Dependable Integration Concepts for Human-Centric AI-based Systems *

Georg Macher[1], Siranush Akarmazyan[8], Eric Armengaud[5], Davide Bacciu[2], Calogero Calandra[10], Herbert Danzinger[5], Patrizio Dazzi[4], Charalampos Davalas[3], Maria Carmela De Gennaro[10], Angela Dimitriou[8], Juergen Dobaj[1], Maid Dzambic[1], Lorenzo Giraudi[6], Sylvain Girbal[7], Dimitrios Michail[3], Roberta Peroglio[6], Rosaria Potenza[10], Farank Pourdanesh[10], Matthias Seidl[1], Christos Sardianos[3], Konstantinos Tserpes[3], Jakob Valtl[9], Iraklis Varlamis[3], and Omar Veledar[5]

[1] Graz University of Technology, Graz, Austria, georg.macher@tugraz.at
[2] University of Pisa, Pisa, Italy
[3] Harokopio University of Athens, Greece
[4] Institute of Information Science and Technologies (ISTI), CNR, Italy
[5] AVL List GmbH, Graz, Austria
[6] Ideas & Motion, Turin, Italy
[7] Thales Research and Technology, France
[8] Information Technology for Market Leadership, Greece
[9] Infineon Technologies AG, Munich, Germany
[10] Marelli Europe S.p.A, Turin, Italy

Abstract. The rising demand for adaptive, cloud-based and AI-based systems is calling for an upgrade of the associated dependability concepts. That demands instantiation of dependability-orientated processes and methods to cover the whole life cycle. However, a common solution is not in sight yet That is especially evident for continuously learning AI and/or dynamic runtime-based approaches. This work focuses on engineering methods and design patterns that support the development of dependable AI-based autonomous systems. The emphasis on the human-centric aspect leverages users' physiological, emotional, and cognitive state for the adaptation and optimisation of autonomous applications. We present the related body of knowledge of the TEACHING project and several automotive domain regulation activities and industrial working groups. We also consider the dependable architectural concepts and their applicability to different scenarios to ensure the dependability of evolving AI-based Cyber-Physical Systems of Systems (CPSoS) in the automotive domain. The paper shines the light on potential paths for dependable integration of AI-based systems into the automotive domain through identified analysis methods and targets.

Keywords: AI · dependable systems · CPSoS · dependability.

1 Introduction

A comprehensive set of methods, tools, and engineering approaches has evolved in the past decades to ensure the correctness of operation and to affirm trust in

* Supported by the H2020 project *TEACHING* (n. 871385) - www.teaching-h2020.eu

automotive systems. However, new challenges are exposed through the embrace of non-deterministic components and their no strict correctness characteristics by dependable systems. Several questions arise concerning dependability and standard compliance, including process and technical engineering aspects.

For dependable system integration, different challenges are linked to deterministic and non-deterministic functions. As a deterministic function assures always the same output for a given input, it is possible to predict and determine system behaviour under all considered circumstances. That assumption is the basis for the construction of sufficiently safe products and state of the art safety argumentation based on evidence for appropriate process engineering during design, development, implementation, and testing.

In contrast, as non-deterministic functions deliver different outputs to the same inputs at different runs, traditional processes and engineering approaches do not guarantee accurate prediction of the system behaviour under all considered circumstances. Hence, it is not possible to pledge necessary evidence for appropriate safety assurance.

TEACHING project tackles the specified issue while focusing on autonomous applications running in distributed and highly heterogeneous environments. It emphasises the relationship between Artificial Intelligence (AI), humans and CPSoS by leveraging human perception for adaptation and optimisation of autonomous applications. The resulting human-centric systems leverages the physiological, emotional, and cognitive state of the users for the adaptation and optimisation of autonomous applications. The implementation is based on the structuring of a distributed, embedded and federated learning system, which is reinforced by methods that improve system dependability. The results are exploited in the automotive and avionics domains. Both domains pose an autonomous challenge with high dependability needs for the system with the human in the loop.

This paper focuses on the automotive sector, which is confronted by four main trends: electrification, ADAS and Autonomous Driving (AD), connected vehicles and diverse mobility [1]. The successful response to these trends depends on openness to changes, skills to execute the same and dedicated implementation [22]. That is especially the case for AD, which relies on smart environment sensing and complex decision making supported by CPSoS. The complexity of autonomous decision making induces the need for embedding AI algorithms. Such algorithms must mimic low-level cognitive skills to enable machines to use available data and generate appropriate decisions [22].

Machine Learning (ML) models enable dynamic extraction of knowledge from historic data to anticipate the effect of actions, plans and interactions within the cyber and the physical realm and provide adaptability to effectively handle human interactions. Hence, ML is a key enabling service providing fundamental adaptation primitives and mechanisms for applications running on the CPSoS [3]. Nevertheless, the neural-based empowerment of the CPSoS requires addressing compelling challenges related to the dependability of Neural Networks (NN).

The apparent poor dependability of AI in critical decision-making environments is one of the key causes of the low level of acceptance of and trust towards new technologies. Thus, there is a need to demonstrate and inform the community of reliability approaches for AI and their benefits.

This paper offers dependability perspectives to consider different application case of non-deterministic systems as described in section 3, which is also one key factor of the TEACHING project. Prior to that, we consider the related work and standardisation activities in section 2. The paper is concluded with the key findings and outlook for the TEACHING project in the reported context.

2 Related work and regulation activity overview

Automotive regulations and working group activities are summarised in this section, which also includes synopsis of consequentially resulting challenges for dependable AI-based systems in the autonomous automotive context. These regulation activities form the framework, in which the proposed approaches need to sustain and provide evidence, as depicted in Figure 1. European manufacturers comply with regulations provided by the United Nations Economic Commission for Europe ('UNECE')[21], which is the legal basis for uniform type approval regulations. The UNECE Regulations contain provisions for (a) administrative procedures for granting type approvals, (b) performance-oriented test requirements, (c) conformity of production, and (d) mutual recognition of type approvals. The regulations related to the automotive sector come from UNECE world forum for harmonisation of vehicle regulations (WP.29)[11].

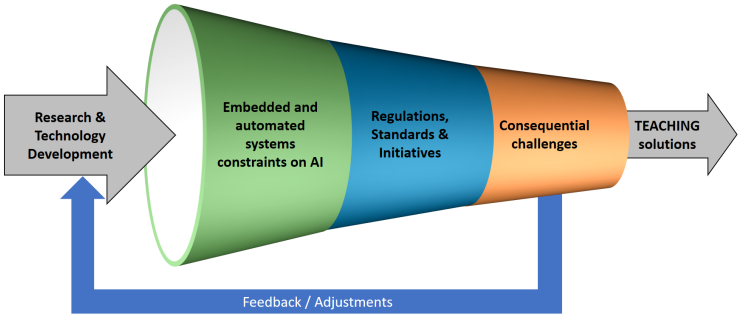

Fig. 1: Regulatory and standardisation constraints on development.

2.1 Research related work

The development of Artificial Intelligence (AI) throughout the years has introduced new concepts and methods for solving complex technical problems. The benefits of AI and machine learning were recognised by many industrial areas, which started to utilise it for their applications. However, ist utilisation in dependable systems brings new challenges into play. Although standardisation bodies from different fields have already started to consider the integration

[11] https://unece.org/wp29-introduction

of AI-based systems in the context of safety-critical applications, this topic is still in an early phase [23] and solutions focusing on the dependability of AI-based systems are seldom. G. Montano et al. [17] proposes a novel naturalistic decision-making support system for complex fault management procedures on board modern aircrafts. The framework is responsible for generating applicable configurations at run-time by using sensor data and autonomously generating effective decision support information for the pilot. The authors showed that instead of the constraint programming paradigm, AI could be effectively utilised for analysing the system and supporting the pilot with decision-making. Nevertheless, an evaluation of the quality constraints of the dependability features is not given. In the work of [14] different methods for uncertainty estimation on metrics which were designed to give more insights on the performance concerning safety-critical applications were described.

In [4] explainability is mentioned as the heart of Trustworthy AI and thus the guarantee for developing AI systems aimed at mission-critical (including safety) applications. The authors focus on approaches with humans keeping the responsibility for the decisions, but relying on machine aids.

The nn-dependability-kit [5] is an open-source toolbox to support safety engineering of NNS for autonomous driving systems. The rationale behind this is a GSN structured approach to argue the quality of NNs. The tool also includes dependability metrics for indicating sufficient elimination of uncertainties in the product life cycle and a formal reasoning engine to avoid undesired behaviours.

Besides these publications, several survey and overview papers [10,15,18,19] provide perspectives and descriptions of the AI and safety landscape [9]. However, there is no common approach to protect the systems against wrong decisions and possible harm to the environment, determination of safety measures for AI-based systems or generic pattern for the AI-based system applications.

2.2 Regulations and standards for automated vehicles

The UNECE regulations include new UN regulation on uniform provisions concerning the approval of vehicles with regards to cyber-security and cyber-security management system (UNECE R 155 2021- the final phase of approval), uniform provisions concerning the approval of vehicles with regards to software update and software update management system (UNECE R 156 2021 – under development), and regulations event data recorder (UNECE WP.29 GRVA – 2020 not frozen). There is a multitude of other regulations dealing with more specific parts of the automated vehicle (e.g. Automated Lane Keeping System (ALKS), Advanced Emergency Braking System (AEBS)).

Standards embody the specific topic's global agreed state of the art, within a particular domain. They are not legally binding, but they offer the agreed design and development practices. The following standards are specific to autonomous vehicle design, development, and testing. This is not an exhaustive list for the whole domain, but a fair representation of specific standards to be considered for autonomous vehicle development and safety and cybersecurity functions.

The most prominent automotive standard is ISO 26262 [11] intended for safety-related systems that include one or more electrical and/or electronic (E/E) components. This document addresses possible hazards caused by malfunctioning behaviour of safety-related E/E systems, including the interaction of these

systems. The included framework is intended to be used to integrate functional safety activities into a company-specific development framework.

From the perspective of process engineering, the non-deterministic system behaviour is addressed by new standards, such as SotIF (Safety Of The Intended Functionality) [13]. SotIF is a technical product safety standard with a focus on how to specify, develop, verify and validate an intended functionality to be considered sufficiently safe.

The absence of unreasonable risk due to hazards resulting from functional insufficiency of the intended functionality or by reasonably foreseeable misuse by persons is referred to as the Safety Of The Intended Functionality (SotIF). ISO PAS 21448 enhances ISO 26262 and is applied to intended functionality where proper situational awareness is critical to safety. These are situations that are derived from complex sensors and processing algorithms; especially emergency intervention systems and systems with levels of automation 1 to 5 on the OICA / SAE standard J3016 automation scale.

The third key automotive standard is ISO SAE DIS 21434 - Road vehicles - cybersecurity engineering [12]. It replaces the SAE J3061 - Cybersecurity Guidebook for Cyber Physical Vehicle Systems, provides guidelines for the organisation management of cybersecurity (CSMS) and performs operative cybersecurity activities for automotive product development. It is accompanied by ISO DTR 4804 – Road Vehicles, Safety and cyber-security for Automated Driving Systems Design, Verification and Validation. This document provides recommendations and guidance on steps for developing and validating automated driving systems based on basic safety principles derived from worldwide applicable publications. These principles provide a foundation for deriving a baseline for the overall safety requirements and activities necessary for the different automated driving functions including human factors as well as the verification and validation methods for automated driving systems focused on vehicles with level 3 and level 4 features according to SAE J3016:2018. ISO/WD PAS 5112 Road vehicles - Guidelines for auditing cybersecurity engineering and VDA - Automotive Cyber Security Management System Audit provide guidelines on how to perform cybersecurity audits and to evaluate the compliance to CSMS defined in the UNECE Reg 155.

2.3 Regulations and standards for AI-based systems

Aside from multiple ethics guidelines for AI-based systems, which are out of scope of this work, the ethics guidelines for trustworthy AI by an EU Independent High Level Expert Group on Artificial Intelligence highlight the need for AI systems to be human-centric.

Also, in the context of AI-based systems, UNECE WP.29 released a first informal document, WP.29-175-21, about artificial intelligence and vehicle regulation. This work connects AI to two automotive-specific applications: (a) HMI enhancements for infotainment and vehicle management and (b) development of self-driving functionalities (building on HD maps, surrounding detection using sensor data fused with deep learning algorithms, and driving policies for automated driving using deep learning). Currently, there are no established UNECE regulations specifically for AI-based systems.

Additionally, in the last two years, the European Commission has been actively studying AI and its impact on citizens' lives. The European Commission

created an independent group of high-level experts for AI. The European Commission released a set of guidelines for AI-based systems and a white paper on AI [20] to create a unique 'ecosystem of trust'. AI technologies may present new safety risks for users when they are embedded in products and services. A lack of clear safety provisions tackling these risks may, in addition to risks for the individuals concerned, create legal uncertainty for businesses that are marketing their products involving AI in the EU.

The new EU regulatory framework would apply to products and services relying on AI. To that aim, the intended regulatory framework will be defined following a risk-based approach. A risk-based approach requires clear criteria to differentiate between the different AI applications, concerning the question of whether they are 'high-risk' or not.

Conformity assessment is needed to verify and ensure compliance of certain mandatory requirements, which address high risks. The prior conformity assessment could include procedures for testing, inspection or certification, as well as checks of the algorithms and data sets used in the development phase.

2.4 Consequential challenges

Automated Driving: Minimising the potential for fault propagation and limiting complexity requires safety-related systems to include dependable and function-specific encapsulated systems. However, the large number of intercommunicating nodes of ADSs limits the ordinary applicability of functional safety. ADSs require new approaches to real-time fault tolerance and reasoning about the consequences of faults because the fault tolerance of ADSs is unlikely to be efficiently solved solely as a software problem due to the need to coordinate complex integrative system comprised of hardware, software and physical elements.

Connected Vehicle End2End Safety: New security risks may be exposed, opening the opportunity for automated remote attacks on vehicle fleets through increased interlacing of automotive systems with networks (e.g. V2X), new features like autonomous driving, and online software updates. Remote cyberattacks can directly affect vehicles' safety-related functions. Hence, a combined approach is needed for safety and cybersecurity analysis.

Safety of the intended functionality: The hazard analysis and risk assessment are followed by triggering condition analysis in line with the safety goals. It would be useful to describe which is the most suitable method to define the SotIF requirements to discover the weaknesses of the system design and reduce Area 3 to an acceptable level already to the first phase of system development without waiting for driving tests, simulation, endurance testing, etc.

Dependability Engineering Methods for AI-based systems: The goal is to manage and evaluate the risk posed by inadequate performance of the NNs. Considering a huge encasement in the number of advanced vehicle functionalities, an acceptable safety level for the road vehicles requires the avoidance of unreasonable risk caused by every hazard associated with the intended functionality and its implementation, especially those due to performance limitations.

However, for the systems, which rely on sensing the external or internal environment, potential hazardous behaviour caused by the intended functionality or performance limitation of the fault-free system is not adequately addressed

in the ISO 26262. Example of such limitations includes ML algorithms and AI-based system. Therefore, when developing a safety-critical AI, the safety case is used as a key tool for determining safety requirements to encapsulate all safety arguments for the AI. That safety case is based on a SotIF standard [13], demonstrating that all necessary safety measures are appropriately applied for AI. So, both ISO 262626 and SotIF are addressed in parallel to evaluate potential risks which can affect vehicle safety. Combining these two dependability domains will result in the definition of a safe function and mean that weaknesses of the technologies have been considered (SotIF) and that possible E/E faults can be controlled by the system or by other measures (ISO 26262).

3 Conceptual approaches for ensuring dependability of AI-based systems

This section presents four conceptual approaches for ensuring dependability features (e.g., safety or security) of AI-based systems for different view points. The conceptualisation is a step closer to identifying key integrated process engineering approaches that support the development of dependable products that rely on non-deterministic algorithms for different application cases. Table 1 provides an overview of the concepts, drawbacks and benefits.

Table 1: Overview of concepts, main intention, drawbacks, and benefits

Concept	Main Intention	Benefits	Drawbacks
A	support of operator	human takes decisions, traditional safety measures guarantee system safety	system operation not autonomously, decision-making mechanism must be qualified adequately
B	selection of policy	policy-based decision making ensures deterministic system behaviour, traditional safety measures guarantee system safety, autonomous system operation possible	only restricted AI algorithm capabilities due finite set of policies
C	taking critical decisions in supervised manner	comparison with deterministic supervisor system, monitor meets classic safety requirements, less restricted application of AI, autonomous operation of system	AI limited by monitor functionalities, two nearly equally sophisticated systems needed, resource usage increased, synchronisation mechanism required
D	AI to monitor system and enhance dependability	conventional system works without AI intervention, AI acts as monitor, AI increases realiability of conventional system	AI must learn normal vs. abnormal behaviour of system, AI must not reduce the system reliability in any case

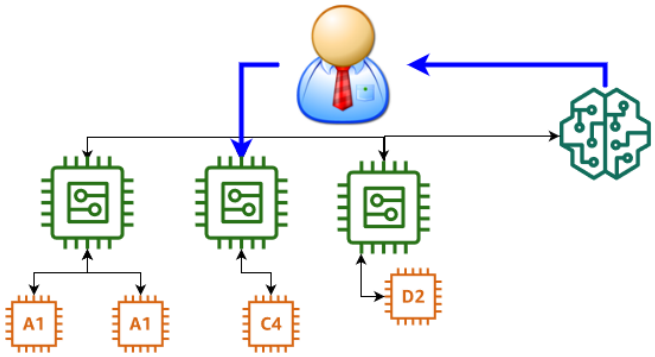

Fig. 2: Conceptual approach A for ensuring dependability of AI based systems.

3.1 Concept A: Human-in-the-Loop

This concept (depicted in Figure 2) uses AI-based system to observe and analyse specific tasks or components and recommend human-readable actions. As a 'safe' decision gate, the human decides whether the AI recommendations should be applied and if so, then how they should be executed. The presence of the human in the loop enables application of traditional safety measures to warrant system safety. The analysis of complex situations and tasks is transferred to the AI algorithm, which frees up human resources, otherwise dedicated to the analysis.

The implication of potentially wrong decision being made by the AI algorithm (i.e., detection or non-detection of a critical situation) can potential violate the system safety, but are monitored by the human. Thus, the system does not operate autonomously because human intervention is required. Kesuma et al. [8] proposed a kit that utilises AI for data anomaly detection. If AI detects unexpected signal behavior, a human is notified. The AI can enhance the monitoring system by observing many sensor signals and signalling the operators if an anomaly is detected at any stage.

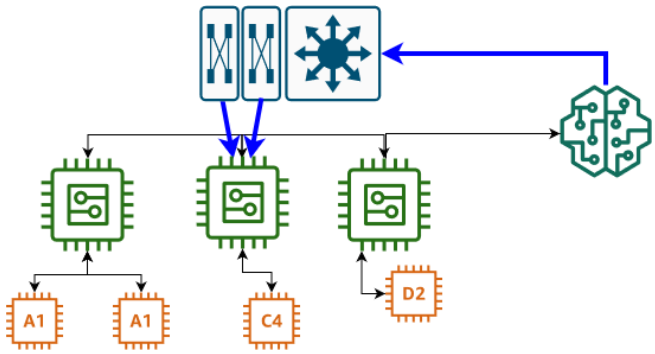

Fig. 3: Conceptual approach B for ensuring dependability of AI based systems.

3.2 Concept B: Policy-based decision integration

This concept (Figure 3) makes the AI-based system responsible for observing and analysing specific tasks or components and recommends machine-readable actions that can be translated into a finite set of policies and objectives. These policies and objectives are then used to influence the set-point generation of the safety-critical system domain. The finite set of policies and objectives (shown as two hard-wired icons in Figure 3) can be analysed for safety, and traditional safety techniques can be applied to guarantee system safety. Thus, the system operates autonomously because no human intervention is required to integrate the actions recommended by the AI algorithm and the policy-based approach can be implemented in a resource-efficient manner. A wrong decision by the AI algorithm (i.e., detection or non-detection of a critical situation) could not violate system safety, since the defined policies and transition between the policies have to be intrinsically safe. Since the set of possible actions is limited to a finite number of policies and actions, the AI algorithm's capabilities might be restricted by this limitation, but the AI algorithm itself is not considered to be a safety-critical component.

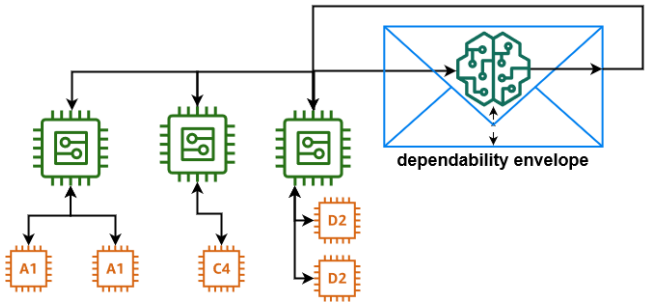

Fig. 4: Conceptual approach C for ensuring dependability of AI based systems.

3.3 Concept C: Model-based decision integration

In this concept, the AI-based system is responsible for observing and analysing specific tasks or components and recommends machine-readable actions. Instead of mapping these actions to a finite set of policies or objectives, the model-based integration approach compares the non-deterministic output of the AI-based system with the output of a deterministic model running along with the AI-based system. The concept (Figure 4) has also been referred to as 'safety envelope' [16]. The AI-based and deterministic models are designed for the same objectives, while the deterministic model is also designed to meet classic safety systems requirements. Hence, the deterministic model can be used to validate the AI-based system's output to ensure system safety.

The concept envisages AI-based system as a replacement of the human driver in fully autonomous vehicles. In such a use case, the system assumes responsibility for perception and interpretation of the vehicle environment for calculating

the input values for the setpoint generator in every specific driving situation. Since the generated inputs have a critical impact on system safety, the vehicle vendor (i.e., the original equipment manufacturer) must therefore guarantee that AI-generated inputs do not violate system safety. That is problematic since the non-deterministic nature of AI-based systems makes them unverifiable with current state-of-the-art safety methods and standards.

The advantage of the system running the deterministic model is that it can be analysed for safety, and traditional safety techniques can be applied to guaranty system safety. As the deterministic model is less restrictive than the policy-based approach, the AI algorithm's capabilities are less restricted. The system operates autonomously as no human intervention is required to integrate the actions recommended by the AI algorithm. Thus, the AI algorithm's wrong decision does not violate system safety. Hence, the AI algorithm is not considered to be a safety-critical component.

The major drawback of this concept is dictated by the limitation of the deterministic model, which might restrict the capabilities of the AI algorithm. This means that two nearly equally sophisticated systems must be developed and the two (possibly) resource-intensive systems must be executed side-by-side in a synchronous manner.

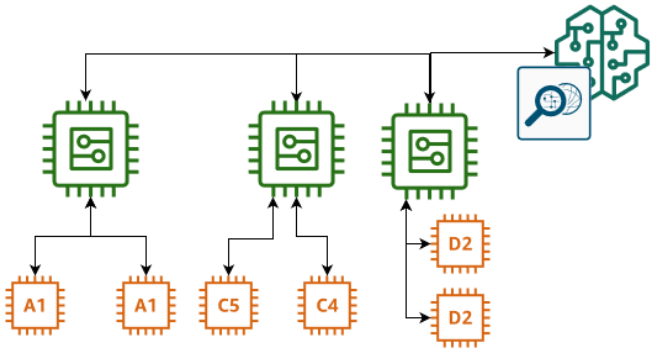

Fig. 5: Conceptual approach D for ensuring dependability of AI based systems.

3.4 Concept D: AI-based system for ensuring dependability

This approach (Figure 5) inverts assumptions for the application case. The AI-based algorithm is not seen as a potential source of harm for the dependability of the CPSoS but as an intelligent monitoring unit. It is used to monitor the conventional system. The AI algorithm learns the normal/expected system behaviour under real operating conditions without influencing the functionality and dependability of the system itself unless the system violates its specification. Consequently, AI algorithm enhances the dependability of systems through monitoring and learning the behaviour of a (dependable) system under observation. If AI detects abnormal behaviour, countermeasures are either recommended or

automatically triggered. Equally, the same is valid for the real-time guarantees of the system. Furthermore, the reliability of the system under observation including the monitoring architecture shall not be lower than the system reliability of the system under observation on its own. Also, the source code of the system under observation shall not be modified by the monitoring architecture [7]. This concept is used in cyber-security applications for anomaly based network intrusion detection [2] or dynamic honeypots [6].

4 Conclusion and Outlook

The assurance of dependability, especially considering novel run-time adaptive AI-based approaches in the automotive domain, is still an open issue that lacks standard solutions for industrialisation. However, there is a necessity to establish the means of delivering a convincing and explicit affirmation that the systems under development are at the appropriate maturity level. To shine a light on possible paths for the dependability of AI systems in the automotive world, this paper presents (I) the body of knowledge of the TEACHING project and related regulatory activities, and (II) four perspectives on dependability architecture concepts and patterns for the adoption of continuously learning AI-based systems into dependable automotive applications. The presented conceptual dependability perspectives support the identification of key integrated process engineering approaches for the development of dependable products in the automotive domain and beyond. We provide an overview of the advantages and drawbacks that are resulting from different perspectives.

In that respect, TEACHING continues to seek the most suitable perspectives and the balance of benefits to continue optimising driving automation applications. Consequently, we are adapting the control strategy to the human in the loop, hence fulfiling the promise of a human-centric approach to driving automation where improvements of the machine itself depend on the human state.

Acknowledgments

The presented work is partially supported by TEACHING, a project funded by the EU Horizon 2020 research and innovation programme under GA n.871385.

References

1. E. Armengaud, B. Peischl, P. Priller, and O. Veledar. Automotive Meets ICT—Enabling the Shift of Value Creation Supported by European R&D. In J. Langheim, editor, *"Electronic Components and Systems for Automotive Applications "*, pages 45–55, Cham, 2019. Springer International Publishing.
2. D. Ashok Kumar and S. Venugopalan. Intrusion detection systems: A review. *International Journal of Advanced Research in Computer Science*, 8:356–370, 2017.
3. D. Bacciu, S. Chessa, C. Gallicchio, and A. Micheli. On the Need of Machine Learning as a Service for the Internet of Things. In *Proceedings of the 1st International Conference on Internet of Things and Machine Learning*, IML '17, New York, NY, USA, 2017. Association for Computing Machinery.

4. R. Chatila, V. Dignum, M. Fisher, F. Giannotti, K. Morik, S. Russell, and K. Yeung. *Trustworthy AI*, pages 13–39. Springer International Publishing, Cham, 2021.
5. C.-H. Cheng, C.-H. Huang, and G. Nuehrenberg. nn-dependability-kit: Engineering neural networksfor safety-critical autonomous driving systems. In *"Workshop on Artificial Intelligence Safety, SafeAI 2020. Proceedings"*, 2019.
6. V. Chowdhary, A. Tongaonkar, and T.-c. Chiueh. Towards Automatic Learning of Valid Services for Honeypots. pages 469–470, 01 2004.
7. A. Goodloe and L. Pike. Monitoring distributed real-time systems: A survey and future directions. Communications in Computer and Information Science. National Aeronautics and Space Administration, 2010.
8. H. Kesuma et. al. Artificial intelligence implementation on voice command and sensor anomaly detection for enhancing human habitation in space mission. In *2019 9th International Conference on Recent Advances in Space Technologies*, 2019.
9. J. Hernández-Orallo. Ai safety landscape from short-term specific system engineering to long-term artificial general intelligence. In *2020 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops*, 2020.
10. T. Hinrichs and B. Buth. Can ai-based components be part of dependable systems? In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 226–231, 2020.
11. ISO - International Organization for Standardization. ISO 26262 Road vehicles Functional Safety Part 1-10, 2011.
12. ISO - International Organization for Standardization. ISO/SAE CD 21434 Road Vehicles - Cybersecurity engineering, under development.
13. ISO - International Organization for Standardization. ISO/WD PAS 21448 Road vehicles - Safety of the intended functionality, work-in-progress.
14. M. Henne et. al. Benchmarking uncertainty estimation methods for deep learning with safety-related metrics. In *"Invited paper - 2019 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)"*, pages 83–90, 2020.
15. Y. Ma, Z. Wang, H. Yang, and L. Yang. Artificial intelligence applications in the development of autonomous vehicles: a survey. *IEEE/CAA Journal of Automatica Sinica*, 7(2):315–329, 2020.
16. G. Macher, N. Druml, O. Veledar, and J. Reckenzaun. Safety and security aspects of fail-operational urban surround perception (fusion). In *International Symposium on Model-Based Safety and Assessment*, pages 286–300. Springer, 2019.
17. G. Montano and J. Mcdermid. Effective naturalistic decision support for dynamic reconfiguration onboard modern aircraft. 05 2012.
18. S. Houben et. al. Inspect, understand, overcome: A survey of practical methods for AI safety. *CoRR*, abs/2104.14235, 2021.
19. Silverio Martínez-Fernández et. al. Software engineering for ai-based systems: A survey. *CoRR*, abs/2105.01984, 2021.
20. The European Commission. White Paper on Artificial Intelligence: a European approach to excellence and trust. *European Commission*, 2020.
21. "UNECE". "task force on Cyber Security and (OTA) software updates (CS/OTA)". https://wiki.unece.org/pages/viewpage.action?pageId=40829521, (accessed 2019-09-07).
22. O. Veledar. New business models to realise benefits of the iot technology within the automotive industry. *WU Executive Academy*, 2019.
23. W. Wahlster and C. Winterhalter. Deutsche normungsroadmap künstliche intelligenz. 11 2020.