Figure 3 (Holistic AI Team, 2023; Raimondo et al., 2023) gives a sense of responsibility, accountability, and ethics as it requires AI systems manufacturers foremost comprehend the objectives, goals, and harms that their AI systems will have on people, infrastructure (ecosystem), and organization (entities, countries, etc.) before developing their AI products. Doing this evaluation ensures manufacturers' awareness of the future impacts and making the final call if they should or should not develop and deploy their AI systems (Raimondo et al., 2023; Holistic AI Team, 2023). The measure core is the next step once the AI systems manufacturer decides to develop and deploy their AI creations or applications. The measures core ensures that AI manufacturers not only identify, quantify, and qualify the AI risks, but also address them to ensure that the AI system developed or about to be deployed is safe, reliable, transparent, and overall enhances privacy (Raimondo et al., 2023; Holistic AI Team, 2023). In the management core, once AI systems manufacturers are quantified, qualified, and addressed, they must apply risk management in place to prevent the existing or new AI threats that may arise (Raimondo et al., 2023; Holistic AI Team, 2023). The NIST AI RMF 1.0 has governance as the driver of all three cores in establishing that AI manufacturers are devoting their efforts to prioritizing AI risk management (Raimondo et al., 2023; Holistic AI Team, 2023).

The existing frameworks and standards reviewed in this section have established grounds for AVs' capabilities expansion due to AI integration and expansion applications in the transport sector. Nevertheless, these standards and frameworks do not include NIST AI RMF 1.0 and AI trust, risk, and governance in their development. Without AI trust, the existing frameworks only scratch the surface of the actual application of AI into AVs, in this case, AI-AVs real threats and risks involved. NIST AI RMF 1.0 and AI TRiSM framework, however, aims to establish AI risk, trust, and security balance between consumers, nations, and manufacturers of AI-AVs (Habbal et al., 2024). Habbal et al. (2024) believes that balancing these three elements (trust, risk, and security) will ensure that ethical conception and deployment of AI-AVs are observed, making manufacturers transparent, accountable, and impartial when dealing with consumers and potential AI threats. Initial integration of these into the framework will ensure the success of AI applications. Because the three features (trust, security, and risks) include protecting data, and assets, and ensuring that AI integration is automatically compliant with laws, standards, and regulations that

traditional (or existent) frameworks do manually causing human errors, scaling concerns and time-consuming (Habbal et al., 2024).

### Theoretical Framework & Methodology

Effective mitigation frameworks that evaluate AI-AV threats must have *risk management, governance, and cross-disciplinary cooperation* as the theoretical approach. A risk management theory will govern how AI-AV risks are being identified, and protection, detection, response, and recovery measures in place. This theory is critical for AI-AV threats because it enumerates through identification, all the threats and challenges introduced by AI-AVs (both domestic and foreign) and states further steps that would guarantee their future detection, security, and means for responding and recovering in the event of present threats. Interdisciplinary or cross-disciplinary collaboration is mandatory and includes every aspect of AI-AVs' conception, deployment, security, risks, trust, and ethical, laws and standards. The cross-disciplinary theory establishes a joint baseline basis where security, risks, trust, ethics, AI-AV manufacturers, lawmakers, existent AV frameworks, and other critical entities such as NHTSA, USDOT, NIST, and ISO. Governance sculptures in both risk management and cross-disciplinary theory. Because it addresses actions that guarantee AI-AV threat mitigations, collaboration between entities, lawmakers, and any other aspects of the mitigation framework's components. Figure 4 shows the proposed theoretical cross-disciplinary collaboration built considering AI-AVs risk management, governance, interdisciplinary approach, and existing cybersecurity and AV frameworks. This makes the first contribution of this research throughout the AI-AV threats mitigation.

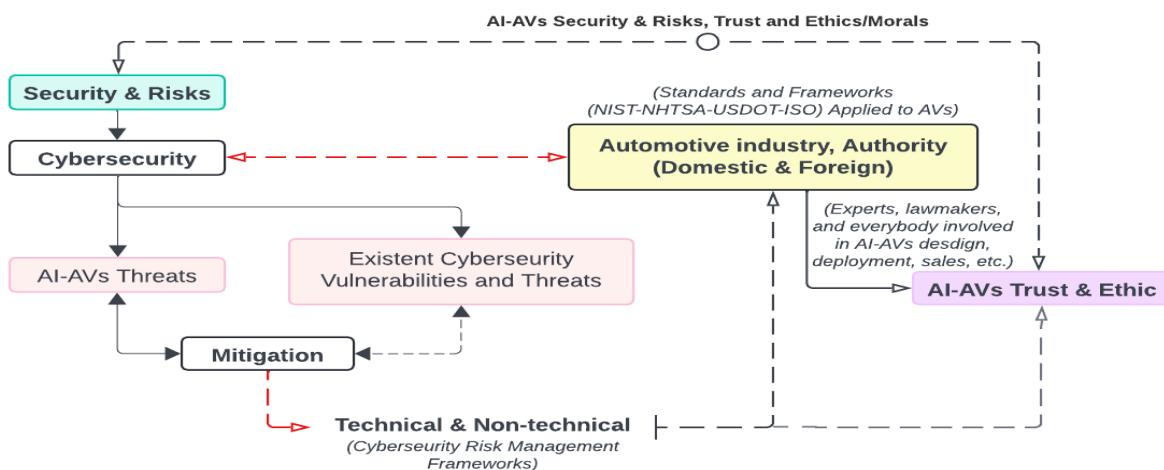


Figure 4: Cross-disciplinary Collaboration Theory

The researchers elaborated this theoretical method by analyzing existing AV frameworks and then adapting their features to the AI-AVs. The threats introduced by AVs and AI-AVs are different from both cyber and technological perspectives. They then analyzed cybersecurity risk management frameworks from a domestic and international perspective and added to that the automotive sector and all entities involved in AI-AVs, both foreign and local, establishing the AI-AVs' security and their threat mitigation while ensuring ethics, trust, and best practices are effectively established.

### Mitigation Framework Development

Key cross-disciplinary stakeholders apart from domestic and foreign AI-AV manufacturers involve from a local standpoint the USDOT, NHTSA, NIST RMF, AI TRiSM, NIST AI RMF 1.0, US lawmakers, and legal experts, and engineers. Cybersecurity has a huge role when defining the mitigation framework for AI-AV due to Autonomous Vehicles' reliance on AI. In conjunction with cybersecurity as the backbone for security, data privacy, and standards comes across-disciplinary approach as shown in Figure 4 to help mitigate other technical and non-technical aspects such as ethics, safety, trust, and morals (see Figure 5).

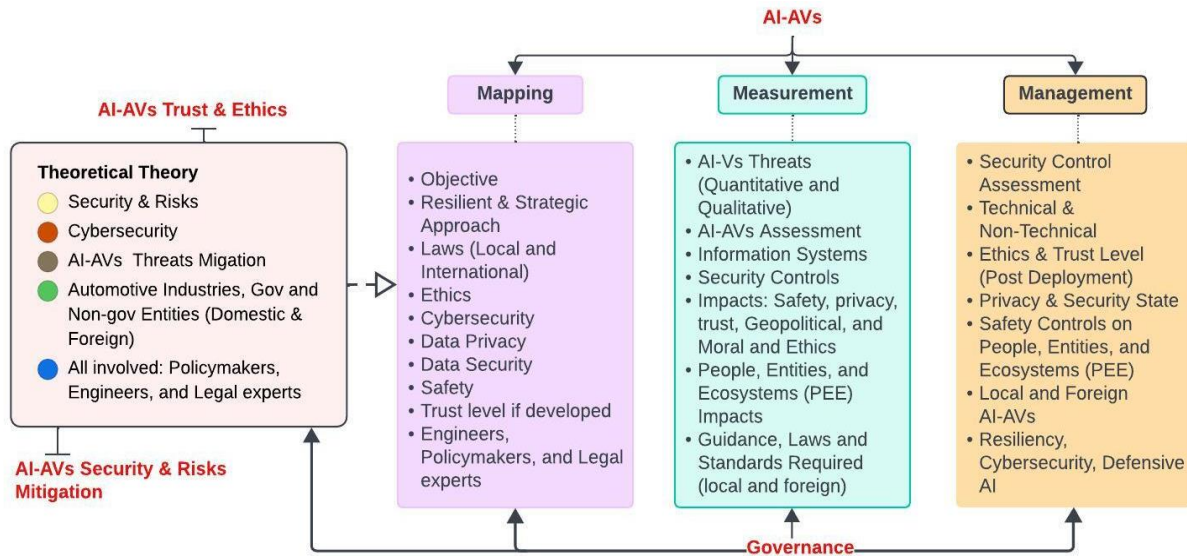


Figure 5: Cross-Disciplinary AI-AV Threat Mitigation Framework (CD-AI-AVs-TMF)

#### CD-AI-AVs-TMF: Governance

Through the governance core, this AI-AV threats framework fosters risk management as a priority throughout the entire threats' mitigation framework cycle. The governance core makes management, measurement, mapping, and even theoretical cores mandatory and indispensable. This encourages knowledge sharing between all entities (local & foreign) involved in all aspects of AI-AVs such as their development, deployment, sales, etc. Governance plays a critical role in this framework as it embeds other frameworks through knowledge sharing from different frameworks, laws, and standards such as NIST, ISO, USDOT-NHTSA, National Security Strategy (NSS), and National Cybersecurity Strategy (NCS). This allows adequate procedures to ensure that AI-AVs are trustworthy, safe, reliable, and overall applied secure structure that protects People Entities (infrastructure), and Ecosystem (PEE).

#### CD-AI-AVs-TMF: Theoretical Theory Core

The theoretical theory core is the matured cross-disciplinary collaboration theory (Figure 4). It puts together critical aspects taking place in the AI-AVs design, deployment, and above all AI applications in AVs. The theoretical theory core is essential in establishing preparatory mitigation anticipation such as AI-AVs trust, ethics, security, and risk mitigation. At this level, the framework establishes baselines on concerns and who should be involved. Individuals, entities, lawmakers, engineers, legal experts, automotive industries,

cybersecurity, and security theories are also involved in supporting governance by prioritizing risk management.

## **CD-AI-AVs-TMF: Mapping**

Developed from the AI RMF, mapping as the initial core, requires that all involved in AV-AVs as established in the theoretical theory core first comprehend before developing AI-AVs, critical aspects such as data privacy, ethics, trust, safety, laws involved, cybersecurity, resilience as well as define expected objectives for their AI-AVs. Everyone involved will help manufacturers with technical and non-technical insights utilizing practical knowledge from NIST RMF to identify AI-AV threats or legal expertise to identify ethical situations. Mapping AI-AVs includes acquiring understandings from the USDOT/NHTSA & NIST (DOT HS 812073) for domestic and ISO/SAE21434 for foreign AI-AVs manufacturers.

## **CD-AI-AVs-TMF: Measurement**

The mapping core identified AV-AV threats, whereas the measurement step qualified and quantified them. The core applies from domestic and foreign standpoints, measuring strategies such as security by design, data safeguard, accountability, security controls, impacts on safety, privacy, morals, and ethics. The quality and quantity of AI-AV threats will determine risk management technical and non-technical methods that should be enforced and prioritized at the management core to ensure AI-AVs are trustworthy & ethical, secure, and safe for people, infrastructure, and the ecosystem.

## **CD-AI-AVs-TMF: Management**

The management is the most significant core of the CD-AI-AVs-TMF. At this core, all the risks and threats quantified and qualified are managed. This is where cross-disciplinary mitigation is more seen in the action. All AI-AV risks and threats are mitigated through the application of risk management frameworks from cybersecurity (NIST RMF, ISO, AI-Trust, NIST AI RMF1.0), automotive (ISO/SAE 21434), federal (FIPS 190/200), while also applying laws and regulations that mitigate ethical and morals. The core is an ongoing risk mitigation because it draws collaboration, and insights from all other cores while including any unforeseen threats (technical and non-technical) that may arise. The more AI-AVs extend AI dependency, the more CD-AI-AVs-TMF's management core will evolve to manage new risks.

## **Discussion**

Locally, the CD-AI-AVs-TMF employs any means necessary to ensure national security, data privacy, ethics and morals, and application of AI-AVs federal laws and standards. Any domestic AI-AV manufacturers will have not only to use the CD-AI-AVs-TMF's mapping, measurement, and management cores, but also the theoretical theory to ensure they stay compliant with local regulations regarding AI-AV threats and risks in all aspects (people, infrastructure, and ecosystem). Domestic and foreign AI-AVs share the same CD-AI-AVs-TMF, however, as they cross internationally, they must comply with the ISO/SAE 21434. Domestic and foreign AI-AVs have similarities, in the threats, and risks they introduce but may vary depending on the country they are deployed in. The proposed CD-AI-AVs-TMF helps mitigate those threats, however, laws and regulations are different from one country to another, and security privacy and even ethics may have different meanings. From a cybersecurity perspective, the domestic and foreign AI-AVs share the same cross-disciplinary approach to mitigate threats in their AI-AVs. The discrepancy comes when considering nations' politics, culture, views, and competitors. The discrepancy, however, does not

prevent the proposed CD-AI-AVs-TMF from mitigating AI-AVs high threats related to cybersecurity and PEE.

## Recommendations

A continuous effort and cross-disciplinary collaboration must be observed to guarantee CD-AI-AVs-TMF effectiveness in mitigating threats and risks associated with AI-AVs. Stakeholders, AI-AV manufacturers, policymakers, engineers, cybersecurity, legal experts, government, and non-government involved in AI-AVs must work together through the proposed cross-disciplinary theoretical collaboration to extend or create theory associated with AI-AVs threats and risks. AI will keep evolving and its application to Autonomous will require a continuous cross-disciplinary complicity to include any other aspects of AI-AV threats that do not figure in the proposed theoretical theory. One way to enhance cross-disciplinary collaboration at local and global levels is by promoting transparency and accountability between all involved in AI-AVs. Through transparency and accountability, all involved in AI-AVs will collaborate to exchange what is needed to address AI-AV threats and risks. That will also make them accountable for failing to establish theoretical theories required for CD-AI-AVs-TMF effectiveness. Non-mandatory aspects of all AV guidelines multiple times recommended by precedent authors still exist, therefore, they must be mandatory. This research adds to the compulsory recommendation, accountability, and transparency through cross-disciplinary collaboration.

## Conclusion & Future Work

The research highlighted how Artificial Intelligence Autonomous Vehicles (AI-AVs) relying on perception, network, and application layers present cybersecurity threats that make AI-AVs untrustworthy, intelligence collection, and espionage moving command centers. The study also highlighted how AI as the decision maker in AVs, makes AVs susceptible to AI attacks that can cause in addition to data privacy and ethical considerations, logistical disasters. The absence of a cross-disciplinary approach to address threats introduced by AI-AVs was not only a gap but a reason to add value and significance by developing the first Cross-Disciplinary AI-AVs Threats Management Framework (CD-AI-AVs-TMF). The researchers believed that cybersecurity is, indeed, crucial for mitigating AI-AV threats, however, binding manufacturers, lawmakers, governments, and any entities involved in the design, and deployment of AI-AV is the only way to have a bigger picture of the actual AI-AV threats introduced.

The CD-AI-AVs-TMF was developed considering all aspects of AI-AVs damage on People, Entities, and Ecosystems (PEE) through its *mapping*, *measurement*, and *management* cores uplifted from the NIST AI RMF 1.0 and AI TRiSM. Considering the PEE baseline, CD-AI-AVs-TMF effectively mitigates any threats at the three layers that AI-AVs heavily depend on. CD-AI-AVs-TMF also establishes an AI trust level and endorsement through the adaptable theoretical cross-disciplinary collaboration that feeds all knowledge and insights required to develop trustworthy, transparent AI-AVs through governance. CD-AI-AVs-TMF introduces accountability and transparency to enforce secure and resilient AI-AVs design. CD-AI-AVs-TMF impact on AI-AVs' future is simply the mandatory aspect multiple times recommended by precedent research. In addition, the CD-AI-AVs-TMF accountability and transparency significance can be researched further through another cross-disciplinary collaboration that could be developed to hold accountable all involved in AI-AVs on both domestic and global scales, as HIPAA does for healthcare sectors operating in the US, for example, but in this case one, for AI-AVs.

## References

- Bhupathiraju, S. H. V., Sheldon, J., Bauer, L. A., Bindschaedler, V., Sugawara, T., & Rampazzi, S. (2023). EMI-LiDAR: Uncovering vulnerabilities of LiDAR sensors in autonomous driving setting using electromagnetic interference. *Association for Computing Machinery*. <https://doi.org/10.1145/3558482.3590192>
- Chen, S., Kuo, H., & Lee, C. (2020). Preparing society for automated vehicles: Perceptions of the importance and urgency of emerging issues of governance, regulations, and wider impacts. *Sustainability*, 12(19), 7844. <https://doi.org/10.3390/su12197844>
- Cunneen, M., Mullins, M., & Murphy, F. (2019). Autonomous vehicles and embedded Artificial intelligence: The challenges of framing machine driving decisions. *Applied Artificial Intelligence*, 33(8), 706–731. <https://doi.org/10.1080/08839514.2019.1600301>
- Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C., Prakash, A., Kohno, T., & Song, D. (2017, July 27). Robust Physical-World attacks deep learning models. *arXiv.org*. <https://arxiv.org/abs/1707.08945>
- Habbal, A., Ali, M. K., & Abuzaraida, M. A. (2024). Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, applications, challenges, and future research directions. *Expert Systems with Applications*, 240, 122442. <https://doi.org/10.1016/j.eswa.2023.122442>
- Holistic AI Team. (2023, January 26). NIST launches AI Risk Management Framework 1.0. *Holistic AI*. Retrieved May 16, 2024, from <https://www.holisticai.com/news/nist-launches-ai-risk-management-framework-1-0>
- ISO. (2021). ISO/SAE 21434:2021. Retrieved May 12, 2024, from <https://www.iso.org/standard/70918.html>
- Li, S., Sui, P., Xiao, J., & Chahine, R. (2019). Policy formulation for highly automated vehicles: Emerging importance, research frontiers, and insights. *Transportation Research Part A: Policy and Practice*, 124, 573–586. <https://doi.org/10.1016/j.tra.2018.05.010>
- Martinho, A., Herber, N., Kroesen, M., & Chorus, C. (2021). Ethical issues in focus by the autonomous vehicles industry. *Transport Reviews*, 41(5), 556–577. <https://doi.org/10.1080/01441647.2020.1862355>
- McCarthy, C., & Harnett, K. (2014). National Institute of Standards and Technology (NIST) Cybersecurity Risk Management Framework Applied to Modern Vehicles. (DOT HS 812073). *Washington, DC: National Highway Traffic Administration*
- Miller, C., & Valasek, C. (2015). Remote exploitation of an unaltered passenger vehicle. Retrieved May 8, 2024, from [https://ioactive.com/pdfs/IOActive\\_Remote\\_Car\\_Hacking.pdf](https://ioactive.com/pdfs/IOActive_Remote_Car_Hacking.pdf)

- National Highway Traffic Safety Administration (NHTSA). (2022). Cybersecurity best practices for the safety of modern vehicles. In *National Highway Traffic Safety Administration (NHTSA) Guidance* [Report]. <https://www.nhtsa.gov/sites/nhtsa.gov/files/2022-09/cybersecurity-best-practices-safety-modern-vehicles-2022-tag.pdf>
- Nie, S., Liu, L., Du, Y., Zhang, W., & Keen Security Lab of Tencent. (2017). OVER-THE-AIR: HOW WE REMOTELY COMPROMISED THE GATEWAY, BCM, AND AUTOPILOT ECUS OF TESLA CARS. In *Keen Security Lab of Tencent* [Journal-article]. Retrieved May 8, 2024, from <https://i.blackhat.com/us-18/Thu-August-9/us-18-Liu-Over-The-Air-How-We-Remotely-Compromised-The-Gateway-Bcm-And-Autopilot-Ecus-Of-Tesla-Cars-wp.pdf>
- NIST. (2024, April 30). AI Risk Management Framework | *NIST*. Retrieved May 16, 2024, from <https://www.nist.gov/itl/ai-risk-management-framework>
- Okubo, D. & The Hoffman Agency. (2020). Model hacking ADAS to pave safer roads for autonomous vehicles [Press-release]. Retrieved May 8, 2024, from <https://0e190a550a8c4c8c4b93-fcd009c875a5577fd4fe2f5b7e3bf4eb.ssl.cf2.rackcdn.com/EINPresswire-510125941-model-hacking-adas-to-pave-safer-roads-for-autonomous-vehicles-2.pdf>
- Petit, J., Stottelaar, B., Feiri, M., & Kargl, F. (2021). Remote attacks on automated vehicle sensors: experiments on camera and LiDAR. In *Security Innovation Services, Cybersecurity and Safety Institute of Distributed Systems* [Journal-article]. Retrieved May 8, 2024, from <https://www.blackhat.com/docs/eu-15/materials/eu-15-Petit-Self-Driving-And-Connected-Cars-Fooling-Sensors-And-Tracking-Drivers-wp1.pdf>
- Povolny, S., & Fralick, C. (2020, February 17). Introduction and application of model hacking. *McAfee Blog*. Retrieved May 8, 2024, from <https://www.mcafee.com/blogs/other-blogs/mcafee-%20%20labs/introduction-and-application-of-model-hacking/>
- Priscila, S. S., Sharma, A., Vanithamani, S., Ahmad, F., Mahaveerakannan, R., Alrubaie, A. J., Jagota, V., & Singh, B. K. (2022). Risk-based access control mechanism for the internet of vehicles using artificial intelligence. *Security and Communication Networks*, 2022, 1–13. <https://doi.org/10.1155/2022/3379843>
- Raimondo, G. M., Locascio, L. E., U.S. Department of Commerce, & National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). In *NIST AI 100-1* [Report]. <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>
- SAE, I. (2014). Automated driving levels of driving are defined in the new SAE international standard J3016. AS: *Warrendale*.
- Seed, K., & Kisow, M. (2023, October 4-7). Vulnerabilities in modern connected vehicles: A call to action. Panel Discussion, *Proceedings of the 63rd International Association for Computer Information Systems Conference*. [https://www.iaicis.org/conference/proceedings/IACIS\\_2023\\_Proceedings](https://www.iaicis.org/conference/proceedings/IACIS_2023_Proceedings)

Taeihagh, A., & Lim, H. S. M. (2018). Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transport Reviews*, 39(1), 103–128. <https://doi.org/10.1080/01441647.2018.1494640>

Toulas, B. (2024, March 8). MiTM phishing attacks can let attackers unlock and steal a Tesla. *Bleeping Computer*. Retrieved May 8, 2024, from <https://www.bleepingcomputer.com/news/security/mitm-phishing-attack-can-let-attackers-unlock-and-steal-a-tesla/>

USDOT Automated Vehicles Activities. (2016). US Department of Transportation. Retrieved May 8, 2024, from <https://www.transportation.gov/AV>