iseAuto as a Testbed for Safe Autonomous Driving: Bridging Formal Verification and Artificial Intelligence

Mauro Bellone*† , Ralf-Martin Soe* and Raivo Sell‡
*FinEst Centre for Smart Cities, Tallinn University of Technology, Tallinn, Estonia
†Universitas Mercatorum, Rome, Italy;
‡Dep. of Mechanical and Industrial Engineering, Tallinn University of Technology, Tallinn, Estonia

***Abstract*—Autonomous driving has reached a stage where technical feasibility is no longer in question, yet large-scale deployment is made difficult by safety concerns. Traditional approaches to system assurance, based on extensive testing or scenario catalogues, are insufficient to capture the open-endedness of real-world environments. At the same time, machine learning (ML) methods, particularly in perception, offer remarkable capabilities but remain opaque and difficult to verify formally. This tension calls for a new paradigm in which formal verification methods and artificial intelligence approaches are not seen as incompatible but rather as complementary. This position paper argues that bridging these domains is both necessary and feasible, and that *iseAuto*, an open-source autonomous shuttle developed in two generations (v1 and v2), provides a well-suited testbed to demonstrate how theoretical methods can translate into real-world applications.**

***Index Terms*—iseAuto, Autonomous Driving, Validation and Verification**

I. Introduction

The safety of intelligent and autonomous vehicles remains one of the central challenges of mass deployment. Existing approaches, based either on extensive testing or on strict formal specifications, fail to address the complexity and openness of real-world environments.

This position paper argues for a bridging paradigm, combining the rigor of formal verification with the adaptability of artificial intelligence. The *iseAuto* shuttle serves as a case study to illustrate how such integration can be achieved in practice.

Dealing with safety concerns in autonomous driving is inherently challenging under several points of view. AI-based perception systems, while remarkably powerful, remain vulnerable to corner cases and adversarial conditions, leading to a persistent uncertainty in perception. At the same time, formal verification techniques provide rigor but struggle to scale when confronted with the complexity of high-dimensional, sensor-driven environments. Beyond these technical barriers, the validation of autonomous systems continues to impose an unsustainable burden, as current industrial practice still relies on billions of kilometers of testing to demonstrate safety. Compounding these difficulties, increasing connectivity introduces cybersecurity risks that directly affect safety-critical functions. Taken together, these issues highlight the pressing need for systematic approaches that move beyond isolated solutions, integrating physics, algorithms, data spaces, and digital twins into a unified safety framework.

II. iseAuto as a Testbed for Development, Validation and Verification

In response to these challenges, TalTech has developed the *iseAuto* autonomous shuttle as a research prototype that not only enables the testing and validation of verification approaches, but also serves as a broader experimental platform for advancing perception [1], [2], digital twin integration [3], and cybersecurity [4] in automated driving. The *iseAuto* shuttle [5] has thus become a key point in our vision of safe and explainable autonomy. Developed as an open-source level-4 autonomous platform for last-mile mobility, *iseAuto v1* demonstrated the feasibility of integrating perception, decision-making, and control in a research-driven vehicle that could be openly shared with the wider community. Building upon this foundation, the recently released *iseAuto v2* significantly extends the platform's capabilities, incorporating more advanced perception pipelines, seamless integration with digital twin environments for scalable validation [6], and dedicated modules for cybersecurity testing and

This research has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 856602 (Finest Twins), from the European Union's Horizon Europe Research and Innovation Programme under grant agreement No. 101135988 (PLIADES: AI-Enabled Data Lifecycles Optimization and Data Spaces Integration for Increased Efficiency and Interoperability) and under grant agreement 101192749 — (XTRUST-6G: Extended zero-trust and intelligent security for resilient and quantum-safe 6G networks and services).

Corresponding author: mauro.bellone@taltech.ee

monitoring. By evolving from a proof-of-concept shuttle into a comprehensive research testbed, *iseAuto* enables the experimental study of methods that bridge machine learning with formal verification, while also providing a reproducible and extensible framework for collaboration between academia and industry.

III. Bridging Formal Verification and Machine Learning

While the narrative around autonomous driving often suggests that technical feasibility has already been achieved, this assumption remains contested. For industry, the dominant concern is scaling and market readiness: deploying systems that are "good enough" to operate under regulatory and liability constraints, and demonstrating safety mainly through accumulated testing mileage. Academia, however, emphasizes that key scientific challenges are still unresolved—most notably in perception robustness, explainability, and the formal verification of learning-based systems. What industry frames as an economic or market-entry hurdle often remains, at its core, a research problem requiring systematic investigation.

This divergence highlights the importance of academic testbeds such as the *iseAuto* shuttle, which enable exploration of safety, verification, and trustworthiness beyond short-term commercial incentives. Such platforms create space for addressing questions that may not yield immediate market advantage but are essential for achieving reliable and transparent autonomy in the long term.

In this spirit, we advocate for a *physically informed validation and verification* paradigm, in which Newtonian laws and physical priors are explicitly integrated into the development and assessment of AI-based models. Grounding perception and decision-making in physical principles makes outputs more interpretable and supports the design of *glass-box* models for automated driving that enhance transparency and explainability [7]. At the same time, physics-informed models are more amenable to formal verification, as demonstrated in environments such as PolyVerif, an open-source framework for autonomous vehicle validation and verification [8]. Embedding physical constraints into the learning process also reduces reliance on purely data-driven generalization, thereby increasing robustness in rare or adversarial scenarios. This stands in sharp contrast to prevailing industrial practices, where validation is typically equated with amassing vast amounts of driving mileage using opaque machine learning models. While mileage-based demonstrations provide empirical evidence of functionality, they cannot guarantee safety or resilience in untested conditions. A physically informed paradigm, by contrast, enables systematic reasoning about system behavior and opens pathways toward explainable, verifiable, and ultimately trustworthy autonomy.

Achieving this vision requires a genuinely integrated safety framework that bridges theory and practice. Algorithms for perception and decision-making must be constrained by physical laws to ensure consistency and alignment with real-world dynamics. These algorithms should then be subjected to rigorous verification, not only at the component level but also across system architectures. Validation must extend beyond conventional road testing to encompass scalable digital-twin environments, tightly coupled with real-world testbeds, in order to explore scenarios that are impractical or unsafe to replicate on public roads. At the infrastructural level, open, secure, and federated data spaces are needed to enable trustworthy data sharing and address the mounting cybersecurity demands of connected vehicles. Finally, safety assurance must be understood not as a one-off certification but as a continuous process, requiring ongoing monitoring, adaptive updates, and feedback loops throughout deployment. Taken together, these elements form a systematic safety framework that moves beyond fragmented solutions toward a comprehensive and sustainable approach to trustworthy autonomy.

References

[1] J. Gu, M. Bellone, T. Pivoňka, and R. Sell, "Clft: Camera-lidar fusion transformer for semantic segmentation in autonomous driving," *IEEE Transactions on Intelligent Vehicles*, pp. 1–12, 2024.

[2] T. Tahves, J. Gu, M. Bellone, and R. Sell, "A novel vision transformer for camera-lidar fusion based traffic object segmentation," in *Proceedings of the 17th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART*, INSTICC. SciTePress, 2025, pp. 566–573.

[3] R. Sell, M. Malayjerdi, E. Malayjerdi, M. Bellone, and H. Pikner, "Autonomous vehicle for industry 5.0: Digital twin for system safety validation," in *Proceedings of the 11th International Conference on Vehicle Technology and Intelligent Transport Systems - Volume 1: VEHITS*, INSTICC. SciTePress, 2025, pp. 660–667.

[4] A. Roberts, M. Malayjerdi, M. Bellone, R. Sell, O. Maennel, M. Hamad, and S. Steinhorst, "Analysis of autonomous driving software to low-level sensor cyber attacks," in *2025 IEEE/ACM 20th Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS)*, 2025, pp. 122–132.

[5] R. Sell, M. Malayjerdi, H. Pikner, R. Razdan, E. Malayjerdi, and M. Bellone, "Open-source level 4 autonomous shuttle for last - mile mobility," in *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2024, pp. 01–06.

[6] R. Sell, M. Malayjerdi, E. Malayjerdi, M. Bellone, and H. Pikner, "Autonomous vehicle for industry 5.0: Digital twin for system safety validation," in *Proceedings of the 11th International Conference on Vehicle Technology and Intelligent Transport Systems - VEHITS*, INSTICC. SciTePress, 2025, pp. 660–667.

[7] M. Bellone, R. Sell, and R.-M. Soe, "Glass-box automated driving: Insights and future trends," in *Proceedings of the 17th International Conference on Agents and Artificial Intelligence - Volume 1: IAI*, INSTICC. SciTePress, 2025, pp. 880–885.

[8] R. Razdan, M. Akbaş, R. Sell, M. Bellone, M. Menase, and M. Malayjerdi, "Polyverif: An open-source environment for autonomous vehicle validation and verification research acceleration," *IEEE Access*, vol. 11, pp. 28 343–28 354, 2023.