Conceptualizations of Safety and Risk in Current Automated Driving Regulation – Recommendations and a Critical Review*

Marcus Nolte[†‡], Leon Johann Brettin[§], Marvin Loba[§], Hans Steege[¶], Nayel Fabian Salem[§], Robert Graubohm[§], and Markus Maurer[§]

Abstract: "Safety" and "risk" are key concepts in the context of automated vehicle deployment. A clear conceptualization of both concepts is not only necessary for safety engineers who are responsible for assurance processes. For the communication with and between technical and non-technical stakeholders (engineers, regulators, lawyers, the general public) in particular, common, or at least compatible, notions of these abstract concepts are crucial for effective communication and for setting mutual expectations.

In the European market, automated vehicles require type approval or test permits which are subject to current regulatory efforts. The resulting regulatory documents are a central means of communication between regulators and the engineers designing and developing the technical systems. Flawed terminology regarding regulatory safety expectations for automated vehicles can unnecessarily complicate relations between regulators and manufacturers and thus hinder the introduction of the technology. In this paper, we set a focus on the European market and review relevant documents at the UN- and EU-level, for the UK, and Germany regarding their (implied or explicit) notions of safety and risk. We map regulatory notions to established as well as more recently developed notions of safety and risk in the field of automated driving, pointing out potential conflicts. Based on the analysis, we provide requirements and recommendations for establishing clear definitions of safety and risk in regulations. Furthermore, we discuss how applying such clear definitions and acknowledging their consequences can support rather than hinder the market introduction of automated vehicles.

*Parts of this work were supported by the German Federal Ministry for Economic Affairs and Climate Action within the project "Automatisierter Transport zwischen Logistikzentren auf Schnellstraßen im Level 4 (ATLAS-L4)" (FKZ 19A21048H). This current version of the paper is a revision of a preprint, extended by recent regulatory changes, and revisited aspects in the mapping of safety and risk notions, as well as delivering concretized requirements and recommendations for regulation and regulators (`https://arxiv.org/abs/2502.06594`). The work leading to this revision was supported by the German Federal Ministry of Transport (BMV) within the project "Automatisierte (Level 4) und teleoperativ unterstützte Lkw-Verkehre im GVZ Bremen (AUTOGVZ)" (FKZ 45AVF8F061).

[†]Corresponding author: mnolte@kth.se

[‡]KTH Royal Institute of Technology, Department of Engineering Design, Unit Mechatronics

[§]Technische Universität Braunschweig, Institute of Control Engineering

[¶]University of Stuttgart, Institute of Economics and Law

1 Introduction

The market introduction of automated driving systems of SAE-Level 3 and 4 in the European Union and the UK requires active regulatory acts to enable and implement type approval, commercial operation, as well as public-road testing. The type approval system is often referred to as a way to ensure a "safe" market introduction of new technologies, while having a reputation of potentially slowing down innovation processes.

In this context it is important to highlight the inherent dual role of regulation: protecting the public interest and supporting public safety is only one aspect. More importantly, regulation always must fulfill an enabling role, creating "opportunities of impact." [1, 2]

For automated vehicles, the European need for a-priori enabling regulation creates three main challenges: First, regulatory bodies must create laws *before* a technology can be introduced in the regulated market. This can become a competitive disadvantage when compared to target markets, such as the US, that employ a system of self-certification. Second, regulatory bodies must establish an appropriate level of technical expertise. This ensures that the enacted laws and regulations support the introduction of new technologies, rather than hindering them unintentionally due to lacking technical feasibility. Third, at least from a technical perspective, the laws and acts should be written in a way that provides sufficient legal certainty for companies developing automated vehicles by providing a coherent legal framework while avoiding overregulation.

This paper focuses on the third challenge introduced above. While regulation in general can be concretized in a variety of ways, we see two options that seem viable to enabling a coherent legal framework regulating automated vehicles: one option is to provide concrete metrics and checklists that must be fulfilled by manufacturers to prove that the systems are "safe enough". A second way of concretization is to introduce concise terminology regarding the concepts of "safety" and "risk": The goal of the European type approval system is to require companies to prove that their systems are "safe" *before* market introduction. A clear definition of what is legally understood by "safe" would be a logical requirement in this respect.

However, conceptualizations of "safety" and "risk" often carry ambiguity: In everyday language, "safety" and "risk" constitute so-called "open" or "empty signifiers" [3, 4] (cf. section 2) which means that the words do not carry precise semantic meanings in general language. Different stakeholders in transdisciplinary (or even inter- and intradisciplinary) discussions may thus attribute their own understanding to both terms. In consequence, stakeholders involved in discussions can seemingly agree on the importance of "safe" automated vehicles while having different understandings of what "safe" actually means. Concretizing regulation would hence require clear and consistent definitions of such open signifiers to ensure a common understanding of terms between regulators and the stakeholders who are subject to the regulation.

Having stated this need for compatible conceptualizations, even in engineering disciplines, established technical notions of "safety" and "risk", have lately come under scrutiny [5]. Technical definitions of "safety" and "risk" are often derived from hazards to the health and life of humans (in short: physical harm). However, as automated vehicles are part of a larger sociotechnical system of systems [5, 6], there are arguments that even technical notions should, e.g., include ethical or legal considerations that go beyond physical harm.

We want to raise awareness of potential issues that can be caused by differing (implied) understandings regarding the concepts of "safety" and "risk" and provide recommenda-

tions and requirements to regulators how regulation can be improved by providing clear terminology.

In this paper, we give an overview of current discussions regarding the concepts of "safety" and "risk" in and beyond the domain of automated driving. We map conceptualizations of "safety" and "risk" in current European regulation (Regulation (EU) 2022/1426, The 2024 UK Automated Vehicles Act, and the German Implementing Regulation (AFGBV)) to established notions in relevant standards and literature. Based on the mapping, we provide requirements and recommendations for making regulation targeting the test and operation of automated vehicles more concise from a technical point of view to contribute to reducing regal uncertainty for developers.

2 Related Work

The need for clear terminology when it comes to communication about "risk" and "safety" is not unique to the field of automated driving. A core concern in the fields of risk research and risk communication [7] is to establish well-grounded, ideally commonly understandable [8], terminology between different stakeholders, e.g., the general public, regulatory bodies, policymakers, industry, or public institutions such as non-governmental organizations (NGOs) [7, 9, 10]. In this context, Fischhoff *et al.* demand "[...] an explicit and accepted definition of the term 'risk' [...]" [9, p. 123], noting that the definition of "risk" is "inherently controversial" [9, p. 124].

Christensen *et al.* note that deviating terminology between stakeholders can derail discussions from their "core issue(s)" [10, p. 182]. They specifically include terminology that is related to "identifying, estimating, regulating, and communicating risk" [10, p. 182], hence including all the above-mentioned stakeholders in their argument. The authors analyze several references from different regulatory bodies and NGOs such as the European Commission, the UN/OECD, the US-EPA, or ISO/IEC (specifically ISO/IEC Guide 51 [11]). Christensen *et al.* [10] discuss, explain, and clarify applications of terms and views related to risk associated sciences to facilitate communication between stakeholder groups. By not providing a fully consolidated terminology, they acknowledge that, while there must be a fundamental consensus among stakeholders, communication always has to be tailored to the communicating parties. In the following sections of the paper, we mainly consider regulators communicating to stakeholders who are implementing automated driving technology.

The need for clear and consistent communication about risk related to autonomous systems in general is highlighted in [12]. The report emphasizes that it is crucial to identify *who* communicates *how* about *what* related to the communication about safety-critical autonomous systems. Similarly to the discussion in [10], the authors explicitly stress that messages related to safety assurance and their content should be tailored to the relevant audiences (i.e., *to whom* shall be communicated). According to [12], this is particularly important for calibrating the expectations of different stakeholders and for raising awareness for the limitations of autonomous systems' capabilities.

Different concepts of "risk" and "safety" for automated driving systems have, e.g., been discussed in [5] and [6]. Koopman and Widen [5] review automotive safety standards (ISO 26262:2018 [13], ISO 21448:2022 [14], ANSI/UL4600 [15]) and additional resources such as the German Ethics Commission [16] or the US National Highway Traffic Safety

Administration for their conceptualizations of "safety". They provide additional examples of "safety problems" related to automated driving systems which are not covered by the purely technical definitions assumed in ISO 26262 and ISO 21448. The authors point out that risk and safety for automated vehicles should be discussed in a more nuanced way beyond considering technical definitions grounded in "net (physical or monetary) harm".[1]

The line of argumentation is largely compatible with previous work [6, 17], discussing a broader view of notions of "safety" and "risk" which focus on ethical questions and the protection of stakeholder values. Building on a similar assessment of technical (automotive) standards as Koopman and Widen, [6, 17] include general safety and risk management standards such as IEC 61508 [18], ISO 31000 [19], or ISO Guide 51 [11] to avoid narrow technical definitions. Particularly, [6, 17] define "risk" and "safety" in relation to prioritized stakeholder values[2].

Such wider notions of safety and risk acknowledge that technology acceptance and adoption is a complex societal process [20] in which people are willing to accept certain risks if the benefits of a technology justify it.

Considering this body of related work, several distinctions can be made when assessing conceptualizations of "safety" and "risk": Notions of "risk" can, e.g., be characterized by underlying "risk sources",[3] or the considered types of "harm"[4]. Notions of safety can be classified by their extent of absoluteness: Such could range from notions of risk which expect the *absence of any risk* to more relative notions of risk that accept cost-benefit, or value-based, tradeoffs.

3 Safety and Risk – Terminology

In the following, we take the needs for clear terminology that have been identified in Section 2 and derive categories for the analysis in Section 4.

As "safety" is typically defined via "risk," an independent discussion of both concepts is necessary. In this respect, safety is, in most disciplines, understood as the "freedom from (unacceptable/not tolerable/unreasonable) risk". Nonetheless, a nuanced consideration of the definitions is crucial for understanding the consequences of uncareful formulations in regulatory contexts.

The most general "risk" definition is the "effect of uncertainty on objectives" [19]. In technical domains, this uncertainty is often reflected in terms of *probability*. The effect on objectives is considered as *harm* with a certain *severity*. This yields a technical definition of "risk" as the "combination of the probability of occurrence of harm [...] and the severity of that harm" [11, p. 2, def. 3.9], with "harm" being the "injury or damage to the health of people, or damage to property or the environment" [11, p. 1, def. 3.1]. While this is a very materialistic interpretation of "harm," as criticized e. g. in [5], [6] and [17], is should be noted that this technical risk definition is easily extensible: when broadening

[1] However, it should be pointed out that such an interpretation beyond physical net harm is covered by existing risk definitions, such as the definitions provided by [9] or [7].

[2] When compared to the formulations of Koopman and Widen, violated stakeholder values are a potential root cause for "negative societal externalities" that express non-acceptance of a technology.

[3] E.g., E/E failures as per ISO 26262 or functional insufficiencies as per ISO 21448. Based on [10], we will argue in section 4, why we prefer the term "hazard sources".

[4] Such as physical or monetary harm.

the concept of "harm" to, e.g., include violations of stakeholder values such as mobility or legality even technical risk definitions allow to capture nuances of "risk" in socio-technical settings (cf. Section 3.3).

"Safety" definitions via "risk" carry different prefixes in the literature such as "unreasonable" as in ISO 26262 [13, p. 21, def. 3.132] or "unacceptable" as in IEC 61508 [18]. Depending on such a prefix, "safety" can take an *absolute* (as in "no risk is tolerable") or a *relative* (as in "a sufficiently low risk is tolerable") notion which we will use in the following to classify risk definitions in regulatory documents.

3.1 Absolute vs. Relative Notions of Safety in Engineering, Law, and Colloquial Use

Regarding the understanding of "safety" in the context of technical systems, the "freedom from untolerable risk" is a generally accepted notion among engineers. Safety-related technical standards refer to a probability of harm, which suggests that, in engineering disciplines, it is clear that complex technical systems will never be fully infallible. The general public is indeed willing to accept certain risks, as long as they see sufficient benefit from this acceptance [20] – How much risk is "tolerable", is then subject to the definition of societal ("risk") acceptance criteria. This characterizes *relative* conceptualizations of safety.

In this respect, other disciplines and the general public may tend towards more *absolute* notions of "safety": The Cambridge Dictionary defines safety as "a state in which or a place where you are safe and not in danger or at risk" [21]. Despite the cyclic definition of safety by "safe", it is clear, that "not being at risk" suggests an absolute understanding regarding the absence of risk. Ballantine's Law Dictionary defines "safe" as "out of harm's way", pointing to concepts such as "reasonably safe" or "adequately safe" for more relaxed definitions. Hence, the basic definition proposes no chance of occurring harm and no possible relaxation of the concept of "safety" that can be related to risk while the references in the definition point to relative notions of safety. German jurisprudence, particularly police and regulatory law, [22] defines that a risk is given by a "theoretically possible, but practically unlikely" harm, also pointing to a rather absolute notion of safety.

Regarding automated driving systems, such nuances are not only theoretically relevant: Automated vehicles are operating in a practically open world. [5, 17] This means, however, that the risk incurred by automated driving systems can only be mitigated (to a tolerable degree), but never be fully eliminated. Developers thus cannot foresee every possible scenario or required system reaction at design time and must rely on well-informed assumptions during development: Designing for absolute safety would require a system design that avoids any probability of occurring harm. In an open world, this implies extremely conservative system behavior at very slow speeds [23], sacrificing the utility of the technology for its users. To avoid this, developers need to find a balance between a system design that is sufficiently safe and provides mobility at the same time.

This reasoning implies that regulation that is written with an absolute notion of safety that forbids the discussed tradeoffs can, in the worst case, hinder the introduction of the technology completely. At the very least, absolute notions of risk cause legal uncertainty for the developers of automated vehicles who design diligently for a balance between safety and utility for the users.

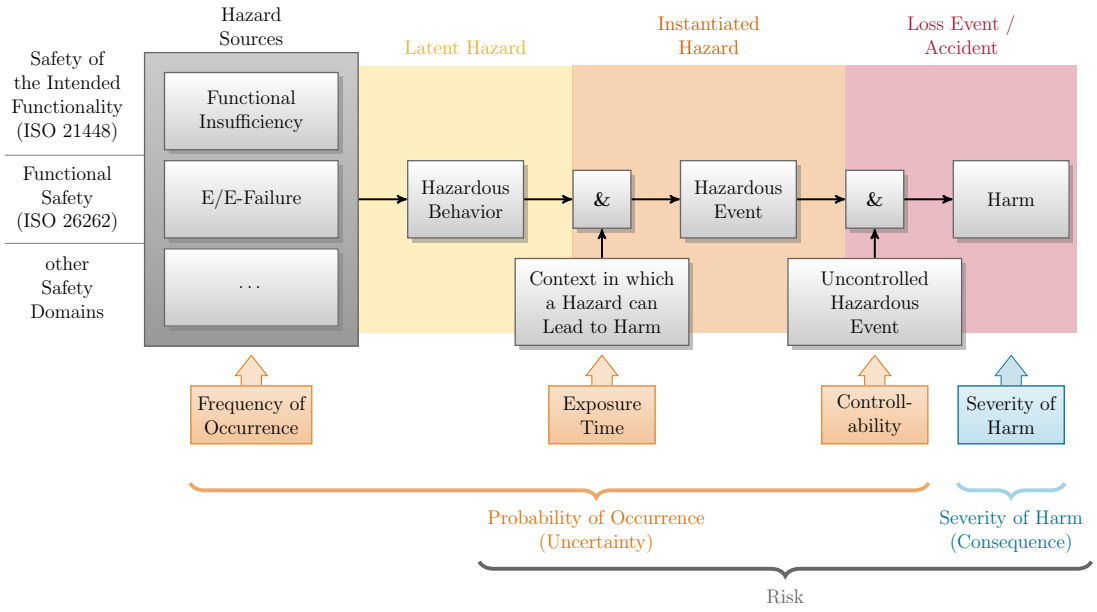

Figure 1: Graphical summary of a taxonomy of risk related to automated vehicles, extended based on ISO 21448 ([14]) and [27]. Top: Causal chain from hazard sources to actual harm; bottom: summary of the individual elements' contributions to a resulting risk. Graphic translated from [17].

3.2 Hazard Sources

Current regulatory documents ([24], [25], [26]) frequently refer to domain standards to formulate requirements for automated vehicles.

Depending on their scope, such standards address differing risk (or rather "hazard") sources whose impact the standards seek to mitigate. Figure 1 provides an overview of such different hazard sources for our reasoning, which is closely in line with the argumentation provided by Waymo in [27].

We prefer "hazard sources" over "risk sources", as a risk must always be related to a *cause* or *source of harm* (i.e., a hazard [11, p. 1, def. 3.2]). Without a concrete (scenario) context that the system is operating in, a hazard is *latent*: E.g., when considering the traffic system, a *collision with a pedestrian* is a possibility of possible (physical) harm for that pedestrian. However, only if an automated vehicle shows (potentially) hazardous behavior (e.g., not decelerating properly) *and* is located near a pedestrian (context), the hazard is instantiated and leads to a hazardous event. If the hazardous event cannot be mitigated or controlled, a loss event or accident is caused in which the pedestrian's health is harmed. Note that this hypothetical chain of events is summarized in the definition of risk: The probability of occurrence of harm is determined by a) the frequency with which hazard sources manifest, b) the time for which the system operates in a context that exposes the possibility of harm, and c) by the probability with which a hazardous event can be controlled. A risk can then be determined as a function of the probability of harm and the severity of the harm potentially inflicted on the pedestrian.

Figure 1 shows two distinct hazard sources, i.e., functional insufficiencies and E/E-failures that can lead to hazardous behavior. ISO 21448 [14] defines functional insufficiencies as insufficiencies that stem from an incomplete or faulty system specification

(specification insufficiencies). In addition, the standard considers insufficiencies that stem from insufficient technical capability to operate inside the targeted Operational Design Domain (performance insufficiencies). Functional insufficiencies are related to the "Safety of the Intended Functionality (SOTIF)" (according to ISO 21448), "Behavioral Safety" (according to Waymo [28]), or "Operational Safety" (according to UN Regulation No. 157 [25]). E/E-Failures are related to classic functional safety and are covered exhaustively by ISO 26262 [13]. Additional hazard sources can, e.g., be related to malicious security attacks (ISO/SAE 21434 [29]), AI-related failures (ISO/PAS 8800, [30]), or even to mechanical failures covered, e.g., in the US, by the Federal Motor Vehicle Safety Standards (FMVSS).

3.3 Types of Harm

Finally, as discussed in the introduction of this section, notions of safety can be classified by the related type of harm. Instead of referring to the additional concept of *loss* as in [5], we extend the concept of harm to the violation of stakeholder *need* or *values*, staying consistent with the majority of safety standards and avoiding the additional concept of loss as in [5].

For automated vehicles, *financial integrity*, *physical integrity* and *mobility* can, e.g., be considered values which need to be balanced to achieve societal acceptance, in line with the discussion of required tradeoffs in section 3. For the analysis of the following regulatory frameworks, we will evaluate if the given safety or risk notions allow tradeoffs regarding underlying stakeholder needs or values.

4 Analysis

In the following sections, we will analyze European type approval and related regulation[5] regarding the underlying notions of "safety" and "risk". We will classify these notions according to their absolute or relative character, underlying risk sources, or underlying concepts of harm.

The categories for the analysis were developed by the corresponding author and refined in the group of authors. The legal documents have been reviewed by the authors by applying a four-eyes principle. The findings were discussed in the author group and summarized by the main author.

[5]Strictly speaking, the German enabling act (AFGBV [24]) does not provide type approval regulation, but regulation for granting test & operating permits, as well as permits for operational design domains for SAE-Level-4 systems. European type-approval for small series follows the Implementing Regulation EU 2022/1426 [26], Type-approval regulation for SAE-Level-3 systems follows UN Regulation No. 157 (UN-ECE-ALKS [25]).

4.1 UN Regulation No. 157 & European Implementing Regulation (EU) 2022/1426

UN Regulation No. 157[6] [25] and the European Implementing Regulation 2022/1426[7] [26] provide type approval regulation for automated vehicles equipped with SAE-Level-3 (UN Reg. 157) and Level 4 (EU 2022/1426) systems on an international (UN Reg. 157) and European (EU 2022/1426) level. EU 2022/1426 has been amended by Commission Implementation (EU) 2026/481 [31], removing the small series limit for the approval of automated valet parking systems, without changes to sections relevant for this paper.

Generally, EU type approval considers UN ECE regulations mandatory for its member states ((EU) 2018/858, [32]), while the EU largely forgoes implementing EU-specific type approval rules, it maintains the right to alter or to amend UN ECE regulation [32].

In this respect, the terminology and conceptualizations in the EU Implementing Act closely follow those in UN Reg. No. 157. The EU Implementing Act gives a clear reference to UN Reg. No. 157 [26, Preamble, Paragraph 1]. Hence, the documents can be assessed in parallel. Differences will be pointed out as necessary.

Both acts are written in rather technical language, including the formulation of technical requirements (e.g., regarding deceleration values or speeds in certain scenarios). While providing exhaustive definitions and terminology, neither of both documents provide an explicit definition of risk or safety. The definition of "unreasonable" risk in both documents does not define risk, but only what is considered *unreasonable*. It states that "unreasonable risk" is considered the "overall level of risk for [the driver, (only in UN Reg. 157)] vehicle occupants and other road users which is increased compared to [a competently and carefully driven (UN Reg. 157)] manual[ly driven ((EU) 2022/1426)] vehicle." Without a clear definition, the concrete underlying notions of safety and risk must hence be derived from the context in which they are used.

4.1.1 Absolute vs. Relative Notions of Safety

In line with the technical detail provided in the acts, both clearly imply a *relative* notion of safety and consistently refer to the absence of *unreasonable* risk throughout, which is typical for technical safety definitions.

Both acts require sufficient proof and documentation that the to-be-approved automated driving systems are "free of unreasonable safety risks to vehicle occupants and other road users" for type approval.[8] In this respect, both acts demand that the manufacturers perform verification and validation activities for performance requirements that include "[...] the conclusion that the system is designed in such a way that it is free from unreasonable risks [...]". Additionally, *risk minimization* is a recurring theme when it comes to the definition of Minimum Risk Maneuvers (MRM).

Emphasizing the relative notions of safety and risk, UN Reg. 157 introduces the concept of "reasonable foreseeable and preventable" [25, UN Reg. 157, Article 1, Clause

[6]Full title: *UN Regulation No. 157 Uniform Provisions Concerning the Approval of Vehicles with Regard to Automated Lane Keeping Systems*

[7]Full title: *COMMISSION IMPLEMENTING REGULATION (EU) 2022/1426 of 5 August 2022 Laying down Rules for the Application of Regulation (EU) 2019/2144 of the European Parliament and of the Council as Regards Uniform Procedures and Technical Specifications for the Type-Approval of the Automated Driving System (ADS) of Fully Automated Vehicles*

[8]As it targets SAE-Level-3 systems, UN Reg. 157 also refers to the driver, where applicable.

5.1.1.] collisions, which implies that a residual risk will remain with the introduction of automated vehicles. [25, UN Reg. 157, Appendix 3, Clause 3.1.] explicitly states that only *some* scenarios that are unpreventable for a competent human driver can actually be prevented by an automated driving system. While this concept is not applied throughout the EU Implementing Act, both documents explicitly refer to *residual* risks that are related to the operation of automated driving systems ([25, UN Reg. 157, Annex I, Clause 1], [26, (EU) 2022/1426, Annex II, Clause 7.1.1.]). [25, UN Reg. 157, Appendix 3, Clause 3.1.]

In this context, (EU) 2022/1426 requires operators to define risk acceptance criteria and gives a quantitative guideline [26, (EU) 2022/1426, Annex II, Clause 7.1.1.]: "For instance based on current accident data on buses, coaches, trucks and cars in the EU, an indicative aggregated acceptance criteria of 10-7 fatalities per hour of operation could be considered for market introduction of ADSs for comparable transport services and situations."

4.1.2 Hazard Sources

Hazard sources that are explicitly differentiated in UN Reg. 157 [25] and (EU) 2022/1426 [26] are E/E-failures that are in scope of functional safety (ISO 26262 [13]) and functional insufficiencies that are in scope of behavioral (or "operational") safety (ISO 21448 [14]). Both documents consistently differentiate both sources when formulating requirements and explicitly reference ISO 26262, ISO(/PAS) 21448, and ISO/SAE 21434 (Road Vehicles – Cyber Security Engineering, [14]).

While the acts share a common definition of "operational" safety ([26, (EU) 2022/1426, Article 2, def. 30.], [25, UN Reg. 157, Annex 4, def. 2.15.]), the definitions for functional safety differ. UN Reg. 157 [25] defines functional safety as the "absence of unreasonable risk under the occurrence of hazards caused by a malfunctioning behaviour of electric/electronic systems [...]", (EU) 2022/1426 [26] drops the specification of "electric/electronic systems" from the definition. When taken at face value, this definition would mean that functional safety included all possible hazard sources, regardless of their origin, which is a deviation from the otherwise precise usage of safety-related terminology.

4.1.3 Harm Types

As the acts lack explicit definitions of safety and risk, there is no consistent and explicit notion of different harm types that could be differentiated.

UN Reg. 157 [25] provides hints regarding different considered harm types. "The absence of unreasonable risk" in terms of human driving performance of a "careful competent driver" is relatable to any chosen performance metric that allows a comparison with a competent careful human driver including, e.g., accident statistics, statistics about rule violations, or changes in traffic flow.

In (EU) 2022/1426 [26], "safety" is, implicitly, attributed to the absence of unreasonable risk to life and limb of humans. This is supported by the performance requirements that are formulated: [26, (EU) 2022/1426, Annex II, Clause 1.1.2. (d)] demands that an automated driving system can adapt the vehicle behavior in a way that it minimizes risk while prioritizing the protection of human life.

Both acts demand the adherence to traffic rules ([26, (EU) 2022/1426, Annex 2, Clause 1.3.], [25, UN Reg. 157, Clause 5.1.2.]). [26, (EU) 2022/1426, Annex II, Clause 1.1.2. (c)]

also demands that an automated driving system shall adapt its behavior to surrounding traffic conditions, such as the current traffic flow. With the relative notion of risk in both acts, the unspecific clear statement that there may be unpreventable accidents (UN Reg. 157 [25]), and a demand of prioritization of human life in (EU) 2022/1426 [26], both acts can be interpreted to allow developers to make tradeoffs as discussed in section 3.

4.1.4 Conclusion

To summarize, the UN Reg. 157 and the (EU) 2022/1426 both clearly apply the technical notion of safety as the absence of unreasonable risk. The notion is used consistently throughout both documents, providing largely unambiguous terminology. Uncertainty remains when it comes to considered harm types: Both acts do not explicitly allow tradeoffs regarding broader notions of safety in the sense of [5] or [6]. Finally, a minor weak spot can be seen in the definition of risk acceptance criteria: Both acts take the human driving performance as a baseline. While (EU) 2022/1426 specifies that these criteria are specific to the systems' Operational Design Domain [26, (EU) 2022/1426, Annex II, Clause 7.1.1.], the reference to the concrete Operational Design Domain is missing in UN Reg. 157. Without a clearly defined notion of safety, however, it remains unclear, how aspects beyond net accident statistics (which are given as an example in [26, (EU) 2022/1426, Annex II, Clause 7.1.1.]), should be addressed, as demanded by [5].

4.2 German Regulation (StVG & AFGBV)

The German L3 (Automated Driving Act) and L4 (Act on Autonomous Driving) Acts from 2017 and 2021,[9] respectively, provide enabling regulation for the operation of SAE-Level-3 and 4 vehicles on German roads. The German Implementing Regulation ([24], AFGBV) defines how this enabling regulation is to be implemented for granting test permits for SAE-Level-3 and -4 and driving permits for SAE-Level-3 and -4 automated driving systems.[10] With all three acts, Germany was the first country to regulate the approval of automated vehicles for a domestic market. All acts are subject to (repeated) evaluation until the year 2030 regarding their impact on the development of automated driving technology. An assessment of the German AFGBV and comparisons to (EU) 2022/1426 has been given in [33] in German.

As UN Reg. 157 and (EU) 2022/1426, neither the StVG nor the AFGBV provide a clear definition of "safety" or "risk" despite "safety" of road traffic is one major goal of the StVG and the German Traffic Code (Straßenverkehrsordnung, StVO). Implicit notions of both concepts can only be interpreted from the context of existing wording. An additional complication that is related to the German language is that "safety" and "security" can both be addressed as "Sicherheit", adding another potential source of ambiguity. Literal quotations in this section are our translations from the German act.

[9] Formally, these are amendments to the German Road Traffic Act (StVG): 06/21/2017, BGBl. I p. 1648, 07/12/2021 BGBl. I p. 3108.

[10] Note that these permits do not grant EU-wide type approval, but serve as a special solution for German roads only. At the same time, the AFGBV has the same scope as (EU) 2022/1426.

4.2.1 Absolute vs. Relative Notions of Safety

The main goal of the German StVO, written for human drivers, is to ensure the "safety and ease of traffic flow," which is a relevant detail when assessing the notions of safety and risk in regulation related to automated vehicles. In this sense, the StVO directly establishes a diametral goal that not only permits, but requires human drivers to make tradeoffs between conservative and permissive driving.

While the StVG enabling regulation contains sections that use wording such as "best possible safety for vehicle occupants" (§1d (4) StVG) and acknowledges that there are unavoidable hazards to human life (§1e (2) No. 2c)), the German AFGBV contains ambiguous statements in this respect: Several paragraphs contain a demand for a hazard free operation of automated vehicles: §4 (1) No. 4 AFGBV, e.g., states that "the operation of vehicles with autonomous driving functions must neither negatively impact road traffic safety or traffic flow, nor endanger the life and limb of persons." Additionally, §6 (1) AFGBV states that the permits for testing and operation have to be revoked, if it becomes apparent that a "negative impact on road traffic safety or traffic flow, or hazards to the life and limb of persons cannot be ruled out". The same wording is used for the approval of operational design domains regulated in §10 (1) No. 1. A particularly misleading statement is made regarding the requirements for technical supervision instances which are regulated in §14 (3) AFGBV which states that an automated vehicle has to be "immediately removed from the public traffic space if a risk minimal state leads to hazards to road traffic safety or traffic flow". Considering the argumentation in section 3, that residual risks related to the operation of automated driving systems are inevitable, these are strong statements which, if taken at face value, technically prohibit the operation of automated vehicles. It suggests an *absolute* notion of safety that requires the complete absence of risk. The last statement above is particularly contradictory in itself, considering that a risk *minimal* state always implies a residual risk and does not eliminate hazards.

In addition to these absolute safety notions, there are passages which suggest a relative notion of safety: The approval for Operational Design Domains is coupled to the proof that the operation of an automated vehicle "neither negatively impacts road traffic safety or traffic flow, nor significantly endangers the life and limb of persons beyond the general risk of an impact that is typical of local road traffic" (§9 (2) No. 3 AFGBV). The addition of a relative risk measure "beyond the general risk of an impact" provides a relaxation (cf. also [33], who criticizes the aforementioned absolute safety notion) that also yields an implicit acceptance criterion (*statistically as good as* human drivers) similar to the requirements stated in UN Reg. 157 and (EU) 2022/1426.

Additional hints for a relative notion of safety can be found in Annex 1, Part 1, No. 1.1 and Annex 1, Part 2, No. 10. Part 1, No 1.1 specifies collision-avoidance requirements and acknowledges that not all collisions can be avoided.[11] Part 2, No. 10 specifies requirements for test cases. It demands that test cases are suitable to provide evidence that the "safety of a vehicle with an autonomous driving function is increased compared to the safety of human-driven vehicles". This does not only acknowledge residual risks, but also yields an acceptance criterion (*better* than human drivers) that is different from the implied acceptance criterion given in §9 (2) No. 3 AFGBV.

[11] The same is true for Part 2, No. 10, Clause 10.2.5.

4.2.2 Hazard Sources

Regarding hazard sources, Annex 1 and 3 AFGBV explicitly refer to ISO 26262 and ISO 21448 (or rather its predecessor ISO/PAS 21448:2019). However, regarding the discussion of actual hazard sources, the context in which both standards are mentioned is partially unclear: Annex 1, Clause 1.3 discusses requirements for path and speed planning. Clause 1.3 d) demands that in intersections, a Time to Collision (TTC) greater than 3 seconds must be guaranteed. If manufacturers deviate from this, it is demanded that "state-of-the-art, systematic safety evaluations" are performed. Fulfillment of the state of the art is assumed if "the guidelines of ISO 26262:2018-12 Road Vehicles – Functional Safety are fulfilled". Technically, ISO 26262 is not suitable to define the state of the art in this context, as the requirements discussed fall in the scope of operational (or behavioral) safety (ISO 21448). A hazard source "violated minimal time to collision" is clearly a functional insufficiency, not an E/E-failure.

Similar unclarity presents itself in Annex 3, Clause 1 AFGBV: Clause 1 specifies the contents of the "functional specification". The "specification of the functionality" is an artifact which is demanded in ISO 21448:2022 (Clause 5.3) [14]. However, Annex 3, Clause 1 AFGBV states that the "functional specification" is considered to comply to the state of the art, if the "functional specification" adheres to ISO 26262-3:2018 (Concept Phase). Again, this assumes SOTIF-related contents as part of ISO 26262, which introduces the "Item Definition" as an artifact, which is significantly different from the "specification of the functionality" which is demanded by ISO 21448. Finally, Annex 3, Clause 3 AFGBV demands a "documentation of the safety concept" which "allows a functional safety assessment". A safety concept that is related to operational / behavioral safety is not demanded. Technically, the unclarity with respect to the addressed standards and by that harm types lead to the fact that the requirements provided by the AFGBV do not comply with the state of the art in the field, providing questionable regulation.

4.2.3 Harm Types

Just like UN Reg. 157 and (EU) 2022/1426, the German StVG and AFGBV do not explicitly differentiate concrete harm types for their notions of safety. However, the AFGBV mentions three main concerns for the operation of automated vehicles which are *traffic flow* (e.g., §4 (1) No. 4 AFGBV), compliance to *traffic law* (e.g., §1e (2) No. 2 StVG), and the *life and limb of humans* (e.g., §4 (1) No. 4 AFGBV).

Again, there is some ambiguity in the chosen wording: The conflict between traffic flow and safety has already been argued in section 3. The wording as written in §4 (1) No. 4 and §6 (1) AFGBV demands to ensure (absolute) safety *and* traffic flow at the same time, which is impossible (cf. section 3) from an engineering perspective. §1e (2) No. 2 StVG defines that "vehicles with an autonomous driving function must [...] be capable to comply to [...] traffic rules in a self-contained manner". Taken at face value, this wording implies that an automated driving system could lose its testing or operating permit as soon as it violates a traffic rule. A way out could be provided by §1 of the German Traffic Act (StVO) which demands careful and considerate behavior of all traffic participants and by that allows judgement calls for human drivers. However, whether §1 is applicable in certain situations is often settled in court cases. For developers, the application of §1 StVO during system design hence remains a legal risk.

While there are rather absolute statements as mentioned above, sections of the AFGBV and StVG can be interpreted to allow tradeoffs: §1e (2) No. 2 b) demands that a system, "in case of an inevitable, alternative harm to legal objectives, considers the significance of the legal objectives, where the protection of human life has highest priority". This exact wording *could* provide some slack for the absolute demands in other parts of the acts, enabling tradeoffs between (tolerable) risk and mobility as discussed in section 3. However, it remains unclear if this interpretation is legally possible.

4.2.4 Conclusion

In contrast to UN Reg. 157 and (EU) 2022/1426, the German StVG and AFGBV introduce openly inconsistent notions of safety and risk which are partially directly contradictory: The wording partially implies absolute and relative notions of safety and risk at the same time. The implied validation targets ("better" or "as good as" human drivers) are equally contradictory. The partially implied absolute notions of safety, when taken at face value, prohibit engineers from making the tradeoffs required to develop a system that is safe and provides customer benefit at the same time. While the more relative notions in other parts of the regulation could permit a more relative interpretation of other absolute formulations, this is not stated explicitly.

This uncertainty creates additional clarification need and effort for manufacturers and engineers who design and develop SAE-Level-3 and -4 automated driving systems. The use of undefined legal terms not only makes it more difficult for engineers to comply with the law, but also complicates the interpretation of the law and leads to legal uncertainty.

4.3 UK Automated Vehicles Act 2024 (2024 c. 10)

The UK has issued a national enabling act for regulating the approval of automated vehicles on the roads in the UK. To the best of our knowledge, concrete implementing regulation has not been issued yet. Regarding terminology, the act begins with a dedicated terminology section to clarify the terms used in the act [34, Part 1, Chapter 1, Section 1]. In that regard, the act defines a vehicle to drive "'autonomously' if — (a) it is being controlled not by an individual but by equipment of the vehicle, and (b) neither the vehicle nor its surroundings are being monitored by an individual with a view to immediate intervention in the driving of the vehicle." The act hence covers SAE-Level-3 to SAE-Level-5 automated driving systems.

4.3.1 Absolute vs. Relative Notions of Safety

While not providing an explicit definition of safety and risk, the UK Automated Vehicles Act ("UK AV Act") [34] explicitly refers to a relative notion of safety. Part 1, Chapter 1, Section 1, Clause (7) (a) defines that an automated vehicle travels "'safely' if it travels to an acceptably safe standard". This acknowledges that absolute safety is not achievable and that acceptance criteria to prove the acceptability of residual risk are required, even though a concrete safety definition is not given. The act explicitly tasks the UK Secretary of State[12] to install safety principles to determine the "acceptably safe standard" in Part 1, Chapter 1, Section 1, Clause (7) (a). In this respect, the act also provides one general

[12]Which means, that concrete implementation regulation needs to be enacted.

validation target as it demands that the safety principles must ensure that "authorized automated vehicles will achieve a level of safety equivalent to, or higher than, that of careful and competent human drivers". Hence, the top-level validation risk acceptance criterion assumed for UK regulation is "*at least as good* as human drivers".

4.3.2 Hazard Sources

The UK AV Act contains no statements that could be directly related to different hazard sources. Note that, in contrast to the rest of the analyzed documents, the UK AV Act is enabling rather than implementing regulation. It is hence comparable to the German StVG, which does not refer to concrete hazard sources as well.

4.3.3 Types of Harm

Even though providing a clear relative safety notion, the missing definition of risk also implies a lack of explicitly differentiable types of harm. Implicitly, three different types of harm can be derived from the wording in the act. This includes the harm to life and limb of humans[13], the violation of traffic rules[14], and the cause of inconvenience to the public [34, Part 1, Chapter 1, Section 58, Clause (2) (d)].

The act connects all the aforementioned types of harm to "risk" or "acceptable safety". While the act generally defines criminal offenses for providing "false or misleading information about safety", it also acknowledges possible defenses if it can be proven that "reasonable precautions" were taken and that "due diligence" was exercised to "avoid the commission of the offence". This statement could enable tradeoffs within the scope of "reasonable risk" to the life and limb of humans, the violation of traffic rules, or to the cause of inconvenience to the public, as we argued in section 3.

4.3.4 Conclusion

From the set of reviewed documents, the current UK AV Act is the one with the most obvious relative notions of safety and risk and the one that seems to provide a legal framework for permitting tradeoffs. In our review, we did not spot major inconsistency beyond missing definitions of safety and risk[15].

The general relative notion of safety and the related alleged ability for designers to argue well-founded development tradeoffs within the legal framework could prove beneficial for the actual implementation of automated driving systems. While the act thus appears as a solid foundation for the market introduction of automated vehicles, without accompanying implementing regulation, it is too early to draw definite conclusions.

[13]Part 1, Chapter 3, Section 25 defines "aggravated offence where death or serious injury occurs" [34].

[14]Part 1, Chapter 1, Clause (7) (b) defines that an automated vehicle travels "'legally' if it travels with an acceptably low risk of committing a traffic infraction"

[15]Note that with the Office for Product Safety and Standards (OPSS), there is a British government agency that maintains an exhaustive and widely focused "Risk Lexicon" that provides suitable risk definitions. For us, it remains unclear, to what extent this terminology is assumed general knowledge in British legislation.

Table 1: Table summarizing the findings in the analyzed documents.

Regulation	Safety Notion	Hazard Sources	Harm Types	Tradeoffs
UN Reg. 157 & EU 2022/1426	relative	• E/E failures[a] • functional insufficiencies • cybersecurity threats	• life & limb • traffic rule compliance • traffic flow	not forbidden
German StVG	relative	—	• life & limb • traffic rule compliance • traffic flow	ambiguous
German AFGBV	relative & absolute	• E/E failures • functional insufficiencies	• life & limb • traffic rule compliance • traffic flow	ambiguous
UK AV Act 2024	relative	—	• life & limb • traffic rule compliance • public inconvenience	permitted

[a] With ambiguities in (EU) 2022/1426

5 Requirements & Recommendations

This paper reviewed UN, EU, German, and UK regulation for the type approval of SAE-Level-$\geq$3 automated driving systems regarding the underlying notions of safety and risk. Our analysis has shown that the considered acts vary significantly in the consistency of the used concepts. The UK AV Act appeared as the regulation with the most consistent, yet generic, conceptualization of safety and risk. The UN and EU regulations adopt a very technical notion of safety and risk, based on the absence of unreasonable risk, while mostly focusing on harm to the life and limb of humans. Compared to the other acts, the German AFGBV provides ambiguous notions of risk and safety. These notions range from the assumption of a *complete* absence of risk, to more technical notions considering the absence of *unreasonable* risk.

In summary, UN and EU regulations as well as the UK AV Act provide frameworks that permit traditional technical interpretations of risk management. Moreover, the UK AV Act could provide a legal framework that allows adopting more holistic perspectives on safety, as discussed by [6] or [5].

Based on the analysis (findings summarized in Table 1), the German StVG and AFGBV create potential of significant legal uncertainty for designers and developers of automated vehicles due to ambiguous or even contradictory wording and requirements. This creates additional effort for engineers to validate assumptions regarding the requirements with the regulators. However, as the regulatory statutes stand as they are, even such clarification can entail the need for court cases, before legal uncertainty is settled. This can unnecessarily impede the market introduction of automated vehicles, where clear regulation could have helped from the beginning.

5.1 Requirements

Given the analysis of the different regulations as well as ongoing challenges regarding the deployment of automated vehicles, regulators should respect the following requirements:

1. AV regulation should apply consistent and explicit relative definition of safety.
2. AV regulation should explicitly acknowledge uncertainty that is related to the inherent risk of public traffic for the safety of technical systems.[16]
3. AV regulation should define clear legal constraints and documentation needs for required development tradeoffs rooted in stakeholder needs and values that are required to resolve that uncertainty.
4. AV regulation should actively resolve potential for regulatory uncertainty that arise from performing such tradeoffs.[17]

5.2 Recommendations

Given these requirements we recommend regulators

a) to clarify current contradictions and ambiguity, particularly regarding open signifiers such as "safety" and "risk",
b) to include broader notions of risk connected to different harm types that force developers to consider negative impacts such as "inconvenience for the public" (see [5] for more examples),
c) to define documentation duties for development decisions that address required tradeoffs between stakeholder needs,
d) to facilitate stakeholder dialogues that support the formulation of clear societal risk acceptance criteria.

Finally, the EU, UN, partially, and the German regulation exhaustively refer to technical standards, which is not an issue per se. The German regulation contains statements that refer at least partially to the wrong standard (cf. Section 4.2.2). At the same time, the current standardization landscape is highly dynamic with new versions of ISO 26262 and ISO 21448 underway. Binding the regulation to specific versions such as ISO/PAS 21448 which has been superseded by the final version of standard with significant changes highlights the issue, while the UK regulation avoids these specific references.

Regulators should hence

e) monitor the maturity of standards relevant to the field,
f) avoid too specific references to evolving standards,
g) avoid false references to standards.

The UK AV Act (implementing regulation is currently expected for 2027) could serve as an interesting example in respect to the recommendations: despite only providing enabling regulation and despite the lack of explicit definitions for risk and safety, it has largely avoided pitfalls with respect to terminology and standardization.

[16] Such as it is acknowledged for human drivers.

[17] This includes tensions between product liability and documentation needs: diligently acknowledging development uncertainty and defining diligent mitigation strategies should not create potential for product liability cases. Furthermore, existing perceived regulatory conflicts (e.g. between the European GDPR and AI Act regarding data protection and data collection requirements, cf. [17]) should be actively mitigated.

Currently available preliminary versions of the Proposal for a new UNECE United Nations Global Technical Regulation on Automated Driving Systems (ADS) [35] by the UN Working Party on Automated/Autonomous and Connected Vehicles (GRVA) suggest that appropriate care is taken at the international level, pending further analysis of the pending final version of the regulation.

5.3 Outlook

The scope of this paper was consciously limited to recommendations that can be directly related to the way that the analyzed European regulations refer to notions of safety and risk. As discussed in [36], one major contributing factor to long-term societal acceptance of automated vehicles lies in an "open information without reservations, and transparent policy" (following [37]). However, legal and business concerns might hinder the open communication of residual risk by companies. Hence, another requirement for regulation might be formulated to demand *ex ante* risk communication by manufacturers—in turn, reducing legal uncertainty. In future work, we want to examine how existing regulation accounts for this demand. Further research hence includes conceptualizing means for effective risk communication regarding automated vehicle technology and reflecting on the potential role of the regulator in this process.

References

[1] W. Hoffman-Riem, "Regulierungswissen in der Regulierung," in *Wissensregulierung und Regulierungswissen*, A. Bora, A. Henkel, and C. Reinhardt, Eds., Weilerswist: Velbrück Wiss, 2014, pp. 135–156.

[2] H. Steege, "Technikregulierung - Der Umgang mit Risiken im Recht," *NZV - Neue Zeitschrift Verkehrsrecht, Verlag H.C. Beck*, vol. 35, no. 6, pp. 257–262, 2022.

[3] T. Fleischer, "Safety and acceptance – a view of two mysteries," Oberseminar Electronic Vehicle Systems, TU Braunschweig, Online, Invited Talk, 2023.

[4] I. Buchanan, *A Dictionary of Critical Theory*, 1st ed. Oxford University Press, Jan. 1, 2010. DOI: 10.1093/acref/9780199532919.001.0001.

[5] P. Koopman and W. Widen, "Redefining safety for autonomous vehicles," in *Computer Safety, Reliability, and Security*, A. Ceccarelli, M. Trapp, A. Bondavalli, and F. Bitsch, Eds., vol. 14988, Series Title: Lecture Notes in Computer Science, Cham: Springer Nature Switzerland, 2024, pp. 300–314. DOI: 10.1007/978-3-031-68606-1_19.

[6] N. F. Salem, S. Le Page, J. Millar, P. Junietz, M. Nolte, R. Graubohm, and M. Maurer, "Safety and risk—why their definitions matter," in *Handbook Assisted and Automated Driving*, H. Winner, K. C. J. Dietmayer, L. Eckstein, M. Jipp, M. Maurer, and C. Stiller, Eds., 4th ed., Wiesbaden: Springer Fachmedien Wiesbaden, 2026, pp. 779–800. DOI: 10.1007/978-3-658-45276-6_40.

[7] O. Renn, "Three decades of risk research: Accomplishments and new challenges," *Journal of Risk Research*, vol. 1, no. 1, pp. 49–71, 1998. DOI: 10.1080/136698798377321.

[8] T. L. Sellnow, R. R. Ulmer, M. W. Seeger, and R. S. Littlefield, *Effective Risk Communication: A Message-Centered Approach*. New York, NY: Springer, 2009. DOI: 10.1007/978-0-387-79727-4.

[9] B. Fischhoff, S. R. Watson, and C. Hope, "Defining risk," *Policy Sciences*, vol. 17, no. 2, pp. 123–139, 1984. DOI: 10.1007/BF00146924.

[10] F. M. Christensen, O. Andersen, N. J. Duijm, and P. Harremoës, "Risk terminology—a platform for common understanding and better communication," *Journal of Hazardous Materials*, vol. 103, no. 3, pp. 181–203, 2003. DOI: 10.1016/S0304-3894(03)00039-6.

[11] "Safety aspects — guidelines for their inclusion in standards," International Organization for Standardization (ISO), International Electrotechnical Commission (IEC), ISO/IEC Guide 51, 2014.

[12] WMG – The University of Warwick, "Cross domain safety assurance for automated transport systems," Warwick, UK, 2023, pp. 1–19.

[13] "Road vehicles — functional safety," International Organization for Standardization (ISO), ISO Standard 26262:2018.

[14] "Road vehicles — safety of the intended functionality," International Organization for Standardization (ISO), ISO Standard 21448:2022.

[15] "Standard for safety - evaluation of autonomous products," American National Standards Institute (ANSI(/ UL Solutions (UL), Northbrook, USA, ANSI/UL Standard 4600, 2020.

[16] U. Di Fabio, M. Broy, R. Jungo Brüngger, U. Eichhorn, A. Grunwald, D. Heckmann, E. Hilgendorf, H. Kagermann, A. Losinger, C. Lütge, A. Markl, K. Müller, and K. Nehm, "Ethik-Kommission Automatisiertes und Vernetztes Fahren," Bundesministerium für Verkehr und digitale Infrastruktur (BMVI), 2017.

[17] M. Nolte, "Werte- und fähigkeitsbasierte Bewegungsplanung für autonome Straßenfahrzeuge – Ein systemischer Ansatz," (in German), Ph.D. dissertation, TU Braunschweig, Braunschweig, 2025.

[18] "Functional safety of electrical/electronic/programmable electronic safety-related systems," International Electrotechnical Commission (IEC), IEC Standard 61508:2010.

[19] "Risk management — guidelines," International Organization for Standardization (ISO), ISO Standard 3100:2018.

[20] A. Grunwald, "Societal risk constellations for autonomous driving. analysis, historical context and assessment," in *Autonomous Driving*, M. Maurer, J. C. Gerdes, B. Lenz, and H. Winner, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2016, pp. 641–663. DOI: 10.1007/978-3-662-48847-8_30.

[21] Cambridge Dictionary. "Safety." (2025), [Online]. Available: https://dictionary.cambridge.org/dictionary/english/safety (visited on 01/27/2025).

[22] F.-J. Peine, "Gefahr, Risiko, Restrisiko – Begiffsbestimmungen und Probleme aus juristischer Sicht," in *Jahrbuch 1999 der Braunschweigischen Wissenschaftlichen Gesellschaft*, (in German), 1999, pp. 151–164.

[23] R. Graubohm, N. F. Salem, M. Nolte, and M. Maurer, "On assumptions with respect to occlusions in urban environments for automated vehicle speed decisions," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, Bilbao, Spain: IEEE, 2023, pp. 738–745. DOI: 10.1109/ITSC57777.2023.10422457.

[24] *Verordnung zur Genehmigung und zum Betrieb von Kraftfahrzeugen mit autonomer Fahrfunktion in festgelegten Betriebsbereichen (AFGBV)*, (in German), Jan. 7, 2022.

[25] *UN regulation no. 157 uniform provisions concerning the approval of vehicles with regard to automated lane keeping systems*, 2021.

[26] *COMMISSION IMPLEMENTING REGULATION (EU) 2022/1426 of 5 august 2022 laying down rules for the application of regulation (EU) 2019/2144 of the european parliament and of the council as regards uniform procedures and technical specifications for the type-approval of the automated driving system (ADS) of fully automated vehicles*, Aug. 26, 2022.

[27] F. M. Favarò, L. Fraade-Blanar, S. Schnelle, T. Victor, M. Peña, J. Engström, J. Scanlon, K. Kusano, and D. Smith, "Building a credible case for safety: Approach proposal for automated driving systems," *Journal of Safety Research*, vol. 96, pp. 181–196, Feb. 2026. DOI: 10.1016/j.jsr.2025.10.019.

[28] Waymo LLC. "Waymo safety report." (2021), [Online]. Available: https://storage.googleapis.com/waymo-uploads/files/documents/safety/2021-03-waymo-safety-report.pdf (visited on 05/25/2026).

[29] "Road vehicles — cybersecurity engineering," International Organization for Standardization (ISO), Society of Automotive Engineers (SAE), ISO/SAE Standard 21434:2018.

[30] "Road vehicles — safety and artificial intelligence," International Organization for Standardization (ISO), ISO Publicy Available Specification 26262:2018.

[31] *COMMISSION IMPLEMENTING REGULATION (EU) 2026/481 of 3 march 2026 amending implementing regulation (EU) 2022/1426 laying down rules for the application of regulation (EU) 2019/2144 of the european parliament and of the council as regards uniform procedures and technical specifications for the type-approval of the automated driving system of fully automated vehicles*, Mar. 3, 2026.

[32] *REGULATION (EU) 2018/858 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 30 may 2018 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles, amending regulations (EC) no 715/2007 and (EC) no 595/2009 and repealing directive 2007/46/EC*, May 30, 2018.

[33] U. Steininger, "Rechtliche Fragestellungen des automatisierten Fahrens," TÜV SÜD Rail GmbH, innocam.NRW, 2022, (in German).

[34] *Automated vehicles act 2024 c. 10, UK AV act*, May 20, 2024.

[35] Working Party on Automated/Autonomous and Connected Vehicles (GRVA), "Proposal for a new united nations global technical regulation on automated driving systems (ADS)."

[36] M. Maurer, "Runder tisch autonomes fahren des BMV – anforderungen an die regulierung," Oberseminar Electronic Vehicle Systems, TU Braunschweig, Online, Invited Talk, 2026.

[37] K. Homann, "Wirtschaft und gesellschaftliche akzeptanz: Fahrerassistenzsysteme auf dem prüfstand," in *Fahrerassistenzsysteme mit maschineller Wahrnehmung*, M. Maurer and C. Stiller, Eds., Springer-Verlag, 2005, pp. 239–244. DOI: 10.1007/b138667.