SYNERGIES BETWEEN ROAD AND RAIL TRANSPORT IN THE DEVELOPMENT OF SAFE SELF-DRIVING VEHICLES

ALEŠ FILIP
Faculty of Electrical Engineering and Informatics, University of Pardubice, Czech Republic

ABSTRACT
In recent years, artificial intelligence (AI) has found numerous applications in medicine, energy, industry and various transport sectors, including rail and road. The use of AI for autonomous train operation is listed as one of the research challenges in the new Master Plan of the European Railway Joint Undertaking (October 2021). Nowadays, AI and machine learning (ML) algorithms are also widely used in connected self-driving cars (SDCs) for detection, classification and localization of objects on roads. Naturally, the rail industry also wants to benefit from recent advances in SDCs. While the current level of safety on the railways is acceptable to society, mass deployment of SDCs is expected to significantly reduce the number of accidents caused by human driver behaviour. Safety is thus currently a major challenge in the development of driverless cars. In contrast, various driverless automatic train operation (ATO) systems supported by automatic train protection with guaranteed high safety integrity level (SIL 4) have been introduced in the last decades, but mainly on segregated networks such as the metro. Therefore, the aim of SDC technology transfer is to go beyond segregated lines and develop fully autonomous driverless trains for open rail networks. In this paper, a comparative analysis was used to show how the required safety is assured in automated driving of trains and cars. The results of the analysis describe the differences, intersections and synergies in these two different application areas, in particular in terms of the basic pillars of safety, the safety standards and regulations used, interoperability requirements, safety demonstration, certification and independent assessment. Finally, the paper summarises how the rail experience in safety could be used to improve SDC safety, or conversely, how the ATO could benefit from transferring the latest AI and ML technologies developed specifically for SDCs.
Keywords: automatic train operation, autonomous vehicles, machine learning, self-driving cars.

1 INTRODUCTION

Railway is traditionally a very safe means of transport. This is because railway safety is based on conservative principles and railway stakeholders have many years of experience in building and operating various signalling systems. Initially, it was mechanical and electromechanical signalling. The breakthrough in the development of railway signalling systems came with the invention of computers and the improvement of radio communication systems. These new technologies not only provided the increased protection against human error and technical system failures, but also enabled the introduction of various degrees of automation on the railway. In addition to higher safety, this has also made it possible to increase the efficiency of railway operations. These automatic systems, which include on-board and trackside equipment, integrate all vital and non-vital functions to ensure safe and dependable train operation. In many cases automatic train operation (ATO) is driverless, especially on segregated rail networks such as the metro, commuter rails, airport lines or some heavy-haul lines. Recent initiatives in various countries around the world to introduce automated and autonomous driverless trains into operation on open rail networks have become a major challenge [1]. For example, autonomous trains (with driver) were put into service on the S-Bahn network in Hamburg in 2021.

The aim is to use artificial intelligence (AI) in combination with advanced sensors and powerful computers to increase punctuality of trains, track capacity (without laying new tracks), operational efficiency and reduce energy consumption. These are similar goals to

ISSN: 2058-8305 (paper format), ISSN: 2058-8313 (online), http://www.witpress.com/journals
DOI: 10.2495/TDI-V6-N3-313-325

those set for road transport 25 years ago, when intelligent transport systems were intended to make better use of roads and highways without the need to build new ones. Today efforts are focused on the use of machine learning (ML) algorithms and neural networks (NN) for autonomous vehicle control. The difference between an automatic/automated and an autonomous system depends on the degree of human intervention. An automated system performs tasks according to predefined rules, whereas an autonomous system can make independent decisions and react to a given operational situation. The term automated is also often used for systems that aim to achieve full or partial autonomy in the future.

AI has also started to be used extensively in recent years in automated driving systems (ADSs) for self-driving cars (SDCs). The reason for the use of AI in SDCs is that ADS must safely cope with millions and millions of different operational situations. Classical deterministic systems with well-traceable algorithms whose behaviour are well-predictable and therefore safe cannot be used in ADS. This is because deterministic systems usually have only a limited number of inputs at which their safe behaviour can be proven.

The safety of SDCs is the number one requirement. ADS with conditional or high driving automation (SAE automation levels 3 and 4) make sense if it is at least as safe as human driving. Today, ADSs can prevent many commonly occurring accidents, because the ADS is not tired, sleepy, distracted, tipsy, etc. However, so-called edge cases (rare dangerous events) may also happen that ADS cannot handle. This can be caused, e.g., by sudden changes in the operating environment, human misuse of the ADS (incorrect takeover of driving), limited sensor performance under fault-free conditions, etc. For these and other reasons, the safety of the current ADSs does not reach the level of safety that a human driver can provide. Liu et al. [2] showed that respondents to a recent survey expect future driverless cars to be four to five times safer than human-driven vehicles. It also means that SDCs should be safe as traveling by train or airplane. SDC safety will therefore need to improve in the coming years and railways should closely monitor the development of the latest ADS technologies and look for ways to use them on the railway.

In this paper, a comparative analysis was used to show how the required safety is assured in automated driving of trains and cars. The results of the analysis describe the differences, intersections and synergies in these two different application areas, in particular in terms of the basic pillars of safety, the safety standards and regulations used, interoperability requirements and safety demonstration. The focus of the paper is to compare railway and automotive safety concepts and standards in terms of autonomous driving, in order to understand where the advantages of railway safety lie, and which can be exploited when implementing autonomous driving. This in turn allows to identify some of the safety issues in automotive ADS and to propose measures for them, which have been in use on the railways for years.

This paper is organized as follows. Section 2 summarizes the state of the art and objectives in the field of automated train operation. The pillars and components of system safety are described in Section 3. Section 4 discusses the use of AI for automated driving. The differences between rail and automotive safety standards and concepts are shown in Section 5. The results from the comparative analysis are summarized in Section 6.

2 AUTOMATED OPERATIONS ON RAILWAY

Autonomous driving in land transport has been associated in recent years mainly with the rapid development of these technologies in automotive transport. However, it is often forgotten that highly automated vehicle control systems in rail transport are not new [1].

The first systems for ATO combined with automatic train protection (ATP) functionalities started to be used in the early 1970s on the Barcelona and London underground. In the

following decades, ATO systems with different grades of automation (GoA) began to be used not only on segregated metro lines in other world cities, but also on open rail networks. It should be noted that the ATO performs the functions of a driver, i.e. automatically drives trains through control of acceleration and braking while the ATP ensures the basic safety of train travel, e.g., avoiding accidents or speeding. ATP is a fail-safe subsystem, e.g., compliant with SIL 4 that supervises ATO [3].

The standard IEC 62290 defines 5 grades of automation from GoA0 to GoA4 with following operational features: GoA0 – on sight train operation under full driver responsibility, GoA1 – non-automated train operation (ATP with driver), GoA2 – semi-automated train operation (ATP and ATO with driver), GoA3 – driverless train operation with train attendant, and GoA4 – unattended train operation.

While the purpose of an ATO with GoA2 is to reduce energy consumption, GoA3 and GoA4 are designed to reduce overall costs. The GoA2 semi-automatic system is based on traditional safety and train control principles, with the driver supervising the correct functioning of the ATO. In a GoA3 system, the train cannot operate safely without the staff member on board, which is responsible, e.g., for door closure. Finally, in the GoA4 system, trains can operate automatically in all circumstances, including door closure, obstacle detection and emergency situations. On-board personnel may be provided for other purposes, such as customer service, but are not required for safe operation. Examples of the GoA4 system on separated lines include the Lille Metro (1983), line 14 of the Paris Metro, the Sydney Metro (2019), as well as Rio Tinto's mining rail network in Western Australia.

However, autonomous train control on open rail networks is still a major challenge, as tasks such as sensing and perception of operational environment, obstacle detection, person detection, safe train position determination, fault detection on rolling stock and rail infrastructure, etc. need to be addressed. These and similar tasks have been solved by the automotive industry for many years, and therefore it is possible that technologies originally developed for SDCs could also be used for autonomous trains.

Railways have one major advantage over the automotive industry when introducing autonomous systems – namely they already have the required safety in place. Current signalling systems can be used for this purpose, e.g., ERTMS/ETCS (European Railway Traffic Management System/European Train Control System) compliant with SIL 4. Individual autonomous applications using AI such as line-of-sight driving, obstacle detection, infrastructure inspection, etc. can be implemented on top of ETCS.

3 SAFETY PILLARS AND COMPONENTS

The safety of systems in industry, energy, transport and other areas is generally achieved by means of safety measures (barriers), designed to prevent the occurrence of a hazardous event with consequent damage. These measures may be technical, operational or organizational [4]. Technical measures are the physical means that must be used for a given application – e.g., the above-mentioned ATP system. Operational measures are actions and activities relating to the system operation, which tell what is to be done. This includes, e.g., control of a train or the operation of station interlocking equipment according to certain rules and procedures. However, the specification of activities alone is not sufficient. It needs to be determined who will carry out the activities. The locomotive is controlled by the driver, the signalling equipment is operated by the dispatcher, etc. This is the content of the organizational measures. Importantly, activities performed within technical, operational and organizational measures are not independent of each other, but are intertwined.

In the context of the design and implementation of technical and operational measures to protect against hazards (e.g. on railways), it is also necessary to consider how safety systems will be operated. This is done firstly by following certain rules and procedures in cases where possible hazards are foreseeable (known), and secondly by managing safety by qualified human operators of the systems in cases where not all hazards are known.

The part of safety that is achieved by best anticipating potential hazards and putting in place rules and means to manage them is called rule-based safety (S_R). The goal of the S_R component is to avoid all foreseeable hazards and thus reduce the frequency of hazards. This is achieved through expertise, technical barriers, rules and procedures. The rules are usually implemented using SW programs in the technical system (e.g. ATP) and are also used by human operators to operate the system. In addition to this, the professionalism and responsibility of the human operator itself is also important for the safe operation of the system – especially in cases where not all hazards and associated hazardous events can be anticipated. This component of safety is called managed safety (S_M) and is based on the skill, experience, learning ability, training and adaptation of the human operators involved in managing the process and safety in real time. They identify the actual situation and react appropriately. The purpose of S_M is to reduce the consequences of accidents. Reducing the frequency of hazards and consequences of the accident by external measures is illustrated by the example of the risk model in Figs. 7 and 8 in Part 2 of EN 50126:2017 [5].

The ratio between S_R and S_M components in the overall system safety depends on the application. Some applications are more regulated, such as nuclear power, aerospace, etc., and therefore the S_R component dominates over the S_M. On the other hand, in applications such as sea fishing or disaster medicine, the S_M component dominates because many regulations would be a constraint to business or services. And which safety component will prevail in rail transport with automated control? The key to the answer is the operating environment. For transport systems operating in a closed environment, such as ATO systems on metro lines, the operating environment is well known. The designers of these safety systems anticipate all possible hazards and dangerous events that could occur and design measures against them in the form of safety functions in the system and rules for operating the system. Therefore, the S_R component prevails. That is why these automatic traffic systems have been implemented in many cities, and with a high level of safety. In contrast, the operating environment on the open rail network is much more complex and not all operational situations and related hazardous events can be foreseen and therefore the share of the S_M component will be higher. In the case of ATO systems in harsh environments, the technical system will also have to provide this part of safety. It must trigger a safe response even when rules are not available. Here one can see an analogy with SDCs in the rare dangerous operational situations that arise from so-called edge cases.

4 ARTIFICIAL INTELLIGENCE FOR CARS AND TRAINS

The fundamental difference between the concept of safety in road and rail transport is as follows. The driver of a road vehicle can drive at any time, unless this is forbidden in any way, e.g. by a traffic sign or lights, by order of a traffic police officer, etc. In contrast, a train can only travel from point A to point B if it is allowed by technical system or dispatcher to do so – i.e. if it receives a Movement authority (MA).

In railway signalling systems up to the current ones based on computers and advanced communications, the rule-based safety component has always prevailed over the managed safety component (S_M). This has been made possible by the relatively simple operating environment due to the reservation of a path by granting an MA for a given train.

4.1 Reasons for using machine learning in self-driving cars

The situation is completely different in road traffic, where the car driver is obliged to carry out all actions in accordance with the rules of the road and to consider all other road users (cars, pedestrians, cyclists) and weather and other conditions (e.g., obstacles, children or animals on the road). The driver uses his/her senses when driving, and of course uses his/her own experience where rules are missing.

If the human driver is to be replaced by an automated driving system (ADS), then the ADS must cope with millions and millions of different operational situations due to the complex operating environment. It is obvious that driving a car with such a huge number of different operational situations cannot be implemented in the form of rules, and therefore the use of AI, different ML software, models and algorithms, e.g., deep NN, have been adopted in the field of ADS.

ML algorithms perform two basic tasks: (1) they process the huge amount of previously recorded data from the car's sensors (e.g. cameras, ultrasonic sensors, GNSS, radar, LiDAR) or data simulators and use these to teach the computer (ML model) with the intention of doing things or learning things as a human driver can and (2) based on prior learning and sensor data acquired in real traffic, they detect, classify, and locate objects in the car's surroundings under all possible traffic situations on the road, as needed for automated driving purposes.

In the last 15 years, ML and NN have found applications in many areas such as medical diagnosis, nuclear energy, industrial production and are also extensively used in a wide range of advanced driver assistance systems (ADAS) with SAE Level 2 and ADSs with Level 3 or higher from different car manufacturers. The question is whether ML can actually replace traditional computer vision algorithms and whether ML models have sufficient potential to ensure the required ADS safety.

4.2 Current limits of ML models for safety applications

At present, one thing is certain – ADS does not provide the level of safety that the average human driver is capable of. Therefore, in certain traffic situations that ADS cannot cope with, it is necessary for a human driver to take over the control of the car from the ADS for a certain time interval or to supervise the correct functioning of the ADS to ensure the safe vehicle state in case of danger – e.g. to stop it. ML technologies, and NN in particular, can learn to recognize patterns, such as facial expressions indicating fatigue or inattention, and so can be used to monitor the human driver for when the driver should take over control of the vehicle from the ADS. Driver monitoring therefore contributes significantly to the safety of SDCs.

The relatively frequent reports coming from various countries of serious accidents caused by ADS failures are indicative of the safety problems of SDCs. What are the main causes? Disengagement accidents which are related to the incorrect handover of vehicle control to a human driver represent a major problem of SDCs. Incorrect outputs from the perception algorithms (e.g. missed detected or wrongly classified objects or traffic signs) have been a major cause of disengagement incidents in SDCs.

The development of classical safety systems essentially involves three basic steps: specifying the system requirements, developing the system (using rules), and proving that the system meets the requirements. Traditional rule-based algorithms are predictable. There are written according to complete specifications. The computer expert developing the SW specifies the key parameters that are needed for decision making. In an algorithm, it is possible to trace the path from input to output and understand all the operations and decisions behind a given

result. Also, comprehensive testing of algorithms and other parts of the system is feasible because traditional safety systems generally have a limited number of inputs.

ML models in SDCs are different. They learn to detect and classify objects from the vast amount of sensor data acquired in the car – corresponding to millions of real traffic situations and billions of simulated situations. The ML model can only make decisions of the quality of the data that the ML model learns with. Although NNs are powerful methods for performing complex tasks compared to humans, they are extremely sensitive to natural noise and to small errors in the test data.

On the other hand, a huge advantage of the ML model is that, unlike the traditional safety, no rules or algorithms defined from requirements are needed. For this reason, the ML model must be considered a black box – it is not possible to trace back from the result at the output of the ML model where the error occurred. This is because the result does not depend on one input, but on all previous inputs to the NN. Therefore, we need to understand how the NN can fail and quantify that failure. Without this, it is impossible to talk seriously about NN applications in safety systems such as SDCs. In addition, ML models must be developed in an application environment (SW) that has been certified for safety systems development. And last of the main issues at the end. The accuracy (certainty) of object classification is usually around 90%, and for well-trained NNs, even 97% accuracy can be achieved. Even if the uncertainty of NN decision making in a car would be about one order of magnitude lower (e.g., 0.1%), the SDC would still be a dangerous machine. To achieve higher safety, data from several different sensors can be combined and independent diagnostics (safety monitor) can be used, which on the other hand considerably limits the much-desired flexibility that ML models provide. The applicability of AI in transport safety systems in terms of safety standards is discussed in the following section.

5 SAFETY STANDARDS FOR AUTOMATED DRIVING

5.1 IEC 61508

IEC 61508 [6] is a basic functional safety standard applicable to safety-related systems in all industries that incorporate Electrical and/or Electronic and/or Programmable Electronic (E/E/PE) devices. It is also the parent standard that has been used to create application-specific safety standards such as EN 5012x [5,7,8] for railways, ISO 26262 [9] for automobiles, IEC 61511 for a process industry, etc.

The fundamental safety concept according to IEC 61508 is that any safety-related system must work correctly or fail in a predictable (safe) way. This safety standard specifically covers hazards that occur when safety functions fail. The main objective of IEC 61508 is therefore to reduce the risk associated with a hazardous failure to an acceptable level. IEC 61508 is built on two fundamental pillars: (i) the safety lifecycle intended to reduce or eliminate failures due to systematic causes during system development and operation and (ii) the probabilistic failure approach to address dangerous random HW failures via safety integrity levels (SILs). This concept is strengthened by the fact that the system must be developed, validated and assessed according to specific requirements which result from the hazard identification and risk analysis. IEC 61508, on the other hand, does not cover in detail the effects of human factors on safety during operation as this goes beyond functional safety.

Predictable behaviour of the system can be achieved in case of both systematic and random failures. Systematic failure is related in a deterministic way to a certain cause, which can only be eliminated by a modification of the design (rules) or of the way the manufacturing process,

operational procedures, documentation or other relevant factors. In case of random failures, the required predictability of the system can be achieved through probabilistic description of the system behaviour. Deterministic means that each event or state is the result of previous events on the principle of causality and fixed rules. Causality and rules are necessary for the predictable behaviour of a system in the event of its failure.

The advantage of AI-based systems for automated driving is their great flexibility, as they can handle tasks involving a huge number of traffic situations using big-data from the car's sensors and learning, without specifying rules for the system's behaviour. However, the absence of rules (probabilistic model description) is on the other hand a major problem for safety assessment of systems based on. IEC 61508 does not recommend the use of AI in systems for fault correction at SIL 2 and above – see Table A.2 in IEC 61508-3 and §3.9 in IEC 61508-7. Nor does this standard take any position for AI applications according to SIL 1. Without a sufficiently described statistical model used by the AI application, it will not be possible to meet the basic safety concept of IEC 61508, which is based on the predictable behaviour of the system in the event of a fault.

5.2 Railway safety standards

The basic framework for ensuring the safety and dependability of railway systems is defined in EN 50126 [5] on the specification and demonstration of RAMS (Reliability, Availability, Maintainability and Safety). EN 50126 considers the railway system in a given physical and operational environment, i.e., including human operators, as well as the factors that influence the railway RAMS – in particular the technical system and the operational and maintenance conditions. The standard specifies in detail the different phases of the system life cycle, i.e. including the role of the human factor in them, and also prescribes methods for managing the RAMS within the system life cycle. Safety shall be demonstrated by means of safety case and independent third-party assessment. The basic framework defined through RAMS can be imagined as an umbrella (Fig. 1) under which a safety-related system is subsequently developed and implemented according to the downstream standards EN 50129 [8] (safety-related system), EN 50 128 [7] (software for safety-related system) and others.

A safety case and its independent assessment alone is still not enough to ensure safety on European railways. Technical interoperability must also be ensured (Fig. 1). In the case of ERTMS, e.g., this means that one manufacturer's on-board equipment works correctly with another manufacturer's track-side equipment. Therefore, certification according to the Technical Specifications for Interoperability (TSIs) must be carried out. But even this may not be enough to ensure safety. In the case of a significant change in the railway system from a safety point of view, the so-called Common Safety Method for Risk Evaluation and Assessment (CSM-RA) according to the Regulation (EU) 402/2013 [10], which harmonises the risk assessment process and safety requirements, must be applied. The safety concept of EN 50129, as well as IEC 61508, is based on the predictable (safe) behaviour of the system in the event of a failure. A causal analysis, i.e. an analysis of the reasons how and why a particular hazard can come into existence, is therefore important part of hazard analysis.

A safety-relevant system is designed for a specific operating environment and therefore the rules for its operation and maintenance as well as external influences (such as climatic, mechanical, electrical, IT-security, etc.) must be clearly defined. The conditions, rules and constraints for the design, manufacture, installation, operation and maintenance of the system (ensuring functional safety) and the way to verify them shall be contained in the document 'Safety-related Application Conditions (SRACs)' according to EN 50129. The safety and

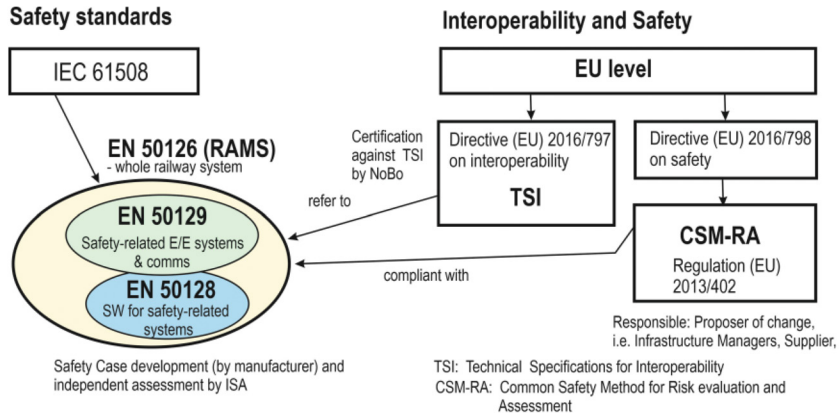

Figure 1: Railway safety standards, interoperability and common safety method.

reliability of system operation with external influences shall be demonstrated in the document 'Operation with External Influences (OWEIs)'. Both documents are part of the safety demonstration. The safety case is valid only within the specified range of external influences, as defined in the system requirements specification.

According to EN 5012x and IEC 61508, the cause of a failure due to the operational environment is a systematic fault in the system design. In contrast to the automotive ISO 26262, any malfunction of the intended functionality of an automotive system due to complex operating environment or gaps in the requirement specifications, is out of scope of ISO 26262 (functional safety) and should be covered by the standard ISO/PAS 21448 (SOTIF) [11], as it is shown below. EN 5012x similarly as IEC 61508 does not recommend using AI as a technique for fault correction for any SILs – see Table A.3 and section D.1 in EN 50128.

5.3 Automotive safety standards

5.3.1 ISO 26262, ISO/PAS 21448 (SOTIF) and UL 4600

A safe ADS means that all hazards associated with ADS operation are fully under control using safety functions with the required safety integrity. The basic functional safety standard used for development and safety demonstration of ADS is ISO 26262. It is an adaptation of the IEC 61508 functional safety standard for automotive Electrical/Electronic (E/E) systems. ISO 26262 aims to eliminate potential hazards caused by malfunctioning E/E systems in vehicle. Malfunctioning behaviour of the system is caused by a failure or unintended behaviour of the system with respect to the intended design. Risk of hazardous operational situations is qualitatively assessed by means of automotive safety integrity levels. Safety measures are defined to avoid or control systematic faults and to detect or control random hardware failures or mitigate their effects.

ISO 26262 covers functional safety of automotive E/E equipment in the event of HW failures and SW faults throughout the life cycle equipment. However, this standard does not apply to vehicle safety in the absence of E/E equipment failure, e.g., in the event of ADS malfunction due to human driver error or unforeseen changes in a complex operating environment. This has led the automotive industry to start addressing hazardous behaviour of systems caused by insufficiencies in the system design and limitations in system performance.

Therefore, the ISO/PAS 21448 standard [11] was developed and is referred to as SOTIF (Safety Of The Intended Functionality). The purpose of SOTIF is to mitigate: (1) risk due to unexpected operating conditions including incorrect user (human driver) behaviour and (2) insufficiencies in requirements specifications. This standard focuses mainly on design guidelines and procedures for validation and verification (V&V) to reduce the residual risk associated with hazards under fault-free (but not error-free) conditions. Safety issues are then resolved by functional modifications.

The system safety according to IEC 61508 or EN 5012x is based on the fact that the behaviour of the system in the event of a failure is predictable. However, this is not the case for ML algorithms, which are considered a black box, because by the nature of ML it is not easy to know what is going on inside. ML for ADS purposes is still under research and there is no technical solution for which the required (high) safety can be demonstrated. SOTIF is mainly aimed at reducing risks in cases referred to as unknown/unsafe. Unknown means a hardly anticipated operational situation and unsafe means the presence of hazards in the system due to limitations of the intended functionality under fault-free conditions.

ISO 26262 and ISO/PAS 21448 prescribe how to design, verify and validate a safety system. Important part of safety demonstration is safety case development and its assessment by an independent third party. More detailed information on V&V and safety case development can be found in the US national standard UL 4600 (Evaluation of Autonomous Products) [12], which prescribes in particular what the safety case for autonomous products should focus on and how the safety case should be assessed. UL 4600 is based on previous automotive standards ISO 26262 and ISO/PAS 21448 and is intended for autonomous driving with SAE Levels from 3 to 5. UL 4600 does not prescribe which technologies or architectures should be used (although it considers the use of ML to be very promising), but on the other hand it does require that the safety case must convincingly argue for the safety claims of the ADS, especially based on analysis, simulation, laboratory testing and testing on public roads.

5.3.2 ISO/TR 4804

A function mitigating risk can be considered safe if ISO 26262 (functional safety) and ISO/PAS 21448 (SOTIF) standards are applied. However, vehicles cannot be in a safe state without secure operation. To cover the whole area of ADS safety, a technical report ISO/TR 4804 (road vehicles – Safety and cybersecurity for automated driving systems – design, verification and validation) was developed [13]. The intention of ISO/TR 4804 is to put together standards ISO 26262 (functional safety), ISO/PAS 21448 (SOTIF) and ISO SAE 21434 (cyber security) under one risk-based approach. It considers safety and cyber security by design, as well as verification and validation methods for ADS with SAE Levels 3–4.

6 SYNERGIES BETWEEN ROAD AND RAIL TRANSPORT

Safety is the most important quality attribute of transport systems. Therefore, possible synergies in the development of ATO and ADS should be discussed primarily in relation to this attribute. While rail and air transport are among the safest modes of transport, the safety of ADS does not currently reach the safety of the average human driver. It is thus natural to discuss how the railways' long experience in safety could be used to improve SDC safety, or conversely, how the railways could benefit from the transfer of the latest AI and ML technologies that have been developed specifically for SDC.

As a first example of the use of synergies between road and rail, we can briefly mention the procedure for estimating a harmonised safety target for self-driving vehicles, which

was developed in the framework of the PosiTrans and HELMET projects [14]. Harmonised safety targets exist for aviation or rail, but none for SDC. Such a broadly acceptable safety target could e.g., outline which level of safety is really needed for SDCs and for this seek appropriate procedures for proving safety. The setting of a safety target could, e.g., be based on the results of a public survey that found SDCs to be as safe as train or air travel [2]. This corresponds to a fatality risk of about 3e-8 deaths/h [15]. Using road accident statistics, this fatality risk can be converted into a probability of dangerous failure for a single car. Now, to harmonise the resulting safety target, a railway Common Safety Method – Design Targets (CSM-DT) can be used [16] . The design targets (DT) represent harmonised safety requirements for the technical system and is fully consistent with the aeronautical safety target levels. If a very small number of persons are affected by an accident and there is one fatality, which is the case for an average fatal car accident, then the DT for a SDC corresponds to 1e-7 dangerous failures per 1 h [14].

A major advantage of railway communications-based safety-related systems is that the required safety integrity level of technical systems (SIL 4) has already been achieved – e.g. in the form of interoperable ERTMS/ETCS. Autonomous train operation (ATO GoA 3 and 4) on open rail networks can be built as a superstructure on top of ETCS. ATO with GoA3 and GoA4 requires track obstacle detection functionality. This and additional functionalities supporting autonomous train operation, such as track environment perception, person detection, supervision of passenger boarding and getting-out, train localization, remote/driverless driving at low speed at depos, punctuality of driving etc. are intended to be solved using AI and ML algorithms.

When AI is to be used for safety-related applications, it must be demonstrated that these applications are safe and meet the required safety standards. Until now, safety systems with predictable failure behaviour in the sense of EN 5012x and IEC 61508 have been used on the railway. Similarly, the behaviour of an AI system must be predictable in a statistical sense. Here we are not dealing with deterministic behaviour but with statistically predictable behaviour. It means that there are measures against systematic failures (based on causality rules) and measures against random failures by describing a statistical model using rules.

AI learns from large amounts of sensor data and makes decisions based on this learning. From this perspective, AI algorithms are considered a black box because we often do not know what's going on inside. Therefore, a correct statistical model of AI algorithms is needed to generate data regarding the process control based on known rules. Introducing causality into an AI-based statistical model would then enable safety analysis.

The safety standards used in rail transport are generic, not technologically oriented. EN 50126 defines the basic framework for specifying and demonstrating safety and reliability in the sense of RAMS. The influence of the human factor (e.g., the operator) on the safety and dependability of the whole railway system is considered at each stage of the life cycle. EN 50126 is subsequently followed by EN 50129 and EN 50128, which are used to implement the safety-relevant system (Fig. 1).

The process of developing automotive safety standards was different. The 1st edition of ISO 26262 was released in 2010 and the 2nd edition in 2018. However, it covers functional safety, i.e., only part of the overall safety. In ISO 26262, very limited attention is paid to the user's influence on system safety. In the use of SDCs, human misuse of the ADS system, as well as difficult-to-predict changes in the physical and operational environment, has proven to be one of the common causes of accidents. Therefore, ISO/PAS 21448 (SOTIF) was released in 2019 as a complement to ISO 26262 to address the above missing points in ISO 26262.

Finally, ISO/TR 4804 on safety and cybersecurity for ADS was released in 2020, which ties together ISO 26262, ISO/PAS 21448 and ISO SAE 21434. ISO/TR 4804 defines a safe and secure function, which means a dependable function. Dependability as defined in ISO/TR 4804 includes Reliability, Availability, Maintainability, Safety and Security (RAMSS). This differs from EN 50126, where dependability includes only RAM attributes.

The automotive safety standards are technology-oriented, as e.g., the purpose of SOTIF is also to support the use of ML algorithms for ADS purposes. However, as mentioned, ML algorithms cannot yet be used for safety applications on railways because their behaviour is not predictable.

If the safety function needs to achieve the required safety integrity, information from two or more independent sources is combined – e.g., within a one-out-of-two (1oo2) or two-out-of-two (2oo2) architecture. IEC 61508 states that the 1oo2 architecture is intended for safety integrity and the 2oo2 for availability. In contrast, railway standard EN 50129 says that 2oo2 is used for integrity and 1oo2 for availability – i.e., the opposite statement. In the literature on functional safety in the automotive industry, the difference between the two architectures is often not stated. Recently, it was explained [17] that the main distinguishing features are the reference diagnostic for the 1oo2 and the external comparator for the 2oo2.

The last example of synergy given in this section concerns cross-acceptance and certification of the EGNOS satellite navigation system used for safe vehicle location in land transport – in particular for ERTMS and SDCs. EGNOS has been designed for safety operations in air transport and therefore the EGNOS safety case developed for aviation cannot be used to certify an EGNOS-based ERTMS according to EN 50129. It has been suggested that EGNOS could be cross-accepted for ERTMS via 'pre-existing' item in the sense of IEC 61508 and EN 50129 standards [18]. All essential information on EGNOS and its integration into ERTMS would be contained in the EGNOS Safety Manual. It is a guidance for designers and system integrators. It is envisaged that EGNOS could also be reused for SDCs, e.g., as a 'proven in use argument' according to ISO 26262. The reuse of EGNOS creates scope for collaboration between rail and automotive sectors.

7 CONCLUSIONS

Recent advances in autonomous driving are usually associated with the development of these technologies in automotive transport. And it is often forgotten that the first automated vehicle control systems in rail transport were put into operation on segregated lines half a century ago. Today, the focus of rail research is on ATO systems with a high grade of automation (GoA3 and GoA4) to ensure autonomous operation on open rail networks. The aim of ATO as a superstructure over ATP is to increase the efficiency of operations. In contrast, the goal of automotive ADS is to increase SDC safety, as it does not currently achieve the safety of the average human driver. The key technologies used by automotive ADS are based on AI, ML algorithms, advanced sensing and environment perception.

The paper suggests how e.g., the rail expertise in safety can be used to specify a harmonised safety target for SDCs or how the rail and automotive industry could work together on cross-acceptance and certification of the EGNOS satellite navigation system for safe vehicle location. Further, the main differences between railway and automotive safety standards were presented and the discrepancies between the 1oo2 and 2oo2 architectures according to IEC 61508 and EN 50129 were explained. The automotive SOTIF concept could also be very useful for ATO development based on AI and ML.

However, it should not be forgotten that current ML models do not yet allow to generate data according to known rules. The behaviour of such ML models is not predictable and therefore it is not possible to perform a safety analysis and demonstrate safety in the sense of EN 5012x standards. Developments in this area need to be further monitored.

ACKNOWLEDGEMENTS

The work was supported from: (1) the ERDF/ESF grant 'Cooperation in Applied Research between the University of Pardubice and companies, in the Field of Positioning, Detection and Simulation Technology for Transport Systems – PosiTrans (2018–2022)', No. CZ. 02.1.01/0.0/0.0/17_049/0008394 and (2) the H2020 HELMET project (2020–2022).

REFERENCES

[1] IRRB Webinar Autonomous technologies in rail – Anticipating expectations, 9.6.2021. Online, https://uic.org/events/IMG/pdf/ato_webinar.pdf. Accessed on: 2.3.2022.

[2] Liu, P., Yang, R. & Xu, Z., How safe is safe enough for self-driving vehicles? *Risk Analysis*, ***39(3)***, pp. 315325, 2018.

[3] Erskine, M., et al., Digital train control functional safety for AI based systems. Presented at *the International Railway Safety Council Conference*, Perth, Australia, 2019.

[4] Richard, P., Boussif, A. & Paglia, Ch., Rule-based and managed safety: a challenge for railway autonomous driving systems. *Proc. of the 31st European Safety and Reliability Conference (ESREL)*, pp. 2363–2369, Angers, France, 2021.

[5] EN50126 (1–2), Railway applications – The specification and demonstration of reliability, availability, maintainability and safety (RAMS) – *European standard*, 2017.

[6] IEC 61508 (1–7), Functional safety of electrical/electronic/programmable electronic safety-related systems, *European standard*, 2010.

[7] EN50128, Railway applications – communication, signalling and processing systems – software for railway control and protection systems. *European standard*, 2011.

[8] EN 50129, Railway applications – safety related electronic systems for signalling. *European standard*, 2018.

[9] ISO 26262 (1–10), Road vehicles – Functional safety. *International standard*, 2018.

[10] Regulation (EU) No. 402/2013 of 30 April 2013 on the common safety method for risk evaluation and assessment and repealing Regulation (EC) No. 352/2009.

[11] ISO/PAS 21448, Road Vehicles – Safety of the intended functionality (SOTIF). *International standard*, 2019.

[12] UL 4600, Standard for safety – evaluation of autonomous products. *American National Standard*, 2020.

[13] ISO/TR 4804, Road vehicles – Safety and cybersecurity for automated driving systems – design, verification and validation. *Technical report*, 2020.

[14] Filip, A., et al., D2.3 System requirements specification. H2020 HELMET project, 2020. Online, https://www.researchgate.net/publication/342673852_HELMET_SYSTEM_REQUIREMENTS_SPECIFICATION. Accessed on: 16 May 2022.

[15] Aviation safety. Online, https://en.wikipedia.org/wiki/Aviation_safety. Accessed on: 16 May 2022.

[16] Jovicic, D., Guideline for the application of harmonised design targets (CSM-DT) for technical systems as defined in (EU) Regulation 2015/1136 within the risk assessment process of Regulation 402/2013. European Union Agency for Railways, 2017. Online,

https://www.era.europa.eu/sites/default/files/activities/docs/era_gui_harmonised_design_targets_en.pdf. Accessed on 16 May 2022.

[17] Filip, A. et al., Clarification of Discrepancies in the Classification of 1oo2 and 2oo2 Architectures Used for Safety Integrity in Land Transport. *Proc. of the 31st European Safety and Reliability Conference (ESREL)*, pp. 2172–2179, Angers, France, 2021.

[18] Filip, A., Certification of EGNOS Safety-of-Life service for ERTMS according to IEC 61508 and EN 50129. *Proc. of the COMPRAIL 2020 conference – Computers in Railways XVII*, 1-3.7.2020, vol 199, on-line, pp. 115–125, 2020.