Rethinking Certification for Higher Trust and Ethical Safeguarding of Autonomous Systems

Dasa Kusnirakova[a] and Barbora Buhnova[b]
Faculty of Informatics, Masaryk University, Brno, Czech Republic

Keywords: Autonomous Systems, Trust, Certification, Regulation, Ethics.

Abstract: With the increasing complexity of software permeating critical domains such as autonomous driving, new challenges are emerging in the ways the engineering of these systems needs to be rethought. Autonomous driving is expected to continue gradually overtaking all critical driving functions, which is adding to the complexity of the certification of autonomous driving systems. As a response, certification authorities have already started introducing strategies for the certification of autonomous vehicles and their software. But even with these new approaches, the certification procedures are not fully catching up with the dynamism and unpredictability of future autonomous systems, and thus may not necessarily guarantee compliance with all requirements imposed on these systems. In this paper, we identified a number of issues with the proposed certification strategies, which may impact the systems substantially. For instance, we emphasize the lack of adequate reflection on software changes occurring in constantly changing systems, or low support for systems' cooperation needed for the management of coordinated moves. Other shortcomings concern the narrow focus of the awarded certification by neglecting aspects such as the ethical behaviour of autonomous software systems. The contribution of this paper is threefold. First, we discuss the motivation for the need to modify the current certification processes for autonomous driving systems. Second, we analyze current international standards used in the certification processes towards requirements derived from the requirements laid on dynamic software ecosystems and autonomous systems themselves. Third, we outline a concept for incorporating the missing parts into the certification procedure.

1 INTRODUCTION

According to the most recent estimates, the majority of human driving will be replaced by autonomous vehicle (AV) technology by the year 2050 (Litman, 2022). Besides the expected benefits of increased road safety, significant cost savings and reduced energy consumption and pollution (Dia et al., 2020), the introduction of self-driving vehicles into public places creates new challenges for safeguarding the mobility ecosystem as a whole in order to ensure vehicles' safe operation. Therefore, alongside the quickly emerging autonomous software advancements, it is critical to develop regulatory frameworks that would be designed to adapt to the technological changes, so that the safety risks are minimized.

Regulations inevitably have to evolve to keep up with technological progress, more so in the face of the increasing software intensity of autonomous-driving systems. This has already occurred in the history of vehicle certification methods. Initially, when vehicles were made only from mechanical components (brakes, tyres), the driving function and all decision-making was in hands of the driver. In this case, classic certification approaches were sufficient. However, following the introduction of ABS[1] and other systems with a greater level of complexity, it became clear that the classical approach was insufficient in assessing all safety-relevant aspects because of the extensive number of potential testing scenarios. As a consequence, process- and functional-oriented safety audits were introduced, one of which is the Annex 6 of the UN Regulation no.79 (Publications Office of the European Union, 2018).

Future road vehicles are expected to gradually overtake more critical driving functions, until manual driving might be fully replaced. This will shift

[a] https://orcid.org/0000-0002-5341-902X
[b] https://orcid.org/0000-0003-4205-101X

---

[1]Anti-lock Braking System; a safety system used on land vehicles activated in case of a skid in order to allow the driver to maintain more control over the vehicle.

the focus and the responsibility from the driver towards the system installed in the AV. With that, the importance and complexity of the electronic control systems utilized in vehicles will continue to increase. This will also significantly increase the number of possible scenario variations. However, the testing phase performed in the same way as for conventional vehicles, that is, verifying the system based on a pre-defined set of tests, will be able to thoroughly examine only a limited subset of all safety areas and scenarios (Kalra and Paddock, 2016). Besides that, autonomous systems, i.e. *"systems changing their behaviour in response to unanticipated events during operation"* (Watson and Scheidt, 2005), among which AVs clearly belong, rely on automatic software updates at runtime in order to adjust to specific environment or context changes (Capilla et al., 2022), or to improve AI components of an AV in general (Baldini, 2020). Again, the traditional certification procedures are not designed to promptly deal with this situation. They do not assume any, or at most a limited number of changes in already certified systems. But just like software engineering does not end with the deployment of the system, future AVs will require novel technological as well as legal approaches for their thorough quality control even during runtime in order to ensure public road safety.

The challenges of the AV certification go even beyond the safety of autonomous ecosystems. Studies have shown that trust plays an essential role in the adoption of automated systems (Cioroaica et al., 2019), from both the societal (people are willing to accept and use the systems) as well as technological (interactions between communicating systems during runtime are trustworthy) point of view (Sagar et al., 2022). However, continual technological progress of AV exceeding the boundaries of previously defined safety requirements hinders trust formation in these systems (Cioroaica et al., 2020), and certification, generally perceived as a trust-building mechanism, is failing to provide sufficient legal guarantees in its current state. It is, therefore, necessary to adapt the certification methods to the envisioned technological advancements and facilitate the adoption process of AV into society.

In this paper, we first explore the necessity to improve the present certification methods for autonomous driving systems, and second, we evaluate recently published standards against the identified requirements for future autonomous systems. Taking into consideration the expertise from social computing as well as the characteristics of dynamic autonomous ecosystems in which such systems operate, we outline a concept for incorporating the missing parts into the certification procedure. To this end, we formulate the requirements for the future certification procedures of autonomous systems, select relevant standards and evaluate them against the defined requirements.

The rest of the paper is organized as follows. Section 2 presents related work in the field of certification of autonomous driving systems. Section 3 lists the characteristics and specifies the certification context for autonomous systems, based on which we define the requirements for standards used for the certification of future autonomous systems. In Section 4, we describe the research methodology and select certification standards for evaluation. Section 5 presents the evaluation results. In Section 6, we outline the suggestions for improvements, needed for the standards to reflect the specifics of future autonomous (eco-)systems.

2 RELATED WORK

The question of the suitability of the current vehicle certification processes for the future has already been to some extent discussed in the literature. However, to the best of our knowledge, no paper has yet highlighted the importance of balancing both trust and ethics in the process of certifying autonomous systems.

Bonnin (2018) has pointed out that certification changes are needed. But their criticism of the standards was mainly directed towards the insufficient reflection of the technological advancement in the connected software development processes. The article does not cover specific characteristics of autonomous vehicles and the ecosystems in which they operate. Other criticisms of certification procedures were detected, too. In (Burzio et al., 2018), which asks for modifications from the standpoint of cyber-security. But the most often addressed certification problems are those related to the safety of AVs, as presented in (Zhao et al., 2022) or (Cummings, 2019). The last mentioned paper specifically criticizes possibly lower vehicle safety unless software upgrades are taken into account in the certification process.

Naturally, the identified shortcomings and criticism of the currently used certification procedures sparked the development of suggestions for improvements. In (International Organization of Motor Vehicle Manufacturers (OICA), 2019), GRVA[2] group

---

[2]Working Party preparing draft regulations, guidance documents and interpretation documents for adoption by the parent body

working under United Nations Economic Commission for Europe (UNECE) in collaboration with experts of the International Organization of Motor Vehicle Manufacturers (OICA) presented a new way of validating autonomous vehicles for the purpose of certification based on a multi-pillar approach consisting of a scenario catalog. Bakirtzis et al. (2022) proposed dynamic certification built on modelling and testing, which are constantly intertwined during the system's life cycle. Besides that, a verifiably-correct dynamic self-certification framework for autonomous systems is discussed in (Fisher et al., 2018), while Dia et al. (2021) introduced the concept of a certification framework for autonomous driving systems based on the Turing test. Digital certificates are proposed to be used in combination with trust and reputation policies for ensuring safety and detection of hijacking vehicles in (García-Magariño et al., 2019). All of the attempts partially cover the shortcomings of static certification, however, the debate regarding certification is grounded solely in the context of safety and security and lacks to consider other aspects, such as ethics and trust of AVs.

Even though trust and ethics are in general frequently debated issues in the context of self-driving cars, the literature presents these concepts more in connection with privacy preservation (Lai et al., 2021), or general calls for adjusting the software development standards and best practices (Kwan et al., 2021a; Myklebust et al., 2020) for building software with social responsibility. Considering trust and ethics directly within the certification process itself does not seem to be covered yet. To this end, if we are to trust the systems that make important decisions for us not only in the area of safety, but also in ethics, we must ensure that these aspects are also considered at the stage of vehicle certification providing legal guarantees.

3 SPECIFICATION OF THE CERTIFICATION CONTEXT

Responsible system certification can be hardly achieved when performed both in isolation from the environment in which the system operates, and without taking into account the characteristics of the system under consideration. AVs are considered cyber-physical systems (CPSs), i.e. *"smart networked systems with embedded sensors, processors and actuators that are designed to sense and interact with the physical world"* (Greer et al., 2019). Therefore, the characteristics of CPSs can aid with the identification of the characteristics of AVs needed for thorough cer-

tification.

The aim of this section is to cover the key characteristics of autonomous CPSs as well as the whole surrounding ecosystems, and specify the certification context based on which we then define the requirements for certification standards for future autonomous systems.

3.1 Autonomous Cyber-Physical Systems

The characteristics of autonomous systems need to be considered when determining requirements for certification frameworks for driving systems. Weyns et al. (2021) defines key principles for future CPS engineering principles, from which the characteristics of future autonomous systems can be derived. The following principles are listed: (1) *crossing boundaries* related to close contact between social-, physical-, and cyber-spaces; (2) *leveraging the humans* and their integration in the design and operation processes instead of treating them only as users of a system; (3) *on-the-fly coalitions* as a way of addressing complex problems through forming multi-agent systems; (4) *dynamically assured resilience* to withstand uncertainty, context changes or any other disruptions and continue in service provision; (5) *learn novel tasks*, that is, utilizing knowledge from the past effectively to deal with novel situations.

3.2 Dynamic Software Ecosystems

Besides that, the ecosystem in which the autonomous system operates must also be considered. The ecosystem identifies its surroundings and entities communicating with the system, specifies the relationships between them as well as determines the system context, that is, the system's main meaning and ways of use.

Capilla et al. (2022) defined ecosystems formed by software autonomous systems as ecosystems supporting dynamic, smart, and autonomous features which are required by modern software systems. In particular, the following features need to be taken into consideration when designing certification suitable for autonomous systems: (1) *automation* understood as automated self-adaptation of the ecosystem on context changes during runtime; (2) *autonomy* meaning there is no longer any human monitoring of the systems inside the ecosystem; (3) *dynamic goal evaluation* as intelligent adaptation to dynamic needs with the intention to achieve the proposed and explicitly defined goals; (4) *automated trust management* as
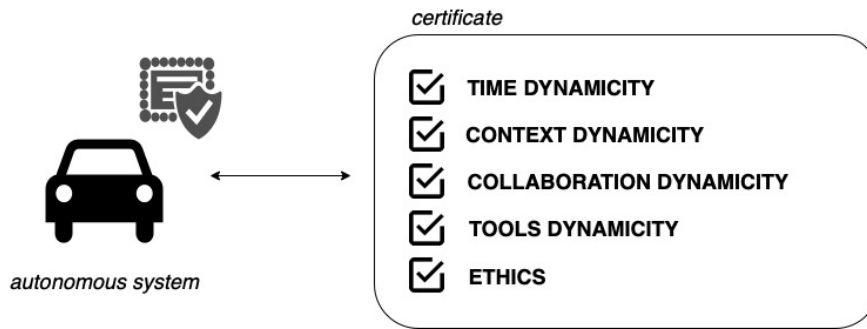
Figure 1: Key certification standards' aspects on future autonomous systems.

a crucial concept driving decision-making[3]; and (5) *architecture implications* as the ecosystem's dynamic nature, driven by the constant (dis-)connecting nodes, impacts the architecture of the ecosystem.

3.3 Ethical Aspects

In case of human-operated vehicles, the drivers themselves are responsible for applying the basic rules of safety and morality when driving. However, integrating autonomous systems into ecosystems shared with humans shifts the ethical aspects of driving on the system. And when humans are not in charge anymore, the system must bear a certain moral obligation instead.

Examples of such circumstances include the well-known trolley problem (Nyholm and Smids, 2016), in which a collision is irreversible and any action that is made results in a tragedy. Other examples where ethical aspects are applied concern an exhibition of altruistic behaviour (e.g., informing other road members about a danger they cannot yet see, such as a person crossing the street on the red light), or a demonstration of solidarity with surrounding entities (e.g., transparency in terms of notifying other drivers of one's change in speed or direction before doing so, in order to maintain smooth traffic and avoid dangerous situations).

Thus besides the primary requirements on autonomous systems, which is assuring safety and security (Ors and Carlson, 2019), ethical considerations might govern the gray space where it becomes clear that some level of harm might happen anyway, or an action can be done to help the overall ecosystem, although such action is not required by law.

The idea of implementation of ethical principles

in autonomous driving elaborates even further on one of the requirements for a CPS, in particular *leveraging the humans*. When ethical standards are applied, people are no longer seen as just users but as equal members of the ecosystem with individual needs and dignity.

The research generally agrees on the fact that applying ethical norms is crucial for trust formation in autonomous systems (Kwan et al., 2021b; Wang et al., 2022; Gerdes, 2020). In human society, ethical diversity is observed, which necessitates the possibility to adjust the moral setting of a system according to one's individual preferences within a personal ethical framework. However, a personal ethical framework with no regulation is insufficient as it would lead to numerous ethical dilemmas. A preferred ethical model should therefore reflect public rational ethical inclinations while at the same time provide users with limited freedom to adjust the particular ethical setting according to their personal preferences (Wang et al., 2022). The contribution to the discussion regarding the form and way of the implementation of specific ethical principles in autonomous systems is out of the scope of this paper, though. Instead, as this paper deals with the readiness of current certification standards for autonomous vehicles, we focus on the examination of whether individual standards take into account ethical issues regarding the driving task.

4 METHODOLOGY

To identify the gaps within the certification procedures of autonomous systems, the requirements on these systems need to be defined and a set of standards for evaluation needs to be established. We derive the requirements based on the characteristics of autonomous systems as well as based on the characteristics in which such systems operate. We draw the knowledge from literature devoted to the study of autonomous systems in dynamic software ecosys-

---

[3]the overarching ecosystem goal derives from the achievement of tactical goals on lower levels and collaboration, which is needed to fulfill the main goal, relies on trust guarantees in such dynamic and hard-to-predict environments

tems. As for the standards selection, we have selected standards published or re-confirmed within the last seven years (2016-2022) in various fields of the automotive domain from international organizations for standardization and evaluated their readiness for fully autonomous driving. The steps are more thoroughly described below.

4.1 Requirements Selection

The characteristics of the dynamic software ecosystems, in which autonomous systems operate, as well as the systems themselves, presented in Section 3, lead to the need for the shift from staticity towards a more dynamic approach. By staticity, we mean the certification standards that shall no longer be static in terms of:

1. *time*, that a vehicle is granted a certificate based on its status at a single time point, typically at the moment of production,

2. *context*, that the certificate is granted based on pre-defined and finite set of tests,

3. *collaboration*, that the standard neither supports dynamic creation of coalitions nor considers communication with other entities in the environment,

4. static *tools* used within the certification process, such as documentation examination, model report assessment or visual inspection.

As we believe future certification standards should be revised to reflect the ecosystems' dynamicity, we suggest the shift the aforementioned static characteristics into their dynamic version, towards which we evaluate the selected standards to check their readiness for autonomous systems. We refer specifically to the dynamicity in the following aspects:

1. *time*, that the standard is able to deal with dynamic system changes, e.g. caused by a software update,

2. *context*, that the standard suggests using tools to thoroughly verify the systems' functionality in dynamic context, such as unforseen and upredictable situations,

3. *collaboration*, that the standard supports dynamic creation of multi-agent systems to enable optimization and solving of complex problems more efficiently, e.g. through on-the-fly coalitions,

4. *tools*, to check whether the tools used to verify the systems' compliance with the standard are dynamic.

For the analysis of the selected standards, in addition to the dynamic characteristics, we also consider

5. *ethics*, meaning whether ethical issues are explicitly discussed or taken into account within the cases addressed in the particular standard under evaluation.

The selected key certification standards aspects on future autonomous systems are visualized in Figure 1.

4.2 Standards Selection

The documents for evaluation were chosen in the following manner. First, we seached for standards issued or re-confirmed in the last seven years (2016-2022) by international organizations for standardization (ASAM, ISO, ITU, SAE, UNECE). We focused on standards published for light-duty vehicles (passenger cars) in the domain of on-road automated driving, regarding software requirements specification, software updates, testing, validation and verification, or guidance for automated driving in general.

From this search, we identified 27 relevant documents. However, not all of the selected items could be analyzed. In addition to documents publicly unavailable at the time of writing of this paper (e.g. ISO/AWI TS 5083 being under development), we also omitted documents defining only the taxonomy and terminology, exchange format or language specifications (ASAM standards), and documents not directly related to automated driving systems (e.g. audit and other control activities guidelines, such as in ISO/PAS 5112). The final list of standards selected for evaluation is in Table 1.

5 STANDARDS EVALUATION

After the selection of relevant documents, we performed an evaluation of the standards' suitability for fully autonomous driving against the requirements defined in Section 4.2. Most of the selected standards are concerned with safety, which is understood in terms of a) functional and system specification, b) identification of triggering events, c) reduction of safety risk, and d) validation and verification of the system's functionality (see Figure 2).

The evaluation results are presented in Table 2, the standards are sorted according to their identifier in alphabetical order. The evaluation results in each of the reviewed aspects are discussed in a more detailed manner in the following paragraphs, with providing the evaluation summary at the end of this section.
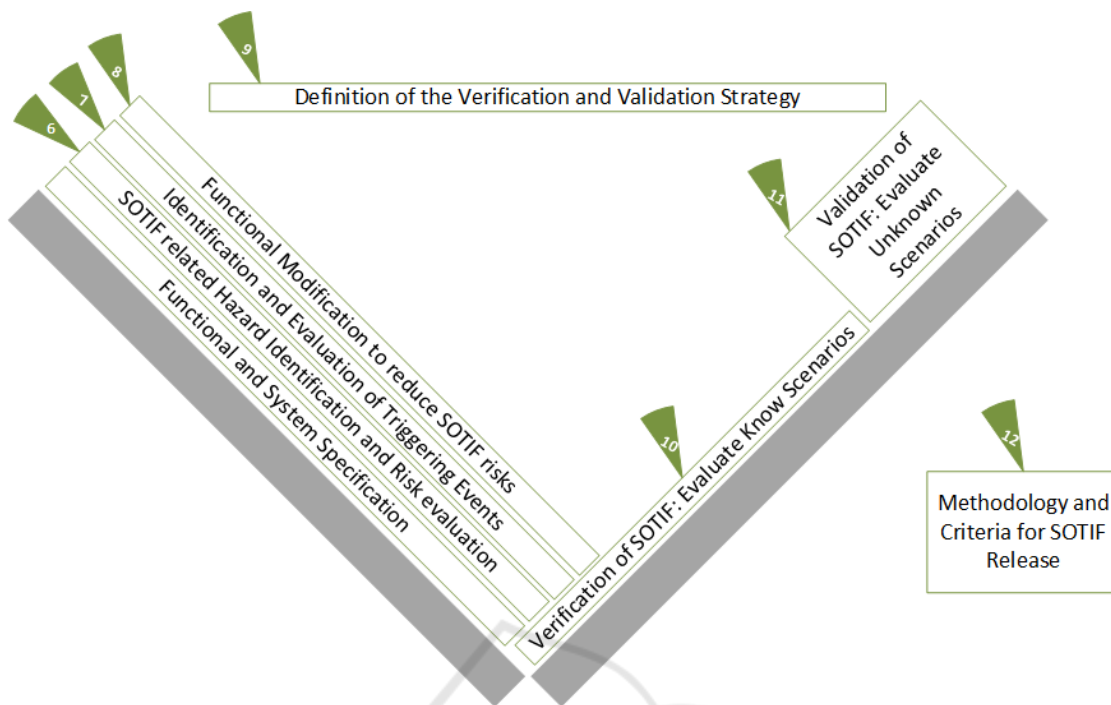
Figure 2: Elements of safety of the intended functionality defined by the standard ISO 21448 (Heicon Global Engineering GmbH, 2020).

5.1 Time Dynamicity

Only seven of the reviewed documents show readiness for the dynamicity in the time context. That is, these standards include measures for the case when there are additional updates of an already certified system, and do not require re-execution of the whole certification process conducted by an authority. Typically, these standards provide guidelines for creating the necessary documentation of changes, performing thorough testing, and arranging on-road monitoring for verifying that the update has not disrupted the intended functionality of the vehicle (UL 4600).

The rest of the standards do not provide any information regarding software updates (e.g. ISO 15037, ISO 22735). This makes it unclear how to proceed in case changes in the software are needed, such as ensuring recognition of a new type of traffic sign, and whether the awarded certificate becomes invalid after an update in the already certified software.

5.2 Context Dynamicity

Unlike all other categories, dynamicity within the context category seem to be relatively well addressed. In terms of system verification, most of the standards require performing simulation tests with unforseen and upredictable situations to thoroughly verify the

systems' safety or security in the development phase of an autonomous vehicle system. Some standards, such as ISO 21448, provide detailed guidance on verifying the system's performance in unknown scenarios. Public road testing is also widely used in the reviewed documents as a testing method in the final stages of the development process. However, techniques and requirements for public road-testing rules differ over countries, and a global regulatory framework for public road AV testing is yet to be developed (Abu Bakar et al., 2022).

Standards not supporting context dynamicity, which are missing a check mark in the table, usually rely on the verification of systems' functionality based only on predefined set of tests (ISO 11270, SAE J3018), or the information about testing tools is missing due to the standard's focus area (ISO/TS 5255).

5.3 Collaboration Dynamicity

As for the collaboration dynamicity, standards ISO/DIS 26262 and ISO 21448 provide a comprehensive specification and design of communication between vehicle and other entities within the surrounding ecosystem. Besides that, ISO/SAE 21434 incorporates distributed cybersecurity activities, assigning cybersecurity responsibilities between multiple parties. Because these three standards explic-

Table 1: List of standards selected for evaluation.

No.	Standard ID	Name	Field
1	ISO 11270	Intelligent transport systems - Lane keeping assistance systems (LKAS) — Performance requirements and test procedures	Safety&Assurance
2	ISO 15037 series	Road vehicles — Vehicle dynamics test methods	Safety&Assurance
3	ISO 21448	Road vehicles — Safety of the intended functionality	Safety&Assurance
4	ISO 22735	Road vehicles — Test method to evaluate the performance of lane-keeping assistance systems	Safety&Assurance
5	ISO 22737	Intelligent transport systems — Low-speed automated driving (LSAD) systems for predefined routes — Performance requirements, system requirements and performance test procedures	Safety&Assurance
6	ISO/DIS 26262 series	Road vehicles — Functional safety	Safety&Assurance
7	ISO/SAE 21434	Road vehicles — Cybersecurity engineering	Cybersecurity
8	ISO/TR 21959 series	Road vehicles — Human performance and state in the context of automated driving	Human factor
9	ISO/TR 4804	Road vehicles — Safety and cybersecurity for automated driving systems — Design, verification and validation	Safety&Assurance, Cybersecurity
10	ISO/TS 5255 series	Intelligent transport systems — Low-speed automated driving system (LSADS) service	Data
11	SAE J2945	On-Board System Requirements for V2V Safety Communications	Safety&Assurance
12	SAE J3048	Driver-Vehicle Interface Considerations for Lane Keeping Assistance Systems	Safety&Assurance
13	SAE J3061	Cybersecurity Guidebook for Cyber-Physical Vehicle Systems	Cybersecurity
14	UL 4600	Standard for Safety for the Evaluation of Autonomous Products	Safety&Assurance

Table 2: Evaluation of certification standards. The numbers of standards match with the list of standards defined in Table 1.

No.	Time dynamicity	Context dynamicity	Collaboration dynamicity	Tools dynamicity	Ethics
1	-	-	-	✓	-
2	-	✓	-	-	-
3	✓	✓	✓	✓	✓
4	-	✓	-	-	-
5	-	✓	-	-	-
6	✓	✓	✓	✓	-
7	✓	✓	✓	✓	-
8	-	✓	-	-	✓
9	✓	✓	unclear collaboration form and purpose	✓	✓
10	-	-	unclear collaboration form and purpose	-	-
11	✓	-	unclear collaboration form and purpose	✓	✓
12	-	-	unclear collaboration form and purpose	-	-
13	✓	✓	-	✓	-
14	✓	✓	unclear collaboration form and purpose	✓	✓

itly address the possibility of collaboration of multiple agents within the ecosystem, they were marked as fully supportive towards dynamic collaboration.

However, even though the dynamic creation of coalitions is covered by these documents, we still see an opportunity for improvement. We notice that none of the reviewed standards considers any form of trust management when collaborating with other entities. Since the entities might have malicious intentions, naively trusting any entity willing to collaborate poses a serious security risk to the operation of the entire ecosystem.

In the UL 4600, ISO/TR 4804, ISO/TS 5255, SAE J2945 and SAE J3048 standards, terms like *"dependencies between items"*, *"vehicle to vehicle communication"*, or *"arrays of systems implementing other vehicle level functions"* have been found. Thus, the possibility of vehicle communication with other entities within their surrounding environment is at least partially addressed. However, it is unclear (1) how and in which directions the communication is carried, and (2) whether joint strategy creation is possible. Because of that, these standards were marked as providing only a partial support. All other standards do not seem to directly support collaboration dynamicity at all, or just to a negligible extent.

5.4 Tools Dynamicity

More than a half of the evaluated documents enable using of dynamic tools for checking the system's compliance with a given standard. Such a check of compliance typically includes on-site testing or execution of further system's verification and validation measures. Even in the cases where dynamic tools are present, a large part of the compliance checks is still performed by static tools, such as manual inspection of documentation, processes and procedures. Standards which are not identified as supportive towards the use of dynamic tools in the assessment table employ only static tools for compliance checks.

5.5 Ethics

The series of ISO/TR 21959 standards is devoted to ethics in a large extent. Except for the ethical considerations, the standards also present the best practices in the field of societal trust formation. In particular, they establish guidelines for a better acceptance of autonomous vehicles based on the analysis of various human factors having an impact on perception and trust formation towards autonomous systems within the automotive domain.

But ethical considerations do not appear to be generally taken into account by the assessed standards, though. Four standards in the Safety& Assurance area refer to the need of absence of unreasonable risk, which is defined as a risk that is *"unacceptable in a certain context according to valid societal moral concepts"* (ISO 21448:2022, en). Other moral aspects do not seem to be addressed, and we find the absence of consideration of moral aspects particularly problematic. We are convinced that autonomous systems such as autonomous vehicles, which have direct responsibility for human lives, bear moral obligation.

The ethics of AI systems in general is a topic that

is currently frequently debated, and the authorities are working on complex market regulation strategies. The activities of the European Commission may serve as an illustration. Starting with the publication *"Policy and Investment Recommendations for Trustworthy Artificial Intelligence"* (High-Level Expert Group on Artificial Intelligence (AI HLEG), 2019), the European Commission laid the foundations of AI regulation in the European Union, covering the fundamental questions on the border of law and ethics. Since then, multiple documents discussing ethics have been published, but many concerns and questions are still unanswered.

Since ethics is a complex issue, it is possible that the authors of standards in the field of autonomous driving are awaiting the central regulatory bodies' further recommendations and will adjust the documents once the direction in which the regulation will evolve is clear. In any case, autonomous systems cannot be developed without ensuring human and societal needs are taken into consideration during the systems' operation. We are convinced that even on a technical level, it is necessary to establish mechanisms which can control the observance of ethical principles, and which can be followed up after further guidelines are published by the central authorities.

5.6 Evaluation Summary

To sum up, the evaluation results are diverse. While some of the standards show clear insufficient preparedness to reflect the dynamicity of future autonomous systems and ecosystems in which they operate (e.g. ISO 22735, ISO 22737, or SAE J3018), other standards show signs of readiness for autonomous driving after meeting all (ISO 21448) or a vast majority (ISO/TR 4804, UL 4600) of the specified requirements. Yet, there is room for improvement even in this case.

Namely, in all the analyzed standards, we identified the absence of considering the creation of dynamic coalitions as well as the absence of trust management in standards with support towards coalition dynamicity as one of the most urgent deficiencies from the evaluated aspects. Besides that, we undoubtedly see shortcomings in the field of ethics, too. We observe that most of the standards are narrowly focused, usually towards safety, while neglecting ethical aspects, which are often directly related.

6 SUGGESTED IMPROVEMENTS

Given the presented evaluation and a relatively long process of preparation and approval of standards, which typically takes three years from the initial proposal to its final publication according to ISO[4], it is not surprising that standards do not fully reflect the advancement in dynamic autonomous systems. Several shortcomings in the field of ethics were identified, too. Yet, there are mechanisms that could help the standards better support the dynamicity of the software ecosystems and the future autonomous cyber-physical systems, while reflecting the moral setting of the society as well.

In this section, we outline five suggested directions that could lead to a better fit of the standards with the requirements of future dynamic autonomous systems and ecosystems, including an advanced assurance of ethical awareness.

6.1 Real-Time Validation of a Certificate and Its Properties

Traditional certificates, whose purpose is to provide certain guarantees about the quality of the certified system, are granted at a specific point in time, usually at the time of production of the certified system. However, in a dynamic and ever-changing environment, such a certificate may not keep its validity over time; instead, its validity may change depending on the context or deteriorate over time.

The certification schemes for dynamic autonomous systems shall account for some form of identification or quantification of the certificate deterioration, in order to address the issue of time dynamicity. In other words, in a dynamic environment, it is necessary to constantly re-check the certificate's validity and decide to what extent the certificate and its guarantees can be trusted.

For better illustration of the need for the implementation of real-time certificate validation, consider a scenario where an AV has a valid certificate (*valid* in the traditional meaning of the certification concept) but its behavior is suspicious, raising the possibility that the vehicle has malicious intentions. Other vehicles will be able to react to such a situation more effectively if there are mechanisms for evaluating suspicious manners that deviate from the expected, typical behavior. In this case, the real-time validity check would show that the vehicle was not certified to handle this specific situation, or that the certificate is outdated because vehicle's software has not been updated

on a version fixing critical bugs that allow attackers to hack the vehicle.

6.2 Certificate Combined with Vehicle's Reputation

The examination of certification standards revealed certain flaws, which might lead to a false sense of security. As discussed above, there may be situations in which the vehicle is certified, yet its safe operation is impaired (e.g. after a faulty software update that is not detected by static tools used to check the system's compliance with a standard). A possible solution to this issue could be linking the certificate to the vehicle's reputation.

In dynamic ecosystems, deciding who to trust becomes a challenging task. As studied in the disciplines of trust, reputation is used as one of the tools for trust-building. Reputation can be defined as *the overall quality of an entity derived from the judgements by other entities in the underlying network, which is globally visible to all members of the network* (Jøsang, 2007), and when such information is propagated through the network of connected entities, it can have a substantial effect on decision-making. Besides that, by providing information allowing distinction between trustworthy and untrustworthy nodes, reputation can also help in dealing with observable misbehavior (Srinivasan et al., 2008) and minimizing damage in case of an insider attack (Li and Song, 2016).

This concept may be particularly useful for addressing the time dynamicity problem within certification. Same as people gain reputation by having their actions evaluated by their peers, an entity's reputation in a smart ecosystem is dependent on how it behaves and interacts with others (Buhnova, 2023). Moreover, a collection of experience during runtime feeding updates of the score of trustworthiness could also be used to promote or demote the certificate's validity.

Tying the certificate to the vehicle's reputation could help other entities respond to changes in a dynamic environment in an even more flexible manner. Demonstrated on a hypothetical scenario, a vehicle would be less trusted if its reputation reported by other vehicles declined despite having a valid certificate, caused by other vehicles reporting its sudden suspicious behavior (indicating a software bug or an attack). In case of a serious reputation drop, the vehicle's certificate could be temporarily or completely revoked to prevent further damage.

---

[4]https://www.iso.org/developing-standards.html

6.3 Extension of the Certificate's Status to a Scale

Another alternative, which is partly related to the linkage of the certificate to the concept of reputation discussed above, is to redefine the concept of certification in terms of the range of potential values it can acquire. The traditional view on a certificate nowadays treats it as a binary value (*valid/invalid*, or *granted/not granted* certificate). Such a perception might be unnecessarily restricting.

Instead, we suggest rethinking the concept and considering it rather a scale to better represent the current certification status of the system installed in the vehicle or other autonomous system. Indeed, there are multiple ways to interpret this newly proposed concept. To mention some of the possible meanings, the scale might represent e.g. the number of software updates installed in the system, or the amount of time elapsed since the last official verification of the system's compliance with a particular standard. The exact interpretation of the certification scale is up for further research and discussion.

6.4 Considering Certificate's Context-Dependant Validity

The implementation of trust management appears to be necessary even in addressing the context and coalition dynamism issue. During runtime, evaluating the acceptable level of trust, or eventually the vulnerability risks, is strongly context-dependent.

Consider two AVs initializing mutual communication in two scenarios. In scenario A, the AVs' intention is to exchange a batch of weather-related data and then stop any further interaction. In scenario B, the entities eshablish a connection with the purpose of creating a vehicle platoon in order to reduce fuel consumption due to lower air resistance. However, close collaboration needed for vehicle platoon formation in Scenario B raises more serious trust concerns about the safety of riding in such close proximity compared to the situation of exchanging data in Scenario A. And even if a vehicle may be certified in correctly handling interaction with other vehicles in some contexts, its behavior may not be verified or guaranteed in the other contexts. Therefore, before engaging in any interaction, it is the system's responsibility to verify whether the awarded certificate, as well as the other party, can be trusted in the given context.

6.5 Certificates Combined with Ethical Concerns

Evaluating the safety of products or their environmental footprint before they are allowed to enter the market is nowadays a common practice. But the review of current standards in this paper has shown that ethical aspects are still not frequently considered. Technology development has to be ethically supervised, though. Otherwise, intelligent systems developed with the intention to help can easily turn to harm or disadvantage certain groups of people.

Our idea to address the certification gap regarding the ethics of autonomous systems is to assess the ethics of an autonomous system in the same way as safety or environmental aspects. In particular, we suggest combining certification with Ethical Digital Identities (EDI), a concept introduced by Cioroaica et al. (2022). Derived from the concept of Digital Identities (Windley, 2005), EDI serve as the basis for safeguarding the evolution of intelligent safety-critical systems in terms of ethics.

7 CONCLUSION

In this paper, we evaluated the readiness of current certification standards for future autonomous driving systems. We analyzed the characteristics of both autonomous systems and the dynamic software ecosystems in which they operate, from which we derived a set of requirements, namely *Time Dynamicity, Context Dynamicity, Collaboration Dynamicity, Tools Dynamicity*, and *Ethics*, which we then used to assess the standards.

The results demonstrate that the present standards are not entirely ready for the expansion of autonomous driving systems, and also assisted us in identifying their primary shortcomings. One of the most serious deficiencies is referring to the *Collaboration Dynamicity* aspect. We criticize mainly the lack of support for the creation of dynamic coalitions among standards, as well as the complete absence of any kind of trust management strategies for establishing communication with other entities. Another shortcoming concerns neglecting ethical aspects in the standards' focus.

In order to address the identified gaps, we outlined a concept for the improvement of certification standards. We present five ideas for rethinking the certification that we believe will help move the discussion towards a complete solution of the identified problems, so that standardization for autonomous systems (not only those in the automotive domain) will

better fit the requirements of future dynamic systems and ecosystems with ethical awareness, which can be trusted. The presented ideas are subject for further research and will be elaborated on in our future work.

ACKNOWLEDGEMENTS

REFERENCES

Abu Bakar, A. I., Abas, M. A., Muhamad Said, M. F., and Tengku Azhar, T. A. (2022). Synthesis of Autonomous Vehicle Guideline for Public Road-Testing Sustainability. *Sustainability*, 14(3).

Bakirtzis, G., Carr, S., Danks, D., and Topcu, U. (2022). Dynamic Certification for Autonomous Systems. https://arxiv.org/abs/2203.10950.

Baldini, G. (2020). Testing and Certification of Automated Vehicles including Cybersecurity and Artificial Intelligence Aspects. Technical report, Publications Office of the European Union, Luxembourg, Luxembourg.

Bonnin, H. (2018). The Certification Challenges of Connected and Autonomous Vehicles. In *9th European Congress on Embedded Real Time Software and Systems (ERTS 2018)*, pages 1–4, Toulouse, France.

Buhnova, B. (2023). Trust management in the Internet of Everything. In *Proceedings of the 16th European Conference on Software Architecture-Companion Volume*, pages 1–13. Springer. Preprint at http://arxiv.org/abs/2212.14688.

Burzio, G., Cordella, G. F., Colajanni, M., Marchetti, M., and Stabili, D. (2018). Cybersecurity of Connected Autonomous Vehicles : A ranking based approach. In *2018 International Conference of Electrical and Electronic Technologies for Automotive*, pages 1–6, Milan, Italy. IEEE.

Capilla, R., Cioroaica, E., Buhnova, B., and Bosch, J. (2022). On autonomous dynamic software ecosystems. *IEEE Transactions on Engineering Management*, 69(6):3633–3647.

Cioroaica, E., Buhnova, B., Jacobi, F., and Schneider, D. (2022). The Concept of Ethical Digital Identities. In *2022 IEEE/ACM 1st International Workshop on Software Engineering for Responsible Artificial Intelligence (SE4RAI)*, pages 17–20. IEEE/ACM.

Cioroaica, E., Buhnova, B., Kuhn, T., and Schneider, D. (2020). Building trust in the untrustable. In *2020 IEEE/ACM 42nd International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS)*, pages 21–24. IEEE.

Cioroaica, E., Kuhn, T., and Buhnova, B. (2019). (Do not) trust in ecosystems. In *2019 IEEE/ACM 41st International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)*, pages 9–12. IEEE.

Cummings, M. (2019). *Adaptation of Human Licensing Examinations to the Certification of Autonomous Systems*, pages 145–162. Springer International Publishing, Cham, Switzerland.

Dia, H., Bagloee, S., and Ghaderi, H. (2020). *Technology-Led Disruptions and Innovations: The Trends Transforming Urban Mobility*, pages 1–36. Springer International Publishing, Cham, Switzerland.

Dia, H., Tay, R., Kowalczyk, R., Bagloee, S., Vlahogianni, E., and Song, A. (2021). Artificial Intelligence Tests for Certification of Autonomous Vehicles. *Academia Letters*. Article 418.

Fisher, M., Collins, E., Dennis, L., Luckcuck, M., Webster, M., Jump, M., Page, V., Patchett, C., Dinmohammadi, F., Flynn, D., Robu, V., and Zhao, X. (2018). Verifiable Self-Certifying Autonomous Systems. In *2018 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*, pages 341–348, Memphis, TN, USA. IEEE.

García-Magariño, I., Sendra, S., Lacuesta, R., and Lloret, J. (2019). Security in Vehicles With IoT by Prioritization Rules, Vehicle Certificates, and Trust Management. *IEEE Internet of Things Journal*, 6(4):5927–5934.

Gerdes, J. C. (2020). The Virtues of Automated Vehicle Safety - Mapping Vehicle Safety Approaches to Their Underlying Ethical Frameworks. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 107–113, Las Vegas, NV, USA. IEEE.

Greer, C., Burns, M., Wollman, D., and Griffor, E. (2019). Cyber-Physical Systems and Internet of Things. Nist special publication 1900-202, National Institute of Standards and Technology.

Heicon Global Engineering GmbH (2020). ISO 21448 – Safety of the Intended Functionality (SOTIF) Why is it required? https://heicon-ulm.de/en/iso-21448-safety-of-the-intended-functionality-sotif-why-is-it-required/.

High-Level Expert Group on Artificial Intelligence (AI HLEG) (2019). Policy and Investment Recommendations for Trustworthy AI. European Commission, Brussels. https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

International Organization of Motor Vehicle Manufacturers (OICA) (2019). Future Certification of Automated Driving Systems. https://unece.org/DAM/trans/doc/2019/wp29grva/GRVA-02-27e.pdf.

ISO 21448:2022(en) (2022). Road vehicles — Safety of the intended functionality. Standard, International Organization for Standardization, Geneva, Switzerland.

Jøsang, A. (2007). Trust and Reputation Systems. In *Foundations of Security Analysis and Design IV*, pages 209–245, Bertinoro, Italy. Springer.

Kalra, N. and Paddock, S. M. (2016). Driving to safety: How many miles of driving would it take to demon-

strate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice*, 94:182–193.

Kwan, D., Cysneiros, L. M., and do Prado Leite, J. C. S. (2021a). Towards Achieving Trust Through Transparency and Ethics. In *2021 IEEE 29th International Requirements Engineering Conference (RE)*, pages 82–93.

Kwan, D., Cysneiros, L. M., and do Prado Leite, J. C. S. (2021b). Towards Achieving Trust Through Transparency and Ethics. In *2021 IEEE 29th International Requirements Engineering Conference (RE)*, pages 82–93.

Lai, C., Du, Y., Guo, Q., and Zheng, D. (2021). A trust-based privacy-preserving friend matching scheme in social Internet of Vehicles. *Peer-to-Peer Networking and Applications*, 14(4).

Li, W. and Song, H. (2016). ART: An Attack-Resistant Trust Management Scheme for Securing Vehicular Ad Hoc Networks. *IEEE Transactions on Intelligent Transportation Systems*, 17(4):960–969.

Litman, T. (2022). Autonomous Vehicle Implementation Predictions. Victoria Transport Policy Institute, Canada, https://www.vtpi.org/avip.pdf.

Myklebust, T., Stålhane, T., Jenssen, G. D., and Wærø, I. (2020). Autonomous Cars, Trust and Safety Case for the Public. In *2020 Annual Reliability and Maintainability Symposium (RAMS)*, pages 1–6, Palm Springs, CA, USA. IEEE.

Nyholm, S. and Smids, J. (2016). The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem? *Ethical Theory and Moral Practice*, 19:1275–1289. https://doi.org/10.1007/s10677-016-9745-2.

Ors, A. O. and Carlson, B. (2019). Safety and Security in Autonomous Systems and Gateways. *ATZelectronics worldwide*, 14:24–31. https://doi.org/10.1007/s38314-019-0077-6.

Publications Office of the European Union (2018). Regulation No 79 of the Economic Commission for Europe of the United Nations (UN/ECE) — Uniform provisions concerning the approval of vehicles with regard to steering equipment [2018/1947]. http://data.europa.eu/eli/reg/2018/1947/oj.

Sagar, S., Mahmood, A., Sheng, Q. Z., Pabani, J. K., and Zhang, W. E. (2022). Understanding the Trustworthiness Management in the Social Internet of Things: A Survey. *ArXiv*, abs/2202.03624.

Srinivasan, A., Teitelbaum, J., Wu, J., Cardei, M., and Liang, H. (2008). *Reputation-and-Trust-Based Systems for Ad Hoc Networks*, pages 375 – 403. Wiley-IEEE Press.

Wang, Y., Hu, X., Yang, L., and Huang, Z. (2022). Ethics Preference Modeling and Implementation of Personal Ethics Setting for Autonomous Vehicles in Dilemmas. *IEEE Intelligent Transportation Systems Magazine*, pages 2–14.

Watson, D. P. and Scheidt, D. H. (2005). Autonomous systems. *Johns Hopkins APL technical digest*, 26(4):368–376.

Weyns, D., Bures, T., Calinescu, R., Craggs, B., Fitzgerald, J., Garlan, D., Nuseibeh, B., Pasquale, L., Rashid, A., Ruchkin, I., and Schmerl, B. (2021). Six Software Engineering Principles for Smarter Cyber-Physical Systems. In *2021 IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C)*, pages 198–203, DC, USA. IEEE.

Windley, P. J. (2005). *Digital Identity: Unmasking identity management architecture (IMA)*. O'Reilly Media, Inc., Sebastopol, CA, USA.

Zhao, T., Yurtsever, E., Paulson, J., and Rizzoni, G. (2022). Formal Certification Methods for Automated Vehicle Safety Assessment. *IEEE Transactions on Intelligent Vehicles*.