

AN EXPLAINABLE CYBER-PHYSICAL ARTIFICIAL INTELLIGENCE FRAMEWORK FOR CRASH PREDICTION AND PREVENTION OF HEAVY-DUTY VEHICLES ON HIGH-RISK TRANSPORTATION CORRIDORS IN NIGERIA

Gilbert I.O. Aimufua, *Buter Daniel Ioryue

Center for Cyberspace Studies, Nasarawa State University, Keffi, Nigeria

*Corresponding Author Email Address: buterdi2014@gmail.com

ABSTRACT

This study proposes an Explainable Cyber-Physical Artificial Intelligence (CPAI) framework to support secure crash-risk prediction and prevention for heavy-duty vehicles on high-risk transportation corridors in Nigeria. The conceptual architecture integrates vehicle and infrastructure data, protected communication, data-trust assessment, machine-learning prediction, explainability and operator response. It is examined through a qualitative assessment of six threat scenarios involving location spoofing, telemetry injection, replay or impersonation, false-data manipulation, availability attacks and privileged-access abuse. The design identifies preventive, detective and recovery controls across the data-to-decision path, including source validation, message integrity, access control, anomaly monitoring, model protection, audit and controlled degraded operation, while a separate protocol specifies future evaluation of prediction performance, explanation quality and cyber resilience. The framework provides a security-oriented reference architecture for subsequent implementation and testing, but its architectural control coverage does not establish measured prediction accuracy, attack-detection effectiveness or crash reduction.

Keywords: Cyber-physical systems; Cybersecurity; Explainable artificial intelligence; Heavy-duty vehicles; Crash prediction; Cyber resilience.

INTRODUCTION

Road traffic crashes remain a major public-safety problem, and the burden is especially severe in low- and middle-income settings. The World Health Organization (WHO) estimates that road traffic crashes cause about 1.19 million deaths each year, while the broader economic and social consequences extend well beyond fatalities (WHO, 2023). Heavy-duty vehicles add a distinct operational risk because of their mass, braking distance, commercial duty cycles and exposure to long-haul corridors. In Nigeria, these risks are compounded by heterogeneous traffic, uneven road conditions, limited instrumented infrastructure and fragmented access to timely transport data. A prevention system that waits for a crash report is therefore structurally late; useful intervention depends on detecting unsafe conditions before they become irreversible.

Machine-learning methods have improved crash-risk modelling by capturing nonlinear interactions among speed, traffic flow, driver behaviour, road geometry and weather. Recent reviews nevertheless show that many published models remain

retrospective, dataset-centred and difficult to transfer into real-time operations (Ali *et al.*, 2024). Studies that combine risky-driving behaviour with traffic variables demonstrate the value of multi-source data. At the same time, explainable approaches show that contextual features such as speed limit, traffic state, road slope and weather can be made visible to decision-makers rather than hidden inside a black-box score (Ma *et al.*, 2023; Masello *et al.*, 2023; Abdurashid *et al.*, 2024). Prediction accuracy, however, is only one part of the problem.

A connected crash-prevention system is also a cyber-physical system. Vehicle telemetry, Global Positioning System (GPS), roadside units, cloud services, edge gateways and Vehicle-to-Everything (V2X) links introduce trust boundaries and attack surfaces that do not exist in an offline statistical model. A forged location update can make a safe vehicle appear to be in a hazardous corridor; altered Controller Area Network (CAN) or On-Board Diagnostics II (OBD-II) values can distort speed, braking or engine-state features; replayed V2X messages can create false traffic conditions; and availability attacks can suppress alerts at the point when they are most needed. Surveys of connected and automated transportation systems therefore treat cybersecurity as a safety requirement, not merely an information-technology concern (Girdhar *et al.*, 2023; Abdo *et al.*, 2024; Gularte *et al.*, 2024).

This study integrates Artificial Intelligence (AI), Cyber-Physical Systems (CPS), Internet of Things (IoT), Explainable Artificial Intelligence (XAI) and cybersecurity within one architecture. However, it makes the security logic explicit rather than placing cybersecurity at the end of the processing chain. Security controls span the complete data-to-decision path through authentication, integrity protection, access control, anomaly detection, secure model handling, audit and recovery. Explainability is retained because operators must be able to challenge a high-risk prediction, especially when cyber compromise is suspected. The resulting design is not presented as a deployed system; it is a security-oriented reference architecture whose control coverage can be examined systematically and whose empirical performance can be tested through the validation protocol specified later in the paper. Machine learning has become central to contemporary crash prediction because it can model nonlinear relationships and interactions that are difficult to represent with conventional parametric methods. Ali *et al.* (2024) show that the field now spans real-time crash occurrence, crash frequency and injury-severity modelling, but also identify recurring weaknesses in event

selection, class imbalance, validation and future forecasting. Ma *et al.* (2023) combined risky-driving behaviour with traffic-flow data and found that richer behavioural context improves freeway crash-likelihood modelling. These findings support multi-source sensing, but they also underline the dependence of the model on the quality and authenticity of incoming data.

Explainability addresses a different but equally important weakness. Masello *et al.* (2023) used SHapley Additive exPlanations (SHAP) to rank contextual contributors to risky driving, while Abdulrashid *et al.* (2024) combined predictive models with system-level explanation of road-accident risk factors. In safety-critical transport, explanations can support operator review, post-incident analysis and model governance. Recent work on trustworthy autonomous driving likewise treats XAI as part of safety assurance rather than a cosmetic visualisation step (Kuznietsov *et al.*, 2024). In the proposed CPAI design, explanations are therefore generated alongside risk estimates and are available to operators before high-consequence interventions are executed.

Connected transportation expands the attack surface from the vehicle itself to roadside infrastructure, wireless links, edge nodes, cloud platforms and management interfaces. Girdhar *et al.* (2023) identify adversarial machine-learning risks in autonomous vehicles, while Abdo *et al.* (2024) extend the cybersecurity problem to the broader connected and automated transportation system, including multi-vehicle and infrastructure-level effects. V2X reviews consistently identify spoofing, replay, impersonation, jamming, eavesdropping and denial-of-service as threats that can undermine the timeliness and integrity of safety messages (Gularte *et al.*, 2024). Fang *et al.* (2024) further show that anomaly diagnosis in connected autonomous vehicles must consider functional, intended-functionality and cybersecurity risks together.

Two standards are particularly relevant to the governance of the proposed architecture. National Institute of Standards and Technology (NIST) Cybersecurity Framework (CSF) 2.0 organises cybersecurity outcomes around Govern, Identify, Protect, Detect, Respond and Recover, encouraging cybersecurity risk to be managed as part of broader organisational risk (NIST, 2024). ISO/SAE 21434 defines cybersecurity engineering requirements across the road-vehicle lifecycle and emphasises threat analysis, risk assessment and traceable risk treatment rather than prescribing a single technical control (ISO/SAE, 2021). These principles suit a CPAI system because no individual defence is sufficient: safety depends on layered controls, evidence, monitoring and recovery.

Research gap and contributions

The literature is mature in three areas separately: crash prediction, connected-vehicle cybersecurity and explainable AI. The gap appears at their intersection. Many crash-prediction studies assume trustworthy input data; many automotive cybersecurity studies focus on vehicle or communication security without linking attack effects to crash-risk analytics; and XAI studies often explain predictions without considering whether the underlying data or model may have been manipulated. For high-risk corridors, these separations are operationally unsafe. A practical architecture must establish trust in the data path, assess model integrity, present interpretable risk evidence and still provide a controlled response when parts of the system are degraded or under attack.

The study makes four focused contributions. First, it reformulates

crash prediction as a cyber-physical safety problem in which compromised data or communications can produce unsafe physical consequences. Second, it proposes a cross-layer CPAI architecture in which cybersecurity assurance spans sensing, communications, integration, analytics and response instead of appearing as a single downstream module. Third, it defines a threat model for attack classes that are credible in connected heavy-duty-vehicle operations and maps each threat to preventive, detective and recovery controls. Fourth, it specifies separate validation measures for crash prediction, explainability and cyber resilience so that future implementations can be assessed with evidence rather than architectural claims alone.

MATERIALS AND METHODS

Study approach

The study presents a conceptual reference architecture and a qualitative scenario-based design assessment. Its materials comprise the transportation and cybersecurity literature discussed in the Introduction, the risk-management principles of NIST CSF 2.0 and ISO/SAE 21434, and the six threat scenarios specified in Table I. No corridor dataset, trained prediction model, implemented testbed, or field deployment forms part of the reported assessment. The assessment therefore concerns the placement and intended functions of architectural controls; the empirical validation protocol describes subsequent work.

Security assumptions and threat scenarios

The framework assumes a mixed environment of trusted and potentially compromised components. Vehicle sensors and roadside devices can fail or be manipulated; wireless links can be observed, delayed, replayed or disrupted; user and service accounts can be misused; and the prediction pipeline can receive poisoned or adversarial inputs. The protected assets are not limited to data confidentiality. The primary safety properties are authenticity, integrity, availability, timeliness and traceability because an apparently valid but stale or manipulated message may be more dangerous than a disclosed message. The design therefore treats every external data source as untrusted until it passes identity, integrity and plausibility checks.

The threat model focuses on attacks with a credible path to an unsafe physical outcome. It does not assume an omnipotent attacker. Instead, it considers compromise of one or more sensors, in-vehicle interfaces, wireless links, roadside units, application programming interface (API) credentials or analytic inputs. The security objective is to prevent a single compromised component from silently controlling the crash-risk decision. Where certainty cannot be restored, the system should degrade safely by quarantining suspicious sources, lowering their trust weight, falling back to verified sources and escalating the decision to a human operator.

The six scenarios concern Global Positioning System (GPS) spoofing, Controller Area Network (CAN) or On-Board Diagnostics II (OBD-II) message injection, Vehicle-to-Everything (V2X) replay or impersonation, false-data or model poisoning, denial of service (DoS) or jamming, and credential or cloud compromise. Application programming interfaces (APIs) and multi-factor authentication (MFA) are included in the access-control considerations. These scenarios define the scope of the design assessment rather than an exhaustive attack inventory.

Design assessment procedure

Each scenario was examined against three questions: whether the proposed architecture provides a detection or containment point before the affected information reaches the decision layer; whether it includes controls intended to limit domination of the risk estimate by a compromised component; and whether it defines a degraded-operation or recovery path. The qualitative assessment links the attack target and potential safety consequence to the proposed controls and remaining exposure. These mappings are presented in Table I. The presence of a control is interpreted as design coverage and not as an experimentally established security outcome.

Proposed empirical validation

A complete empirical evaluation should be conducted in three linked experiments. The first should measure crash-prediction performance on representative heavy-duty-vehicle or corridor data using a time-aware train, validation and test design to reduce leakage from future information. The second should evaluate explanation quality and stability for high-risk cases. The third should inject controlled cyber events into the data and communication path to measure how the security controls affect prediction integrity and service continuity. Validation should compare normal operation with attacked operation so that resilience is measured as degradation and recovery, not merely as attack classification accuracy.

For a proof-of-concept implementation, tree-based ensembles such as XGBoost or Random Forest offer a practical baseline because they perform well on heterogeneous tabular transport data and integrate readily with post-hoc explanation methods. Deep models can be evaluated where image, sequence or multimodal inputs justify their complexity, but they should not be selected simply because they are more sophisticated. Security testing should include GPS spoofing, CAN/OBD-II manipulation, V2X replay or impersonation, false-data injection, denial-of-service and privileged-account misuse. Each test should record the direct security outcome and the downstream change in crash-risk prediction.

RESULTS

Threat and control mapping

The design assessment identifies control points for each of the six selected threat classes. Table I summarises the affected assets, potential safety consequences, proposed controls and residual exposure. The findings describe architectural coverage; detection rates, response latency and operational effectiveness were not measured.

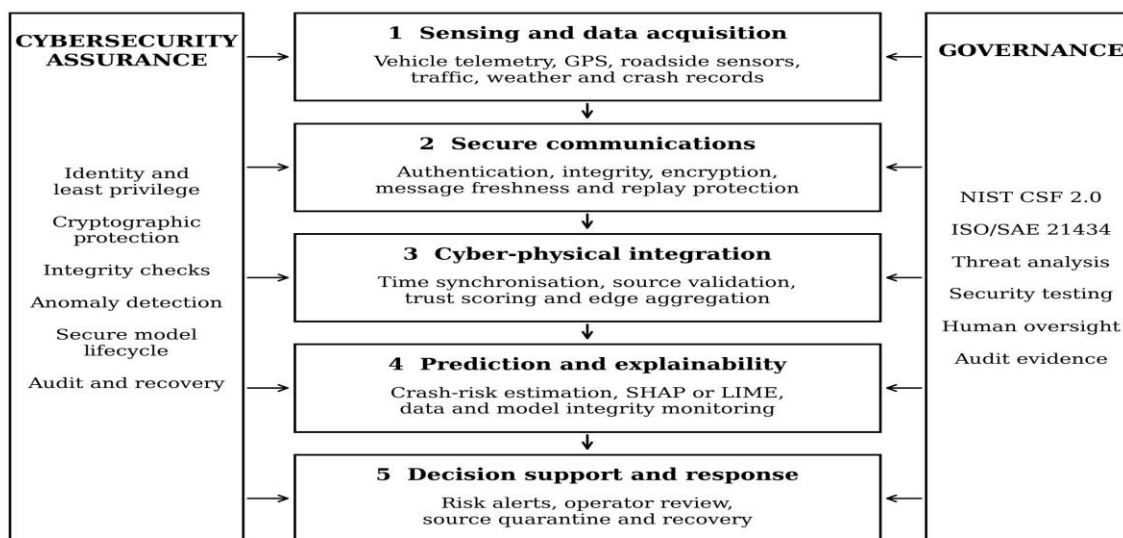
Table I. Threat scenarios and proposed security responses in the CPAI architecture

Threat scenario	Primary target	Safety effect	Required controls	Residual risk
GPS spoofing	Location/time source	Incorrect corridor position or hazard context	Multi-source location validation; signed messages; plausibility checks; anomaly detection	Degraded accuracy if alternate positioning is unavailable
CAN/OBD-II injection	Vehicle telemetry	False speed, braking or engine-state features	Gateway isolation; message allow-lists; integrity checks; behavioural anomaly detection	Legacy vehicle interfaces may lack native authentication.
V2X replay/impersonation	Wireless safety messages	False congestion, hazard or vehicle-state information	Mutual authentication; freshness tokens; anti-replay windows; certificate validation	Certificate or key compromise can bypass identity checks
False-data / model poisoning	Training or inference data	Biased or manipulated crash-risk output	Source provenance; data versioning; outlier checks; model signing; controlled retraining	Slow poisoning may resemble legitimate concept drift.
DoS/jamming	Communication and edge services	Delayed or missing warnings	Rate limiting; redundant links; local edge fallback; service health monitoring; fail-safe mode	Wide-area jamming may still reduce service availability.
Credential or cloud compromise	APIs, dashboards, model store	Unauthorized configuration, data access or model replacement	MFA; least privilege; secrets management; signed artefacts; audit logging; rapid revocation	Privileged insider abuse remains a governance risk

Architecture and operational workflow

Fig. 1 presents the proposed architecture. Cybersecurity assurance is cross-cutting. The sensing layer collects vehicle telemetry, GPS, roadside, traffic, weather and historical information, but raw data are not passed directly to the model. Source identity, message integrity, time consistency and

plausibility are checked before aggregation. Secure communication uses authenticated and encrypted channels where supported, with anti-replay protection and service-health monitoring. The cyber-physical integration layer synchronises data, assigns trust scores and resolves conflicts between sources before the prediction stage.



Security assurance and governance apply to every processing layer.
Suspected compromise triggers source isolation and operator review.

Fig. 1. Proposed Explainable Cyber-Physical Artificial Intelligence architecture with security assurance and governance across the processing layers.

The prediction and explainability layer produces a crash-risk estimate together with evidence about the variables driving that estimate. This layer also monitors data and model integrity. Large, abrupt changes in feature distributions or explanation patterns can be treated as signals for further investigation rather than accepted automatically as genuine traffic change. The response layer converts validated risk information into alerts, operator recommendations or controlled interventions. It also records security and decision events so that post-incident analysis can reconstruct which data were trusted, which controls fired and why a particular action was taken.

Governance sits outside the real-time path but constrains it. Model versions, security policies, certificates, access rights and recovery procedures require approval, traceability and periodic review. This aligns the architecture with the risk-management orientation of NIST CSF 2.0 and the lifecycle approach of ISO/SAE 21434. The design therefore separates operational speed from governance authority: real-time components can detect and contain anomalies quickly, but security policy and model changes remain controlled and auditable.

Operation begins with authenticated acquisition from vehicle, infrastructure and environmental sources. The framework then performs schema validation, time synchronisation, missing-value handling and plausibility checks. Security-relevant metadata such as source identity, message age, integrity status and device health are retained as part of the record. Communication services enforce transport protection and anti-replay controls, while the integration layer fuses only data that meet minimum trust requirements. Suspicious inputs are isolated rather than silently repaired when doing so could hide malicious activity.

The analytics stage applies a crash-risk model to the trusted feature set. Because crash events are usually rare, future implementations should avoid relying on accuracy alone. They

should use recall, precision, F1-score and precision-recall analysis in addition to the area under the receiver operating characteristic curve (ROC-AUC). Explanations are generated for high-risk outputs and checked for consistency with the observed context. A prediction that depends heavily on a feature currently flagged by the security monitor is not treated as an ordinary high-risk event; it is escalated as a potentially compromised decision. This coupling between XAI and security telemetry is central to the framework because it gives operators evidence about both risk and data trust. If an attack or severe anomaly is detected, the system enters a controlled degraded mode. The compromised source is quarantined, its trust score is reduced, and prediction continues only if sufficient verified inputs remain. When data quality falls below an operational threshold, the system should stop issuing automated safety recommendations and require operator confirmation. This fail-safe behaviour avoids the dangerous assumption that uninterrupted automation is always preferable to reduced functionality.

Explainability, privacy and model security

Explainability should not create a new information leak. Operational explanations need to be detailed enough to justify a safety decision but should not expose sensitive driver identifiers, credentials or internal security rules. Global explanations can be used for model governance and bias review, while local explanations can support investigation of individual high-risk events. Techniques such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) are suitable candidates because they can explain complex predictive models without replacing the underlying classifier (Ribeiro *et al.*, 2016; Lundberg & Lee, 2017). Their outputs should be versioned with the model and retained as audit evidence when a decision leads to an intervention.

Model security also requires controls outside inference. Training

data should have documented provenance, retraining should occur in a controlled environment, model artefacts should be signed, and deployment should verify the approved hash or signature before activation. Access to the model registry and feature pipeline should follow least-privilege principles. These measures reduce the risk that a valid software service runs a malicious or unauthorised model.

Validation requirements

Table II specifies the empirical evidence required to evaluate the proposed architecture. It is a validation plan and does not report completed experiments. Prediction and cyber-resilience outcomes should be compared under normal and attacked operation, while explanation quality and operational recovery require separate evaluation.

Table II. Proposed empirical validation dimensions for the CPAI framework

Dimension	What should be tested	Recommended measures	Minimum evidence
Crash prediction	Normal-operation ability to identify elevated crash risk	Precision, recall, F1, ROC-AUC, PR-AUC, calibration, confusion matrix	Held-out chronological test set
Explainability	Whether explanations are stable, intelligible and consistent with model behaviour	Local fidelity, rank stability, explanation consistency, operator review	Documented high-risk cases and explanation audit
Cyber resilience	Effect of spoofing, replay, injection, poisoning, DoS and credential abuse	Detection rate, false-positive rate, time to detect, prediction degradation under attack	Controlled attack injection with baseline comparison
Operational resilience	Ability to continue safely when components or links fail	Fallback success, recovery time, alert latency, service availability	Testbed or simulation with fault and recovery scenarios

In Table II, F1 denotes the harmonic mean of precision and recall; ROC-AUC denotes the area under the receiver operating characteristic curve; and PR-AUC denotes the area under the precision-recall curve.

DISCUSSION

Interpretation of the design findings

The strongest design feature is the separation of trust evaluation from prediction. Spoofed or stale data can be rejected before model inference, while communication failures can trigger local fallback instead of feeding incomplete inputs into the model as if they were normal. The main residual risks are concentrated in legacy interfaces, compromised credentials, slow data poisoning and wide-area availability attacks. These risks cannot be eliminated by model accuracy and therefore require operational controls, key management, secure update processes, monitoring and incident response.

The proposed contribution lies in connecting data-trust assessment with prediction, explanation and response. The crash-prediction and contextual-explanation studies discussed earlier motivate the analytics component (Ma *et al.*, 2023; Masello *et al.*, 2023; Abdurashid *et al.*, 2024), while the connected-transportation reviews motivate explicit treatment of compromised inputs and communications (Girdhar *et al.*, 2023; Abdo *et al.*, 2024; Gularte *et al.*, 2024). The architecture brings these concerns into one decision path, but the present assessment cannot establish superiority over an implemented alternative.

Practical implications for Nigerian deployment

A Nigerian deployment should prioritise controls that remain effective under constrained infrastructure. Edge processing is important because safety alerts should not depend entirely on continuous cloud connectivity. Device identity and key management must be operationally maintainable, not only

cryptographically strong. Roadside and vehicle devices should support remote health monitoring, but critical configuration changes should require authenticated and auditable workflows. Where older vehicle interfaces cannot provide native message authentication, gateway-based filtering and behavioural anomaly detection become compensating controls rather than optional enhancements.

The framework also supports staged deployment. A corridor can begin with passive monitoring and advisory alerts before progressing to automated interventions. During early deployment, human operators can compare model explanations, security telemetry and observed traffic conditions, creating evidence for model calibration and security tuning. This staged approach is more defensible than introducing a fully automated safety function before the attack surface and failure modes are understood.

Limitations

The principal limitation is the absence of an implemented testbed, corridor dataset or field deployment. Accordingly, the study does not claim measured improvement in crash frequency, prediction accuracy or cyber resilience. The scenario-based assessment establishes that the architecture contains appropriate control points, but it cannot quantify effectiveness. Empirical closure of this limitation requires representative transportation data and controlled cyber-attack experiments using the protocol in Table II. A second limitation is that vehicle platforms differ substantially in their support for secure diagnostics, authenticated CAN gateways and V2X security services; implementation details will therefore vary by fleet and corridor.

The qualitative assessment does not establish the completeness of the selected threat scenarios or independently validate the proposed controls. Trust-score thresholds, the crash-risk prediction horizon and criteria for withholding recommendations remain implementation requirements. The intended Nigerian application

has not been evaluated using corridor-specific observations, and alignment with the cited frameworks does not constitute certification or demonstrated standards compliance.

Conclusion

This paper presents a cybersecurity-oriented Explainable Cyber-Physical Artificial Intelligence framework for heavy-duty-vehicle crash prediction and prevention on high-risk transportation corridors in Nigeria. The design does not treat cybersecurity as an isolated layer; security controls instead apply across sensing, communication, cyber-physical integration, analytics and response. The threat model links credible attacks—GPS spoofing, telemetry injection, V2X replay, false-data manipulation, denial-of-service and credential compromise—to controls for authentication, integrity, anomaly detection, least privilege, audit, fail-safe operation and recovery.

The central argument is that a crash-prediction system cannot be considered trustworthy solely because its classifier is accurate. Safety depends on whether the data are authentic, whether the model is the approved model, whether communications are timely and available, whether high-risk outputs can be explained, and whether the system can fail safely when trust is lost. The architecture is aligned with NIST CSF 2.0 and ISO/SAE 21434 and provides a concrete validation protocol for prediction, explainability and cyber resilience. Its next research step is empirical implementation on a representative dataset or testbed, where the proposed controls can be stress-tested under both normal and adversarial conditions.

REFERENCES

- Abdo, A., Chen, H., Zhao, X., Wu, G., & Feng, Y. (2024). Cybersecurity on connected and automated transportation systems: A survey. *IEEE Transactions on Intelligent Vehicles*, 9(1):1382–1401. <https://doi.org/10.1109/TIV.2023.3326736>
- Abdulrashid, I., Zanjirani Farahani, R., Mammadov, S., Khalafalla, M., & Chiang, W.-C. (2024). Explainable artificial intelligence in transport logistics: Risk analysis for road accidents. *Transportation Research Part E: Logistics and Transportation Review*, 186:103563. <https://doi.org/10.1016/j.tre.2024.103563>
- Ali, Y., Hussain, F., & Haque, M. M. (2024). Advances, challenges, and future research needs in machine learning-based crash prediction models: A systematic review. *Accident Analysis & Prevention*, 194:107378. <https://doi.org/10.1016/j.aap.2023.107378>
- Fang, Y., Min, H., Wu, X., Wang, W., Zhao, X., Martinez-Pastor, B., & Teixeira, R. (2024). Anomaly diagnosis of connected autonomous vehicles: A survey. *Information Fusion*, 105:102223. <https://doi.org/10.1016/j.inffus.2024.102223>
- Girdhar, M., Hong, J., & Moore, J. (2023). Cybersecurity of autonomous vehicles: A systematic literature review of adversarial attacks and defense models. *IEEE Open Journal of Vehicular Technology*, 4:417–437. <https://doi.org/10.1109/OJVT.2023.3265363>
- Gularte, K. H. M., Vargas, J. A. R., Costa, J. P. J. da, Silva, A. S. da, Santos, G. A., Wang, Y., Müller, C. A., Lipps, C., Sousa Júnior, R. T. de, Vidal Filho, W. de B., Slusallek, P., & Schotten, H. D. (2024). Safeguarding the V2X pathways: Exploring the cybersecurity landscape through systematic review. *IEEE Access*, 12:72871–72895. <https://doi.org/10.1109/ACCESS.2024.3402946>
- International Organization for Standardization & SAE International. (2021). *ISO/SAE 21434:2021 Road vehicles—Cybersecurity engineering*.
- Kuznietsov, A., Gyevar, B., Wang, C., Peters, S., & Albrecht, S. V. (2024). Explainable AI for safe and trustworthy autonomous driving: A systematic review. *IEEE Transactions on Intelligent Transportation Systems*, 25(12):19342–19364. <https://doi.org/10.1109/TITS.2024.3474469>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30:4765–4774.
- Ma, Y., Zhang, J., Lu, J., Chen, S., Xing, G., & Feng, R. (2023). Prediction and analysis of likelihood of freeway crash occurrence considering risky driving behavior. *Accident Analysis & Prevention*, 192:107244. <https://doi.org/10.1016/j.aap.2023.107244>
- Masello, L., Castignani, G., Sheehan, B., Guillen, M., & Murphy, F. (2023). Using contextual data to predict risky driving events: A novel methodology from explainable artificial intelligence. *Accident Analysis & Prevention*, 184:106997. <https://doi.org/10.1016/j.aap.2023.106997>
- National Institute of Standards and Technology. (2024). *The NIST Cybersecurity Framework (CSF) 2.0*. U.S. Department of Commerce. <https://doi.org/10.6028/NIST.CSWP.29>
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why should I trust you?” Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. <https://doi.org/10.1145/2939672.2939778>
- World Health Organization. (2023). *Global status report on road safety 2023*. World Health Organization.