# TARA 2.0 for Connected and Automated Vehicles

Benyahya, Meriem; Collen, Anastasija; Lenard, Teri; Nijdam, Niels Alexander

**How to cite**

BENYAHYA, Meriem et al. TARA 2.0 for Connected and Automated Vehicles. In: IEEE transactions on intelligent transportation systems, 2025, p. 15. doi: 10.1109/TITS.2025.3574638

This publication URL: https://archive-ouverte.unige.ch/unige:185960
Publication DOI: 10.1109/TITS.2025.3574638

# TARA 2.0 for Connected and Automated Vehicles

Meriem Benyahya, *Member, IEEE*, Anastasija Collen, *Member, IEEE*, Teri Lenard, and Niels Alexander Nijdam

***Abstract*— Connected Automated Vehicles (CAVs) represent a transformative shift in transportation, offering enhanced safety, and efficiency. However, achieving full automation at levels four and five of the Society of Automotive Engineering (SAE) scale poses significant cybersecurity and privacy risks. To address these risks, United Nations Economic Commission for Europe (UNECE) regulations and ISO/SAE 21434 mandate Threat Analysis and Risk Assessment (TARA) as a core methodology for cyber risk management. Existing TARA frameworks, designed for conventional vehicles, fall short for higher automation levels, neglecting complexities such as the absence of human control and data-driven decision making concerns. This work, conducted within ULTIMO, a project tackling the CAVs deployment challenges, introduces TARA 2.0, an enhanced framework addressing cybersecurity, privacy, and expert subjectivity in risk assessment. A step-by-step experimental implementation demonstrates its feasibility, compliance with standards, and potential to secure the deployment of fully automated vehicles.**

***Index Terms*— Autonomous vehicles, road vehicle cybersecurity risk assessment, automotive threat analysis and risk assessment, intelligent transport systems, ISO/SAE 21434.**

## I. Introduction

FOR the last decade, the automotive sector has witnessed a major switch from mechanical to cyber-physical systems (CPS) where Information Technology (IT) components have become dominant. Recent technological enablers such as Artificial Intelligence (AI) [1], advanced environmental sensors (like Light Detection and Ranging (LiDAR) [2]), Electronic Control Units (ECUs), as well as Vehicle-to-everything (V2X) communications [3] support in steering the automotive domain further to achieve highly automated driving where human interventions are reduced to a bare minimum. The SAE proposed six levels of automation, where vehicles of levels four (L4) and five (L5), can drive autonomously within limited or unlimited Operational Design Domain (ODD) [4]. The CAV is a subset of the broader Internet of Vehicles (IoV), which itself stems from the Internet of Things (IoT) [5]. This evolution has transformed conventional vehicles into smart agents that remain continuously connected. However, by incorporating safety-critical systems, software, hardware and constant data exchanges, CAVs bring out a litany of attack vectors [6].

A compromised component from the CAV's environment implies inappropriate driving operations, sensitive data infringements and even fatal accidents jeopardising the passengers safety and privacy. CAV's potential attacks vary from manipulating perception systems leading to blinding the vehicle vision, falsifying navigation data causing a go off course, infiltrating communication channels leading to malicious messages injections to code alteration impacting the vehicle motion control capabilities [7]. On that note, the literature recorded several attacks such as spoofing [8], jamming, data tampering, AI models poisoning [9], malware injection, Man in the Middle (MitM) or sybil attacks [10], and V2X messages manipulation [11]. Moreover, any malicious access to the CAV, where numerous of directly or indirectly identifiable data are exchanged, is a subject to a potential data breach. Consequently, the sniffed or maliciously processed data that can likely embed Personally Identifiable Information (PII) such as name and phone number, taken routes as well as departure and arrival addresses can be the source of privacy attacks like identity theft, location tracking and profiling [12].

Such threats guided the core regulatory and standardisation bodies to harmonise the cybersecurity governance. The UNECE R155 [13] and R156 [14] mandate TARA, referring to ISO/SAE 21434 [15], as an automotive cybersecurity governance tool for detecting, evaluating, mitigating and monitoring potential threats throughout the vehicle life-cycle from design to end of life stages. As of now, security actors from industry, academia and standardisation bodies proposed numerous TARA methodologies such as EVITA [16], HEAVENS [17], TVRA [18], SARA [19] and PIER [20]. However, these methods lack the customisation of the assessment process required for L4 and L5, as they are primarily designed for conventional automotive systems. For instance, EVITA and HEAVENS are well-suited for traditional in-vehicular architecture but they struggle with the evolving threats induced by highly autonomous systems.

Comparative investigations [21] highlighted these gaps, demonstrating that the existing TARA methodologies insufficiently tackle the specific properties of L4 and L5 CAVs and the related challenges. First, existing methodologies do not consider data privacy threats at the forefront of secure CAV's implementation [12]. Second, while the TARA process depends heavily on experts evaluation, existing methodologies do not advertise any confidence factor supporting in determining the objectivity of the assessment outcomes [22]. Third, existing methodologies lack detailed demonstrations to

Received 31 May 2024; revised 23 October 2024, 17 February 2025, and 3 April 2025; accepted 22 May 2025. The Associate Editor for this article was S. Garg. *(Corresponding author: Meriem Benyahya.)*

The authors are with Geneva School of Economics and Management, Centre Universitaire d'Informatique, University of Geneva, 1227 Carouge, Switzerland (e-mail: meriem.benyahya@unige.ch; anastasija.collen@unige.ch; teri.lenard@unige.ch; niels.nijdam@unige.ch).

This article has supplementary downloadable material available at https://doi.org/10.1109/TITS.2025.3574638, provided by the authors.

Digital Object Identifier 10.1109/TITS.2025.3574638

TABLE I
DIFFERENTIATION OF SAE AUTOMATION LEVELS

| Properties | L0 | L1 | L2 | L3 | L4 | L5 |
|---|---|---|---|---|---|---|
| **Driving automation** | No | Driver assistance | Partial | Conditional | High | Full |
| **ODD** | N/A | Domain specific | Domain specific | Domain specific | Domain specific | Unlimited |
| **DDT fallback** | Driver | Driver | Driver | Fallback ready-operator | ADS or fallback ready-operator | ADS |
| **Connectivity** | Not required | Not required | Not required | Recommended | Recommended | Extended V2X |

facilitate the entire TARA process replication. Therefore, it is necessary to resolve these shortcomings in the current TARA framework to ensure that L4 and L5 CAVs can withstand continuously evolving cybersecurity and data privacy threats. Consequently, the following hypothesis has guided this research:

*Hypothesis:* TARA 2.0, enhanced with the inclusion of privacy goals, SAE level and the experts objectivity, will prompt the detection of new risk factors specific to different SAE levels influencing automated transportation.

This work aims to reply these Research Questions (RQs):

*RQ 1:* Is it feasible to extend TARA methodology for L4 and L5 CAVs while improving the focus on privacy threats?

*RQ 2:* How the assessed risks from TARA can depict scrupulously the CAV's SAE automation level?

*RQ 3:* To what extend the TARA process can be automated to reduce its reliance on experts opinion?

The present paper contributions are the following: (i) Identifying and formulating the improvement avenues in the existing TARA framework, mandated by ISO/SAE 21434 (hereby denoted as TARA 1.0), to address the specific challenges of L4 and L5 CAVs. (ii) Proposing TARA 2.0 as an improved framework which enhances focus on privacy risks, differentiates between SAE automation levels, and reduces reliance on subjective expert judgement. (iii) Demonstrating TARA 2.0 through a Proof of Concept (PoC) using a step-by-step approach to showcase the applicability of the framework.

The remainder of this paper is organised as follows: Section II presents key concepts, Section III captures the related work. Section IV proposes TARA 2.0 by outlining the methodology for the construction of each step of the assessment. Section V demonstrates the use of TARA 2.0 through an assessment of the Automated Driving System (ADS) processing unit of an L4 CAV. Section VI presents the limitations and the foreseen future work. The paper concludes in Section VII.

## II. KEY CONCEPTS AND DEFINITIONS

### A. CAVs Ecosystem

Cybersecurity and data privacy challenges are intrinsically connected to the SAE automation levels, as higher levels of automation reduce human interactions while introducing more complex interactions between automation units, and external systems. Each level from Table I is differentiated by its reference to the automation of the Dynamic Driving Task (DDT), encompassing both human and the ADS engagement, along with the ODD, which describes the driving conditions and limitations [23]. In case of a cyber attack, and depending on the SAE level, the driver, operator or the ADS has to take over or relinquish the DDT [4]. In the instance of an attack targeting L4 perception sensors, a prepared fallback operator, whether remote or in-vehicle operator, can guide the vehicle into a stable and safe state, referred to as the Minimal Risk Condition (MRC) in standardised terminology [23]. Projecting an equivalent scenario over an L5 CAV, the ADS per se must achieve the MRC independently of any type of human intervention. As showcased in Table I, several features support distinguishing the properties of each SAE level including the ODD limitation, how the MRC can be conducted, as well as the amount of V2X connections. Therefore, cybersecurity and data privacy considerations, including threats modelling and risk governance have to address such differences respectively.

With the prevalence of such challenges, the core standardisation bodies, as per International Organization for Standardization (ISO) and SAE, elicited norms reflecting risk management applications into the automotive domain. The most dominant standard to comply with is ISO/SAE 21434 [15] since it has been proposed as a key reference in the mandated regulations R155 and R156 by the UNECE Article 29 Working Party (WP29). Both regulations are requiring the Cyber Security Management System (CSMS) [13] and the Software Update Management System (SUMS) [14] certificates respectively. The CSMS comes with the obligations of integrating cybersecurity governance to the automobile manufacturers (OEM)'s organisation and over the entire value chain [22]. Similarly, the SUMS aims to demonstrate that any vehicular software update is not extorting further cyber risks or impacting the overall cybersecurity governance [14]. Both certificates have been set as pre-requisites of the vehicle type approval with a three years renewal cycle by the presentation of an assessment evidence [13]. In this context, ISO/SAE 21434 aligns with UNECE regulations in two perspectives. First, it offers guidelines for organisational audits. Second, it sets the path for obtaining type approval evidence by executing TARA 1.0 [24].

### B. TARA 1.0

TARA 1.0 is a risk-based automotive testing approach which aims to identify threats, evaluate their impact and feasibility, and combine them to derive and prioritise the risks [25]. TARA is used as an acronym for multiple terms including a systematic testing approach, a method, a methodology or even a framework [24]. In this paper, methodology and framework terms are used interchangeably while referring to TARA. TARA is distinct from Hazard Analysis and Risk Assessment (HARA) [26]. While both incorporate standardised Risk Assessment (RA) principles, the former, from ISO/SAE 21434, addresses intended harm conducted by malicious attackers, whereas the latter, from ISO 26262, assesses accidental and hazardous harm [27]. Given the potential overlap of safety and security concepts in the CAV's domain, the present research specifically focuses

on TARA 1.0, centring on cybersecurity and data privacy concerns, yet with safety implications. Additionally, it is important to highlight that TARA 1.0 is the most relevant methodology for CAVs, as informed by a systematic review conducted in a previous study [21].

Figure 1 provides a visual representation outlining the systematic procedures evoked by ISO/SAE 21434 and which are discussed in the following subsections:

*1) Item Definition:* Consists of determining the item boundary, the extent of analysis for the asset under evaluation, and its related functionalities [15]. For item connections and data exchanges with the other components, a generic architecture as well as a Data Flow Diagram (DFD) should be sketched.

*2) Asset Identification:* Lists valuable and attractive targets (components or services) for attackers [15], within the predefined item boundary. Damage Scenarios (DSs) are then determined for each asset showcasing the consequences of a compromise. Then, assets and DSs are associated using the Confidentiality, Integrity and Availability (CIA) model [28].

*3) Threat Scenario Identification:* Investigates each threat type in relation to the DFD's elements [29]. It is accomplished by naming the action required to accomplish each DS determined within the previous step. ISO/SAE 21434 endorses STRIDE but allows other tools [15].

*4) Attack Path Analysis:* Attack Paths (APs) describe how attackers exploit identified threats using three approaches: (i) top-down (from historical vulnerabilities and visualised in attack trees or graphs); (ii) bottom-up (from post-development vulnerability analysis); and (iii) a combination of both.

*5) Impact Rating:* This step represents an estimation of magnitude of damage conveying the severity associated to a DS [15]. The standard computes the impact of each DS, based on the sum of four parameters: safety($i_s$), financial($i_f$), operational($i_o$) and privacy($i_p$) whose values are assigned using a numerical scale (0, 1, 10, 100), depending on the severity (negligible, moderate, major, severe) [26]. The impact rating in ISO/SAE 21434 as well as in most common TARA methodologies [16], [17], [30] is approached using the Equation 1 [26]:

$$I = 10(i_s + i_f) + i_o + i_p, \tag{1}$$

where each parameter is associated to an assigned weight that is set to 10 for safety and financial impacts and to 1 for operational and privacy impacts. According to the standard, this is justified with the criticality of safety and financial impacts over the other impacts. The derived $I$ is mapped to determine the overall impact, per DS, through the impact level $L_I$ according to Table II.

*6) Attack Feasibility Rating:* Each AP details the attacker, attack surface and method used to assess attack feasibility as quantifiable parameters. ISO/SAE 21434 outlines three approaches to translate such information into quantifiable parameters: (i) attack potential-based (as broken down in Section IV-E), (ii) Common Vulnerability Scoring System (CVSS)-based, and (iii) attack vector-based.

Despite the chosen method, an attack feasibility level ($L_F$) is derived. Table II illustrates how the aggregation is conducted using the attack potential-based method where an attack feasibility sum ($F$) is computed and mapped afterwards to $L_F$.

TABLE II

IMPACT RATING (LEFT) AND ATTACK FEASIBILITY RATING (RIGHT) AS DEFINED BY ISO/SAE 21434 IN TARA 1.0

| **Impact sum ($I$)** | **Impact level ($L_I$)** | **Attack feasibility sum ($F$)** | **Feasibility level ($L_F$)** |
|---|---|---|---|
| 0 | 0 - none | $\geq$10 | 0 - very low |
| 1-19 | 1 - negligible | 7-9 | 1 - low |
| 20-99 | 2 - moderate | 4-6 | 2 - medium |
| 100-999 | 3 - major | 2-3 | 3 - high |
| $\geq$ 1000 | 4 - severe | 0-1 | 4 - critical |

TABLE III

RELATED WORK COMPARISON

| **Related work** | **ISO/SAE 21434** | **Enhancing avenues** | | | | | **Demonstration** |
|---|---|---|---|---|---|---|---|
| | | Privacy | SAE Lx | Objectivity | Safety | Others | |
| [31] | ✗ | ✗ | ✓ | ✗ | ✗ | ✓ | ●●○ |
| [32] | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ○○○ |
| [19] | ✗ | ✓ | ✓ | ✗ | ✗ | ✗ | ●●○ |
| [33] | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ●○○ |
| [34] | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ | ●○○ |
| [35] | ✗ | ✗ | ✗ | ✓ | ✓ | ✗ | ●●○ |
| [36] | ✓ | ✗ | ✗ | ✗ | ✓ | ✗ | ○○○ |
| [37] | ✓ | ✗ | ✗ | ✗ | ✓ | ✗ | ○○○ |
| [38] | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ●○○ |
| [39] | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ●●○ |
| [40] | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ●●○ |
| [30] | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ●●● |
| [11] | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ●○○ |
| [41] | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ●○○ |
| [42] | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ●●○ |
| [12] | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ●○○ |
| [43] | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ●●○ |
| [44] | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ●●● |
| [45] | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ●●○ |
| **This work** | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ●●● |

○○○: not demonstrated; ●○○: limited demonstration; ●●○: moderate demonstration; ●●●: extensive demonstration

*7) Risk Determination:* The risk value and level determination proposed by ISO/SAE 21434 combines the impact level ($L_I$), derived from the impact rating, and the feasibility level ($L_F$), derived from the feasibility rating, as follows:

$$R(L_I, L_F) \tag{2}$$

The levels and values are retrieved from Table II to construct the risk matrix.

*8) Risk Treatment Decision:* Determines if the risk can be: (i) reduced (using security or privacy controls), (ii) accepted (without additional measures), (iii) shared (delegated to a third party), and (iv) avoided (by eliminating the source).

Despite the decision taken, it is crucial that all the CAV stakeholders [27], including OEM, get involved in the treatment decision process as associated technological, operational or financial costs may apply. Additionally, the risk should remain monitored within any TARA reiteration [15].

*9) Cybersecurity Goals and Claims:* This step relies on TARA outcomes. While the risk treatment decision addresses the identified risks, cybersecurity goals and claims focus on verifying and validating the treatment.

## III. RELATED WORK

Improving TARA relates to understanding five research fields gathering privacy assessments, the inclusion of SAE level, the involvement of experts knowledge, enhanced frameworks and the quality of the process demonstration.

### A. Privacy Assessment

With the significant volume of data exchanged within the CAV's environment, privacy assessments have recently begun to attract increasing research attention. For instance, Chah et al. [11] and Azam et al. [12] conducted privacy-focused threat modelling using LINDDUN and STRIDE, respectively, but they were limited to qualitative analyses and did not encompass complete risk assessment processes. Closely aligned with the current research, Monteuuis et al. [19] introduced a systematic TARA framework for L3 CAVs with an extended threat model incorporating privacy aspects like linkability. However, it used inconsistent terminology and redundant categories, reflecting limitations of earlier TARA iterations. Consequently, these studies highlight the need for a comprehensive approach to privacy motivating the development of TARA 2.0.

### B. Security Assessments and the SAE Automation Level

Following the discussion on TARA methodologies preceding the final ISO/SAE 21434, some researchers showcased the correlation between the SAE level and the attacks severity. Dominic et al. [31] and He et al. [34] are among the pioneering researchers raising such concerns. Nevertheless, the authors discussed the significance of the SAE levels in a cause-and-effect manner, lacking quantified measurements. Besides, Bolovinou et al. [33] proposed TARA+ as an improved TARA framework that was intended for highly automated vehicles of SAE L3 onward. The methodology suggested an enhancement of the preliminary TARA 1.0 version by considering a driver controllability factor within the impact rating step. While the viability of the proposed approach is sound, its main limitation is that TARA+ is applicable to CAVs under a driver intervention which discards L4 and L5 CAVs settings. In contrast, TARA 2.0 not only addresses the limitations of previous approaches by incorporating the automation level factor, but also accounts for the in-vehicle and remote human intervention under different SAE levels, hence different ODDs.

### C. Experts Subjectivity

Assessing the experts subjectivity has been an ebb and flow at the general scope of risk assessment domain. Wen et al. [35] proposed a flexible safety risk assessment framework incorporating subjective and objective weights to address hesitant information provided by experts. From the statistics domain, O'Hagan [46] introduced a knowledge elicitation principle that quantifies expert uncertainty, independence, and multiplicity through probability distributions. In a closer work to the CAV's domain, Khastgir et al. [32] highlighted the influence of expert experience and cultural biases on HARA results through a conducted workshop, emphasising the need to automate risk analysis and minimise expert reliance. While the experts objectivity has been extensively addressed in several critical infrastructures domains like energy and aeronautics [47], there is a lack of literature tackling its application to CAVs as well as its consideration within TARA. TARA 2.0 directly addresses this gap by reducing expert subjectivity through the introduction of quantified metrics and providing a more transparent representation of the level of expert involvement.

### D. Further Enhanced TARA Frameworks

Being highly aware of the TARA limitations within the CAV's landscape [30], researchers oriented their efforts towards enhancing TARA either by combining it to other risk assessment methodologies, by tweaking its threat modelling or by extending its risk factors. Agrawal et al. [36] proposed Threat/Hazard Analysis and Risk Assessment (THARA) to unify security and safety concepts by integrating the controllability metric from Automotive Safety Integrity Levels (ASIL) [48] into HARA for addressing safety-critical attacks in L3 onwards CAVs. Similarly, Vogt et al. [37] combined Failure Mode and Effects Analysis (FMEA) from HARA with TARA for quantitative risk analysis, while Dobaj et al. [38] merged TARA 1.0 with Automotive Software Process Improvement and Capability dEtermination (A-SPICE) [49] to derive cybersecurity requirements iteratively. However, these joint models lack concrete demonstrations. With a focus on CAV's perception systems, Ghosh et al. [43] combined TARA 1.0 with STPA-Sec [50], adding a robustness factor for AI performance and a mitigation factor to refine risk prioritisation. This approach addresses specific threats to perception systems but complicates risk calculations.

While these research works opted for joint models, others improved TARAs through the automation of certain subprocesses. Schmittner et al. [39] and Ebrahimi et al. [41] proposed a refinement of the threat modelling stage by automating the threat scenarios and attack paths generation respectively using ThreatGet tool. Similarly, Zelle et al. [42] proposed a semi-automated attack paths generation using a different software entitled ThreatSurf. Nevertheless, the suggested tools rely on additional manual work for the input preparation. TARA 2.0 aligns with the stated research in recognising the need to expand the traditional TARA 1.0 by incorporating additional concepts; however, it intentionally avoids merging with the safety assessment process in order to maintain its focus on cybersecurity and data privacy threats.

### E. Demonstrated TARA

Understanding how to execute a TARA in CAV's environment is crucial for valid risk treatment and cybersecurity requirements elicitation. However, with the multiple sub-steps and the collection of factors including impact, likelihood, and attacker profiles analysis, TARA 1.0 is foreseen as complex and no-ready-to-use methodology [37]. Several works aimed to address this shortcoming through demonstrations (Table III). For instance, Abuabed et al. [44] proposed a framework with in-depth threat modelling, and Plappert et al. [40] focused on fine-grained attack feasibility analysis. However, while these authors demonstrated TARA 1.0 on modern vehicles, they did not provide detailed analysis for the entire process. Loskin [45] offered insights into implementing TARA 1.0 with clear documentation, though the resulting assessment remained confidential. In a more comprehensive work, Lautenbach et al. [30] enhanced HEAVENS [17] to align with ISO/SAE 21434 requirements, though the level of automation and human intervention remain at a high level of abstraction. To that end, TARA 2.0 highlights the process complexity where unclear steps can cause misinterpretations

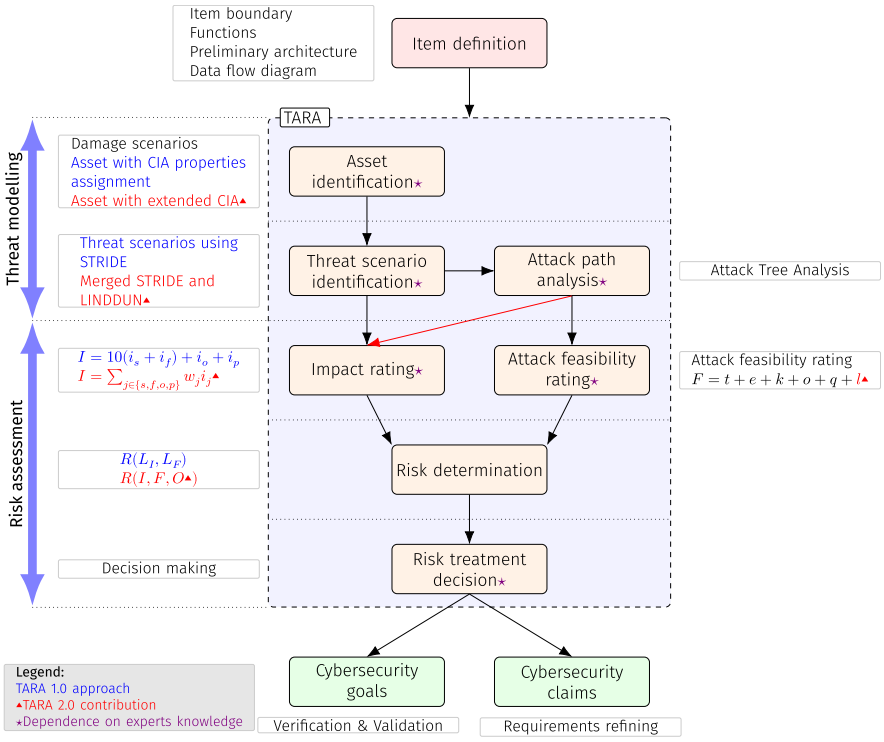



Fig. 1. TARA 2.0 improvement avenues.

and affect decisions. It also offers step-by-step guidance for CAV's stakeholders to replicate the framework.

## IV. TARA 2.0

### A. Methodology

The methodology was built based on the analytics and experimental findings upon TARA 1.0 limitations. First, originating from systematic review of Benyahya et al. [21], TARA 1.0 was established as the most prominent approach to assess CAVs vulnerabilities. However, an appropriate adaptation to tackle the SAE L4 and L5 CAVs' specifications including privacy threats and automation level implications is required. The same review asserted that TARA 1.0 outcomes rely on experts' opinion, prone to the assessment's subjectivity. To overcome such pitfalls, every step from TARA 1.0 has been evaluated and experimented as a baseline assessment [51]. The results showcased that: (i) Asset and threat scenario identification steps model mainly security threats with little attention to privacy. (ii) The privacy impact is underestimated with lower weights than safety and financial impacts (Equation 1). (iii) The SAE level is not considered in the assessment. (iv) 6 out of 7 TARA 1.0 steps depend on experts opinion (as notated by symbol ⋆ in Figure 1).

To that end, by providing a comprehensive assessment of privacy threats at the threat modelling phase; proposing an agile weight to the privacy impact depending on the nature of the data; considering the SAE level as one of the RA metrics; and quantifying the experts' objectivity then the overall process will be improved to address L4 & L5 specifications.

The contribution brought forward by TARA 2.0, highlighted in red and marked with ▲ within Figure 1, consists of integrating improvements to several steps of TARA 1.0. At the asset identification step (IV-B), privacy and security goals are consolidated, to extend the CIA model, for an in-depth identification of DSs. At the threat scenario identification (IV-C), the threat modelling is more extensive by considering eleven threat classes as an outcome of fusing Spoofing, Tampering, Repudiation, Information disclosure, Denial-of-service and Elevation of privilege (STRIDE) and Linkability, Identifiability, Nonrepudiation, Detectability, Disclosure of data, Unawareness, and Noncompliance (LINDDUN) techniques instead of limiting the analysis just to STRIDE classes. The impact rating is improved by expanding the privacy impact assessment through the integration of a weight representing data sensitivity and Privacy Enhancing Technologiess (PETs) solutions in place (IV-D). The attack feasibility rating step is enhanced by incorporating the SAE level (IV-E). The inclusion of the experts' objectivity index elevates further the risk determination milestone to consider any expert's subjective influences (IV-F). The provided improvements apply to core steps of the TARA excluding the item definition, cybersecurity goals and claims steps.

### B. Asset Identification

Guided by the privacy risks within the CAV ecosystem and the General Data Protection Regulation (GDPR) data processing principles [52], the ISO/IEC 27000 [28] privacy goals and the privacy and data by design principles from European Union Agency for Cybersecurity (ENISA) [53], the CIA model is extended by incorporating three privacy protection goals which are unlinkability (U), accountability (Ac) and compliance (Com). Unlinkability is compromised when privacy-related data can lead to identify the data owner. Accountability is intertwined with transparency and non-repudiation concepts, where the originated entities of a claimed

action or event can be proven [12]. Compliance indicates the appropriate integration of relevant privacy policies such as the GDPR. To that end, joining privacy goals to the CIA model is perceived as building blocks towards constructing both security and privacy threat modelling in the TARA's following steps.

### C. Threat Scenario Identification

A holistic threat modelling for health data context was successfully demonstrated by Treacy et al. [54]. The same principles are applied in TARA 2.0 by combining STRIDE [55], for its automated functionality in instantiating security threats, and LINDDUN, for its pertinence to privacy threats [11], as detailed in Appendix A. The consolidation of both approaches led to a categorisation of eleven threat classes where repudiation/ non-repudiation and information disclosure/ disclosure of information classes are merged respectively. Then, every threat scenario is mapped to one or several threat classes. Every intersection between a threat class, which can be one of the categories of STRIDE or LINDDUN classes, and the threat scenario is further developed and sketched through the attack path analysis. It is important to highlight that the attack path analysis step is unchanged in TARA 2.0, justifying a direct transition to the impact rating step.

### D. Impact Rating

The ISO/SAE 21434 proposes an impact rating per DS. Therefore, as soon as DSs are defined, the impact can be computed. In contrary to the standard, the present work puts forth an impact rating per AP, which occurs after accomplishing both the threat scenario identification and attack path analysis steps. Consequently, a fine-grained risk determination can be also performed afterwards per AP. The detailed examination allows also for targeted mitigation strategies to be determined for each individual AP instead of being derived per DS. Considering the one-to-many relationship between DSs, threat scenarios and APs, eliciting decisions per DSs mirror a more generalised assessment while an assessment per AP provides more granularity. Additionally, assessing the risk for individual APs allows fewer aggregations at both impact rating and attack feasibility rating phases.

Considering that TARA 1.0 follows the ISO 26262 impact rating where the privacy impact is underestimated (Section II-B.5), TARA 2.0 advocates a weighted impact rating, as in [31] and in Equation 3, where every impact category is joined to a specific weight. While in Equation 1 safety weight $w_s$ and $w_f$ are fixed to ten and $w_o$ and $w_p$ are set to one, each impact weight is set in the range [0, 10] (Equation 3).

$$I = \sum_{j \in \{s,f,o,p\}} w_j i_j \tag{3}$$

While focusing on the $i_p$ and $w_p$, by extending the ISO/IEC 29100 [56], which elicited the privacy rating for privacy damage as referred in the ISO/SAE 21434 appendix, to bring forth a granular privacy impact assessment addressing three factors: data sensitivity, linkability to PII and the PET solutions implemented in place. As defined in Table IV, data sensitivity levels are established to be: highly sensitive, medium and not sensitive just as the ISO/IEC 29100 [56] classification. Similarly, the linkability to PII is scaled as easy or difficult to link. The PET factor, determining the presence of methods such as anonymisation and pseudonymisation [57], is classified into: none, partial or strong implementation. The $i_p$ and $w_p$ are generated afterwards depending on the combination, of the three privacy factors, which are appropriate to the assessed context. The derived $i_p$ and $w_p$ are aggregated to the other impacts $i_s$, $i_f$ and $i_o$ with their relevant weights ($w_s$, $w_f$ and $w_o$ ) which are assigned based on gathered consensus from stakeholder and panel of experts [35]. To that end, Equation 3 determines the overall impact ($I$) from which the impact level is derived, if required, according to Table II.

TABLE IV
TARA 2.0 PRIVACY IMPACT SCORING

| Level | Privacy factors | | | $i_p$ | $w_p$ |
|---|---|---|---|---|---|
| | Sensitivity | Linkability | PET | | |
| Severe | Highly sensitive | Easy | None | 100 | 10 |
| | Highly sensitive | Easy | Partial | 100 | 9 |
| | Highly sensitive | Easy | Strong | 100 | 8 |
| Major | Highly sensitive | Difficult | None | 10 | 9 |
| | Medium | Easy | None | 10 | 8 |
| | Highly sensitive | Difficult | Partial | 10 | 7 |
| | Medium | Easy | Partial | 10 | 6 |
| | Highly sensitive | Difficult | Strong | 10 | 5 |
| | Medium | Easy | Strong | 10 | 5 |
| Moderate | Medium | Difficult | None | 1 | 5 |
| | Not sensitive | Easy | None | 1 | 4 |
| | Medium | Difficult | Partial | 1 | 3 |
| | Not sensitive | Easy | Partial | 1 | 2 |
| | Medium | Difficult | Strong | 1 | 2 |
| | Not sensitive | Easy | Strong | 1 | 1 |
| Negligible | Not sensitive | Difficult | None | 0 | 1 |
| | Not sensitive | Difficult | Partial | 0 | 1 |
| | Not sensitive | Difficult | Strong | 0 | 0 |

### E. Attack Feasibility Rating

This study adopts the attack potential-based approach from ISO/IEC 18045 [58], as it considers attackers' capabilities and intentions, unlike CVSS and attack vector-based methods limited to software and network vulnerabilities. The attack potential-based approach captures factors leading to a successful attack [15] which are: (i) Elapsed time ($t$) indicating the attack exploitation time in months or years; (ii) Specialised expertise ($e$) determining the attacker(s) capabilities; (iii) Knowledge of the item/component ($k$) evaluating the attacker's acquired information; (iv) Windows of opportunity ($o$) defining the target accessibility type and duration; and (v) Equipment ($q$) indicating the tool properties supporting the attack execution.

These factors are extended further by including the CAV automation level (SAE Lx ($l$)) as an additional constituent impacting the attack feasibility rating as depicted in Table V. The $l$ in this context considers the human factor, which can be either an in-vehicle driver or a remote operator, in reducing the associated risk. With this notion, the attack feasibility encompasses the transition between human and machine,

TABLE V
TARA 2.0 FEASIBILITY RATING

| **Elapsed time** ($t$) | **Specialised expertise** ($e$) | **Knowledge of the item/component** ($k$) | **Windows of opportunity** ($o$) | **Equipment** ($q$) | **SAE Lx** ($l$) | **Assigned value** |
|---|---|---|---|---|---|---|
| $\leq$ 1 month | Layman | Public | Unlimited | Standard | L5 | 0 |
| $\leq$ 6 months | Proficient | Restricted | Easy | Specialised | L4 $\star$ | 1 |
| $\leq$ 3 years | Expert | Confidential | Moderate | Bespoke | L4 $\star\star$ | 2 |
| $\geq$ 3 years | Multiple | Strictly confidential | Difficult | Multiple bespoke | L3 | 3 |

$\star$: Remote operator; $\star\star$: In-vehicle operator

TABLE VI
THE SCALED ATTACK FEASIBILITY RATING UPON THE INCLUSION OF THE SAE LX METRIC IN TARA 2.0

| **Attack feasibility sum** ($F$) | **Feasibility level** ($L_F$) |
|---|---|
| $\geq$12 | 0 - very low |
| 8-10 | 1 - low |
| 4-7 | 2 - medium |
| 2-3 | 3 - high |
| 0-1 | 4 - critical |

exploring how this transition is likely to influence a successful attack.

Each factor in Table V has four possibilities which are associated to numerical values varying from zero to three. The smaller the value, the more likely the attack is to occur. The proposed attack feasibility rating is aligned to the ISO/IEC 18045 and ISO/SAE 21434, but uses a light-weighted numerical scaling (from 0 to 3) as in [30] unlike the larger scale (from 0 to 19) proposed in the standard [58]. For every perceived AP, the attack feasibility ($F$) is computed by summing up the scaled values for every parameter as in Equation 4:

$$F = t + e + k + o + q + l \tag{4}$$

Thereupon, Table II values are scaled to the inclusion of the automation level metric to allow a correct elicitation of the five feasibility levels from the new attack feasibility sum $F$. An interpolation is used for a linear transformation to reflect the change of $F$'s composition. Table VI depicts how the relationship between $F$ and $L_F$ is preserved through the new scaling which differs from Table II on the boundaries of medium, low and very low $L_F$. While the Table proposes the new mapping, it is noteworthy to mention that for the purpose of this research, TARA 2.0 exploits mainly $F$ without compiling the $L_F$ for the risk determination and visualisation discussed in the following subsection. Such decision is adopted as the $F$ provides the needed granular value to construct the 3-D plot (SectionV-H). Consequently, Table VI is incorporated to the present work just to support any further replication aiming a classical risk matrix with the use of $L_F$.

### *F. Risk Determination*

In addition to the impact sum $I$ and the attack feasibility sum $F$, TARA 2.0 extends further the risk determination by incorporating a third metric $O$, referred to as the experts' objectivity index:

$$R(I, F, O) \tag{5}$$

The experts' objectivity index ($O$) takes into account the experts' subjectivity while making assumptions throughout the TARA process. Following the experts' knowledge elicitation principle [46], commonly used in the statistics domain, along with the ISO/IEC 17065 requirements on impartiality [59], the $O$ index compiles four factors AP: (i) $c$: certainty (representing the experts' confidence); (ii) $r$: peer review (determining if the analysis is conducted by one experts' group or several groups); (iii) $m$: measurable tools (indicating the usage of measurable metrics or automation tools); and (iv) $p$: impartiality (demonstrating if experts may provide any unfair or biased inputs due to their affiliations).

A rate from the [0,1] interval is assigned by the experts to each factor where values close to 1 represent higher confidence and hence objective opinion while values close to 0 illustrate low confidence and hence subjective opinion. A mean value, representing the objectivity index $O$, of the four factors is computed as follows:

$$\bar{O} = \frac{\sum(c, r, m, p)}{n} \tag{6}$$

The risk matrix combines the three values ($I$, $F$, $O$), using a three-dimensional plot where the x-axis represents the experts' objectivity index ($O$), the y-axis depicts the attack feasibility sum ($F$) and finally, the z-axis draws the impact sum ($I$) as demonstrated in Figure 5. Such visualisation simplifies the process of risk prioritisation where the tallest dots with the smallest $y$ values and the highest $x$ values indicate most potential APs or threats to prioritise.

Similar to the attack path analysis step, no enhancements were introduced at the risk treatment decision step, justifying its omission. Notably, TARA 2.0 adheres to the same requirements as TARA 1.0 for this step.

## V. PROOF OF CONCEPT (POC)

This section demonstrates TARA 2.0's usage through illustrative example of its applicability over the ADS as a PoC. Following the proposed methodology, the demonstration considers a real context inside the ULTIMO project [60] where L4 CAVs are designated for people and goods transportation as an integral solution for public transport systems. To that extent, the PoC consists of executing TARA 2.0 at the design stage, over a reference architecture envisioned to fulfil the ULTIMO project objectives.

### *A. Materials and Tools*

To implement the PoC, specific tools were selected to conduct an assessment from the outset including (i) Google sheet and MS Excel scripts: used as an inventory tool (for

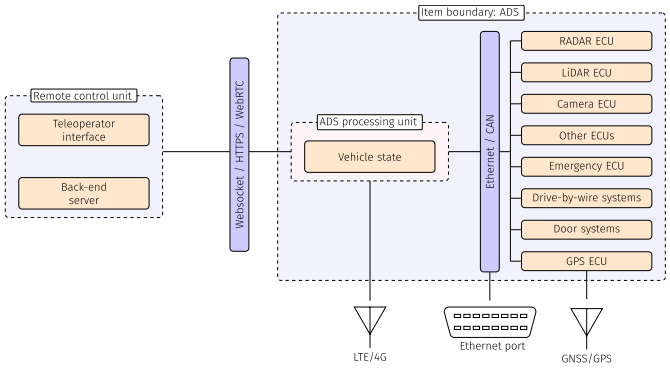



Fig. 2. Preliminary SAE L4 architecture for TARA 2.0.

damage and threat scenarios elicitation) as well as for risk calculation and prioritisation; (ii) Microsoft Threat Modeling Tool (TMT7): used to build DFDs and derive STRIDE threat scenarios; and (iii) LINDDUN documentation [61]: used to elaborate LINDDUN threat scenarios and their related Attack Tree Analysis (ATA). The remainder of this section describes the outcome of every TARA 2.0 step.

### B. Item Definition

As depicted in Figure 1, the item definition step is the starting point of any TARA process. According to the ISO/SAE 21434 recommendations, the following work products are determined at that step:

1) Item boundary: the ADS processing unit.
2) Functions: the ADS takes autonomous motion decision by merging and cross-checking data from all sensors. The output includes vehicle motion wrapping up localisation status, object detection and path planning, along with doors control and the transition between manual or remote driving.
3) Preliminary architecture: Figure 2 depicts a generic architecture for the L4 CAV including the buses (Controller Area network (CAN) or Ethernet), and ECUs connected to the ADS processing. The architecture depicts external connections such as (4G/WiFi) and GNSS/GPS. The figure highlights the crucial connection to the manual control unit which is triggered once the human intervention (in-vehicle or remote) is required.
4) Data Flow Diagram: Figure 3 describes the different data flows exchanged with the ADS processing unit. Each data flow is represented with a one way arrow while processes are depicted on a circle design. Devices and data stores are drawn using two parallel horizontal lines. The interactors that can be either a passenger, developer, maintenance user, service customer or a third party user are depicted in a rectangular shape. The trust boundary, where trust flows occur without the use of encrypted connections to the ADS, are drawn using red dotted rectangle. Interfaces like vehicle user interface and internet are represented using a dotted line boundary. The DFD was elaborated using TMT7 tool and with an alignment to the software standards symbols [55]. The sketched DFD is built using the automotive template [62].

### C. Asset Identification

The required work product for the asset identification consists of the identification of assets, their association to cybersecurity properties and their mapping to DSs. Based on the system architecture and the assets derived from the DFD diagram, Table VII lists a sample from the 14 valuable assets within the ADS boundary which can be a safety-critical function (like A.1) or a data set (A.14). For every asset, DSs are defined, where multiple DSs can be related to a single asset. For instance, DSs: D.1, D.2, D.3 relate all to A.1. As introduced in Section IV-B, the table leverages the CIA model further by mapping the assets to additional privacy goals: unlinkability (U), accountability (Ac) and compliance (Com). To illustrate, compromising the Global Navigation Satellite System (GNSS) processing (A.1) does not only impact the CAV's integrity and availability but it affects the unlinkability privacy goal if the vehicle location data can be linked to end-users' sensitive data leading to a privacy-related DS (D.3).

### D. Threat Scenario Identification

For a holistic threat modelling, an automated interaction-based analysis was considered with STRIDE and TMT7, combining its findings to the element-centric analysis from LINDDUN. On the one hand, 550 threats were reported from the TMT7 which are defined using an ID, a description and a categorisation through the STRIDE threat classes. For our analysis, the report is exported to a comma-separated values (CSV) file where the threats' list is filtered to remove redundant entries and is grouped per asset to match the predefined DSs. The filtering criteria consists of omitting threats which: (i) are duplicated for the two ways (in/out) of data flows between the same components; and (ii) can be adequately mitigated through existing security controls. The threats synthesising and grouping led to 40 threats scenarios as sampled in Table IX. On the other hand, as in LINDDUN threats must be determined by the type of DFD elements, every retained threat, from the 40 threat scenarios, is then extended with an explicit linking to the DFD elements to define the privacy threat class related to it. Following [63], a template is constructed (Table VIII) suggesting which threat class, from LINDDUN categories, is relevant to each DFD element based on data types descriptions provided with the system architecture. Hence, every threat scenario is mapped to both security and privacy threat classes at the end of that process.

In a nutshell, Table IX illustrates the mapping between DSs and the eleven threat classes, as a combination of the STRIDE and LINDDUN results. Every x in the table implies that the corresponding threat scenario is susceptible to the selected threat classes. Additionally, each x represents an interaction that needs to be elevated in the form of an ATA. The interactions annotated with ⓧ are selected to be showcased in the following subsection. For instance, a spoofing scenario was chosen as a well documented cybersecurity threat [8] and a linkability scenario is selected for being a prominent privacy threat in the CAV landscape [11].

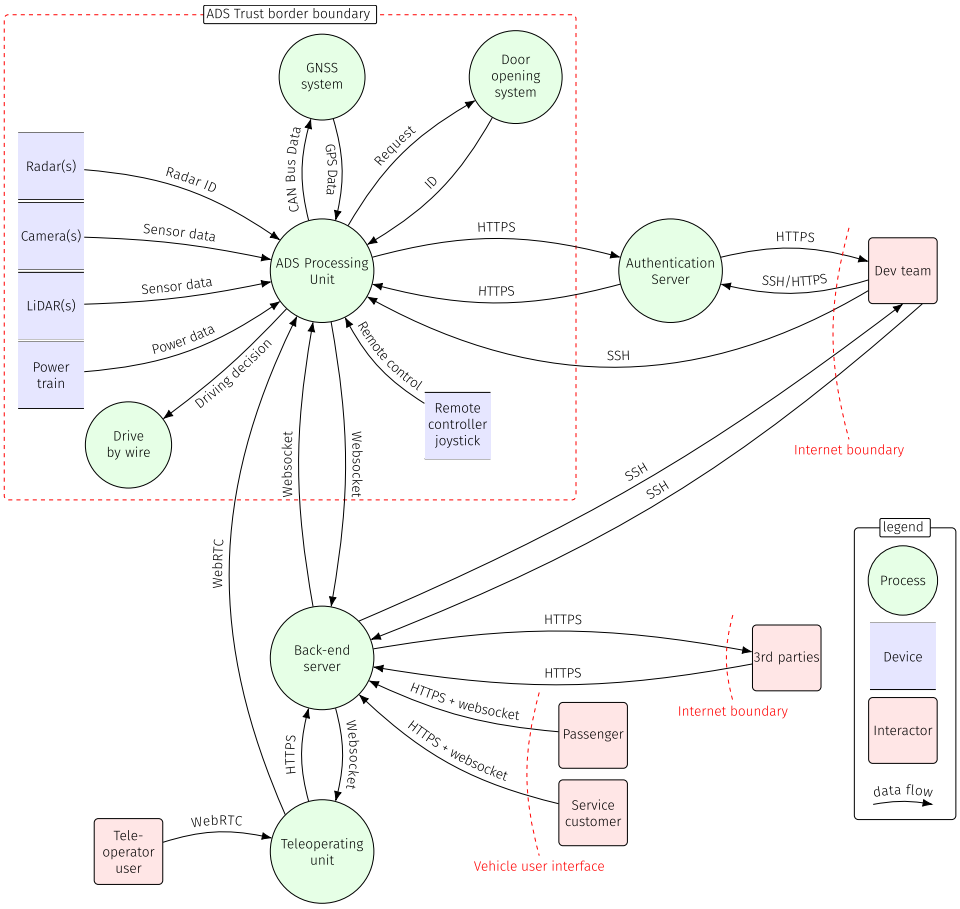



Fig. 3. ADS data flow diagram.

TABLE VII
TARA 2.0 ASSET IDENTIFICATION

| Asset # | Asset description | Cybersecurity properties C | I | A | *U* | *Ac* | *Com* | Damage # | Damage Scenario (DS) |
|---|---|---|---|---|---|---|---|---|---|
| A.1 | GNSS capturing and processing | | x | x | x | | | D.1<br>D.2<br>D.3 | Compromising GNSS signal<br>Lost of GNSS signal<br>Unauthorised location tracking |
| A.2 | Image capturing and processing | x | x | | x | | x | D.4<br>D.5<br>D.6 | Blinded vision<br>Presume non-existent obstacles<br>Unauthorised facial image capturing |
| ... | ... | | | ... | | | | ... | ... |
| A.14 | Passenger data | x | | x | x | x | x | D.25<br>D.26<br>D.27 | Disclosure of PII without consent<br>PII de-anonymisation<br>Malicious data manipulation |

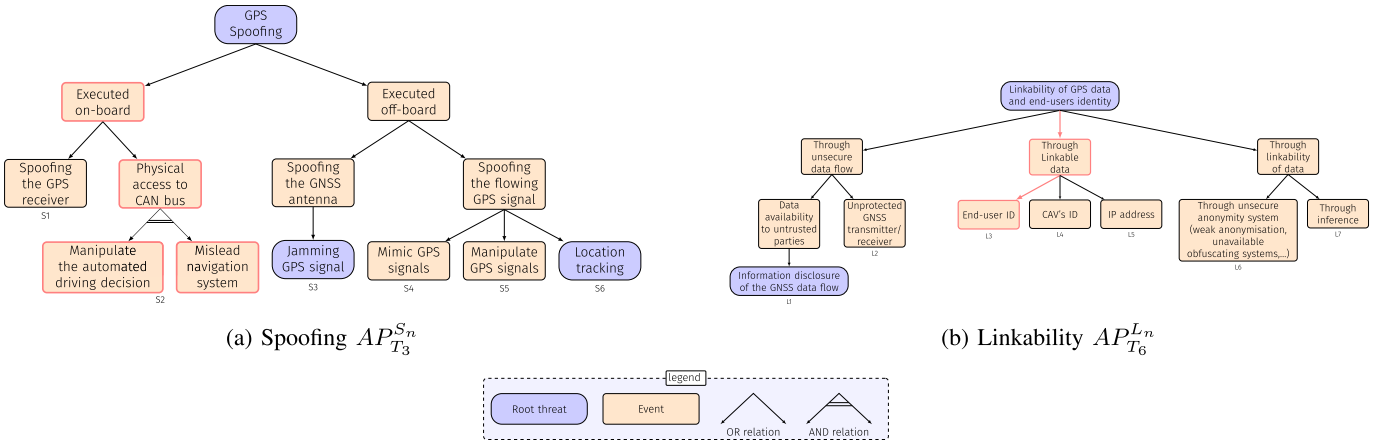
TABLE VIII
MAPPING LINDDUN THREATS TO DFD ELEMENTS

| DFD element | L | I | N | D | D | U | N |
|---|---|---|---|---|---|---|---|
| Data flow | x | x | x | x | x | | x |
| Process | | | x | x | | | x |
| Entity | x | x | | | | x | |

## *E. Attack Path Analysis*

For the purpose of the present work, the top-down approach is selected through ATA (Section II-B.4). To maintain simplicity, attack trees are adopted over attack graphs as the former are easier to understand and depict simple event flows showcasing the different ways an attacker can follow to achieve the attack, while the latter is more resource-intensive and involves the interconnected relationship between vulnerabilities which is more appropriate for highly complex systems [64].

ATA helps in identifying the significance of a threat to the system. By considering the 11 threat classes, every threat class per threat scenario should generate several attack paths leading to the threat execution. Such attack tree elicitation is conducted based on knowledge of the system architecture from the item definition step, the list of threats in the




(a) Spoofing $AP_{T_3}^{S_n}$ (b) Linkability $AP_{T_6}^{L_n}$

Fig. 4. Spoofing and linkability attack trees.

TABLE IX
TARA 2.0'S THREAT SCENARIO IDENTIFICATION ILLUSTRATION USING STRIDE AND LINDDUN

| Damage # | Threat # | Threat scenario | DFD element | Threat classes S | T | R / N† | I / D†† | D | E | L | I' | D' | U | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| D.1 | T.1 | Compromise GNSS to deliver malicious updates | Data flow | | | x | x | | x | x | x | x | | x |
| | T.2 | Flood GNSS with invalid data | Data flow | | | x | x | x | | x | x | x | | x |
| | T.3 | Spoof GPS signals to manipulate the vehicle | Data flow | ⓧ | | x | x | | | x | x | x | | x |
| D.2 | T.4 | Take the GNSS offline | Process | | | x | | x | | | x | x | | x |
| | T.5 | Jam the GPS signal, causing a DOS on the antenna | Data flow | | | x | x | x | | x | x | x | | x |
| D.3 | T.6 | Link CAV location to the end-user identity | Entity | | | | x | | | ⓧ | x | | x | |
| ... | | | | | | | | | | | | | | |
| D.27 | T.40 | Tamper with data in transit to the ADS | Data flow | | x | x | x | | | x | x | | | x |

†: merged **R**epudiation and **N**on repudiation. ††: merged **I**nformation disclosure and **D**isclosure of information. X: an interaction showing a mapped threat scenario to a threat class. ⓧ: a demonstrated interaction in Figure 4 and Table X.

UNECE R155 annex [13], known Common Vulnerabilities and Exposures (CVEs) [65] relevant to CAVs and attacks taxonomies [6]. For simplicity purposes, we scrutinise a spoofing interaction demonstrating a security threat (T.3) and a linkability interaction illustrating a privacy threat (T.6) as depicted in Figure 4a and 4b respectively. Consequently, every path from the parent to the child node indicates a valid attack path where AND relation yields to one AP (as exemplified by our unique AND case in $AP_{T_3}^{S_2}$) while the OR relation produces distinct APs. For instance, the spoofing tree generates six attack paths starting from $AP_{T_3}^{S_1}$ to $AP_{T_3}^{S_6}$ and the linkability tree demonstrates seven attack paths from $AP_{T_6}^{L_1}$ to $AP_{T_6}^{L_7}$. The upper script on $AP$ designates the threat class combined to the AP number within the evaluated threat scenario while the lower script shows the threat scenario ID. To illustrate, the $S_1$ in $AP_{T_3}^{S_1}$ refers to the first AP of spoofing as a threat class related to the third threat scenario $T_3$.

### F. Impact Rating

Table X illustrates how the impact value is calculated for the two selected APs: $AP_{T_3}^{S_2}$ and $AP_{T_6}^{L_3}$ using Equation 3. First, the value is defined for safety($i_s$), financial($i_f$) and operational($i_o$) factors mirroring the experts assessment (Negligible- 0, Moderate- 1, Major- 10 or Severe- 100). Second, the weights are distributed by the experts respectively with regard to the impact factor importance to the AP. For instance, the operational weight ($w_o$) is at its highest value (10) for $AP_{T_3}^{S_2}$ in case an attacker is spoofing the GNSS data directly from the CAN bus compromising the ADS integrity and its well functioning. However, $w_o$ has a lower weight of 1 in $AP_{T_6}^{L_3}$ as the compromise of the unlikability property does not immediately impact the vehicle route. Regarding the privacy, both impact ($i_p$) and weight ($w_p$) are retrieved using Table IV based on the assessed combination of privacy factors. By summing up all the impacts, the $I$ leads to the impact level ($L_I$) identification according to Table II.

While Table X depicts the impact rating for both selected APs, we discuss here the impact rating performance for $AP_{T_6}^{L_3}$ in comparison to TARA 1.0. TARA 2.0 leads to an amber Moderate impact of level 3 for $AP_{T_6}^{L_3}$. However, by following TARA 1.0 impact calculation in Equation 1, the $I$ would be equal to 10=10x(0+0)+0+10 (assuming a Negligible $i_s$, $i_f$ and $i_o$ with a Major $i_p$) resulting into a yellow Negligible $L_I$ of level 1 (according to Table II) discriminating the privacy importance in such attack path. To that end, our example illustrates how an appropriate privacy impact and weight can change the entire impact rating outcome.

### G. Attack Feasibility Rating

Following the attack potential-based approach [15], and with the purpose to address the SAE automation level within the assessment, TARA 2.0 proposes a simplified attack feasibility rating which wraps up the SAE Lx ($l$) as an additional metric with an alignment to Table V. By applying

TABLE X
TARA 2.0 ANALYSIS ON TWO ILLUSTRATED ATTACK PATHS

| | | |
|---|---|---|
| **Damage Scenario** | D.1 Compromising GNSS signal. | D.3 Unauthorised location tracking. |
| **Threat scenario** | T.3 Spoof GPS signals and deliver malicious GPS data or to manipulate the vehicle. | T.6 An attacker relating the CAV location to the end-user identity for unauthorised location tracking. |
| **Attack Path** | $AP^{S_2}_{T_3}$ Physically connecting to the vehicle CAN bus to manipulate the automated driving function and mislead the navigation, compromising the CAV integrity and availability. | $AP^{L_3}_{T_6}$ Linking GPS data and end-users identity through linkable end-user ID. |
| **Impact** | $w_s$ = 10, $i_s$ = 10 (Major)<br>$w_f$ = 5, $i_f$ = 10 (Major)<br>$w_o$ = 10, $i_o$ = 10(Major)<br>$w_p$ = 9, $i_p$ = 10 (highly sensitive/difficult to link/none)<br>$I = \sum_{j\in\{s,f,o,p\}} w_j i_j$= 340<br>$L_I$ = 3 - Major | $w_s$= 5, $i_s$ = 0 (Negligible)<br>$w_f$ = 1, $i_f$ = 0 (Negligible)<br>$w_o$ = 1, $i_o$ = 0 (Negligible)<br>$w_p$ = 9, $i_p$ = 10 (highly sensitive/difficult to link/none)<br>$I = \sum_{j\in\{s,f,o,p\}} w_j i_j$= 90<br>$L_I$ = 2 - Moderate |
| **Attack feasibility** | $F = t + e + k + o + q + l$ = 0+1+0+2+0+2=5<br>$L_F$ = 2 - medium | $F = t + e + k + o + q + l$ = 1+2+1+2+0+2= 8<br>$L_F$ = 1 - low |
| **Objectivity index** $O$ | Mean $\bar{O} = \frac{\sum(c,r,t,p)}{n} = \frac{1+1+0.75+0.5}{4}$ = 0.81 | Mean $\bar{O} = \frac{\sum(c,r,t,p)}{n} = \frac{0.5+1+0.25+0.5}{4}$= 0.56 |
| **Risk treatment** | Encrypt the CAN bus flow.<br>Implement authentication with certificates among all ECUs.<br>Consider GPS corrector from NTRIP service providers [66]. | Mask end-users' ID using differential privacy.<br>Encrypt GPS data using zero-knowledge proofs. |

equation 4 to every defined AP, every metric from the equation is enumerated based on the background knowledge of the CAV environment as depicted in Table X for the two selected APs. Our findings show that the success of an attack depends not only on the attacker knowledge, expertise and equipment but also on the presence and reactivity of a supervising operator (who can be in or out the vehicle). Such information is conveyed through the $o$ and $l$ values which are correlated especially for attacks that can be executed on-board. Such human intervention has a direct impact on $L_F$ which can turn into high instead of medium if the same attack is conducted over an L5 rather than an L4 CAV for the case of $AP^{S_2}_{T_3}$.

### H. Risk Determination

TARA 2.0 risk determination relies on three parameters: the impact sum ($I$), the attack feasibility sum ($F$) and the experts objectivity index ($O$). The compilation of $O$ is demonstrated for the two selected APs in Table X where the experts' subjectivity is assessed through the four predefined metrics: certainty ($c$), peer review ($r$), measurable tools ($m$) and impartiality ($p$). Different $c$ and $m$ values were compiled impacting the $O$ scores of the two APs. Such variation is caused by the usage of the automated tool STRIDE for the $AP^{S_2}_{T_3}$ while $AP^{L_3}_{T_6}$ was elicited manually using LINDDUN. Similarly, higher certainty value is assigned to GPS spoofing than linkability threat as the former attack is well reported and simulated within the cybersecurity community [13]. The $r$ and $p$ values are identical for both APs where $r$ value is assessed to be 1 as multiple experts participated in the analysis while $p$ got a score of 0.5 as the involved experts consisted of OEMs, whose expertise may impact the provided opinion.

The risk is determined in TARA 2.0 using a 3-D plot (Figure 5). The graph compiles the rates for all APs derived from T.3 and T.6. Furthermore, it reveals, with high confidence, that $AP^{S_3}_{T_3}$ (for its node's height) and $AP^{S_2}_{T_3}$ (for its $F$ proximity to 0) are the most critical attacks requiring mitigation efforts.

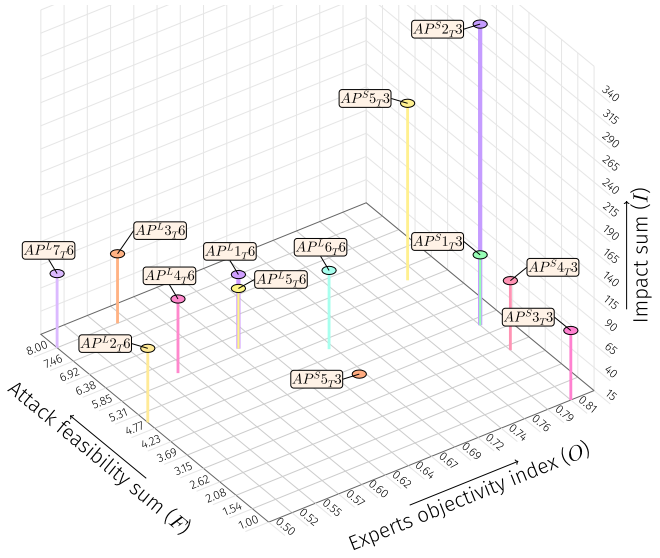



Fig. 5. 3D risk visualisation from TARA 2.0.

### I. Risk Treatment Decision

Based on LINDDUN supporting documentation mapping privacy threats to PET, and according to the UNECE appendix, we advocate robust encryption and redundancy for the GPS data flows while anonymisation solutions are recommended for data flows incorporating end-users IDs.

## VI. DISCUSSIONS AND FUTURE WORK

This section summarises the findings. To assess the validity of this work, the research hypothesis was verified through a detailed analysis of each RQ and by comparing TARA 1.0 and

TARA 2.0 performance. It also acknowledges the study's limitations and outlines directions for future research.

### A. Research Questions Analysis

*RQ1- Is it feasible to extend TARA methodology for L4 and L5 CAVs while improving the focus on privacy threats?* Capturing privacy threats in parallel to cybersecurity issues was the prime motivation of the present research which prompted the formulation of RQ1. The thorough assessment of privacy threats was successfully demonstrated from three perspectives: (i) the extension of the CIA model to encapsulate further privacy goals as per unlinkability, accountability and compliance; (ii) the combination of STRIDE and LINDDUN threat modelling; and (iii) the adjustment of the privacy impact rate calculation to consider PII processing as well as the implemented PETs on the system. This grants the possibility to model privacy threats equally to cybersecurity ones and acknowledge their coexistence within CAVs ecosystem.

*RQ2- How the assessed risks from TARA can depict scrupulously the CAV's SAE automation level?* Our work considers the human presence in controlling the risk. Integrating the SAE level into the attack feasibility rating refines, the risk assessment to focus on automated driving components, allowing decision makers to set appropriate mitigation and develop cost analysis for shifting from an in-vehicle to remote and from an L4 to L5 operations.

*RQ3- To what extend the TARA process can be automated to reduce its reliance on experts opinion?* Our work evaluates the TARA steps requiring experts knowledge, existing tool-assisted solutions and the standardised measurements. Our findings concluded that a full automated TARA is still lagging behind with the consideration of the required experts involvement. Consequently, TARA 2.0 proposes the experts objectivity index, which determines the experts subjectivity and confidence, enabling an efficient risk prioritisation and guiding auditor to address risks with higher confidence.

### B. TARA 1 0 and TARA 2 0 Qualitative Comparison

A comparison of TARA 1.0 and TARA 2.0's performances can be summarised as follows:

*1) Privacy Modelling:* TARA 2.0 models threats using privacy and security goals. It evaluates threats over eleven threat classes compared to TARA 1.0's six classes from STRIDE. Furthermore, the privacy impact is underestimated in TARA 1.0 with higher weight on safety and financial impacts than privacy and operations. Unlike TARA 1.0, TARA 2.0 evokes a weighted impact formula (Equation 3) associating a weight to every impact depending on the assessed context.

*2) SAE level:* TARA 1.0 omits the human controllability factor and applies broadly to all automotive assets, while TARA 2.0 is designed specifically for L4 and L5 CAVs.

*3) Metrics Aggregations:* TARA 2.0 provides a granular assessment of impact and the attack feasibility per AP, while TARA 1.0 computes the impact rating at the DS level. Albeit, the attack feasibility is assessed by AP, it is aggregated afterwards using upper bound analysis based on the maximum ratings to assign a value per DS. This is valid also for the risk treatment step where TARA 2.0 recommends mitigation strategies for each AP, instead of DS, offering a more fine-grained countermeasures.

*4) Risk Compilation:* TARA 1.0 combines the impact and attack feasibility levels in a 2-D matrix to provide a risk value, whereas TARA 2.0 uses a 3-D visualisation, gathering the impact sum, attack feasibility sum and experts objectivity index, facilitating the risk prioritisation.

To that end, TARA 2.0 not only aligns with the core principles of ISO/SAE 21434 but also ensures a more comprehensive and granular evaluation of cybersecurity and privacy risks specific to L4 and L5 CAVs. However, it is essential to acknowledge that both frameworks remain dependent on knowledge from OEMs, cybersecurity and privacy experts. Human insights remain crucial in any TARA where collaborative efforts are required to properly: (i) define the level of abstraction at the asset identification stage; (ii) map DSs to threat scenarios and then to threat classes; (iii) set attack trees and attack paths; (iv) rate impact factors, values and weights as well as feasibility metrics; and (v) analyse risk priorities with recommendations proposal. This limitation require further exploration in future work, as delineated in the next subsection.

### C. Experimental Validation

In the context of the ULTIMO project, experimental validation was implemented by comparing the execution of the TARA 1.0 over the real L4 vehicle [51] and application of TARA 2.0 PoC over generalised reference architecture accustoming the L4+ vehicles for passengers and goods. Several CAVs' stakeholders contributed with their expertise stemming from various backgrounds: OEMs, and Public Transport Operators (PTOs) all targeting the deployment of L4+ vehicles as ashort-term strategy. The released dataset (available at https://isec.unige.ch/) highlights risks specific to SAE L4 and L5, which were previously neglected by TARA 1.0. Furthermore, by merging STRIDE and LINDDUN methodologies, TARA 2.0 identifies a broader set of threats, leading to a more comprehensive risk model. Finally, stakeholders confirmed that 3D visualisation of risks simplifies the process of the risks prioritisation, allowing risk analysts to focus on the most relevant risks specific to L4 and L5 CAVs.

### D. Further Limitations and Future Work

Future efforts will focus on automating TARA 2.0 sub-processes. First, a tool-assisted LINDDUN would reduce the experts involvement at the threat modelling phase. Thus, it is planned to explore automated privacy threat modelling as initiated by other researchers [61]. Second, further automation of the threat modelling phase is planned by integrating Machine Learning (ML) and Deep Learning (DL)-based Intrusion Detection System (IDS) [67] for real-time asset monitoring. The aim here is to link dynamic inputs from CAVs to the threat scenario identification step. Another technological limitations identified is regarding the STRIDE analysis is the automotive template [62]. The threat model from TMT7 inherits vehicular threats and does not fully incorporate autonomous driving features. Thus, it is necessary to customise the template stencils. Therefore, it is intended to develop a dedicated autonomous driving template for

TABLE XI
STRIDE'S SECURITY & LINDDUN'S PRIVACY THREAT CATEGORIES

| ID | Class | Compromised property | Designation |
|---|---|---|---|
| **S** | Spoofing | Authenticity (Confidentiality) | Impersonating an entity to interact with a system. |
| **T** | Tampering | Integrity | Unauthorised data or functions modification. |
| **R** | Repudiation | Accountability | Not being able to trace back the author of a performed action. |
| **I** | Information Disclosure | Confidentiality | Exposing confidential data. |
| **D** | Denial of Service | Availability | Degrading service and making it unavailable to legitimate users. |
| **E** | Elevation of Privilege | Authorisation (Confidentiality) | Conducting unauthorised actions. |
| **L** | Linkability | Unlikability | Inferring items of interest about data subjects from protected data. |
| **I'** | Identifiability | Unlikability | Identifying data subjects identity. |
| **N** | Non-repudiation | Accountability | Being able to trace claimed events as well as their action owner. |
| **D'** | Detectability | Accountability | being able to detect the existence of an item of interest relted to a data subject. |
| **D** | Disclosure of Information | Confidentiality | Exposing confidential data. |
| **U** | Unawareness | Confidentiality | Not being aware about the consequence of sharing their own sensitive data. |
| **N** | Non-Compliance | Compliance | Not complying with data protection legislations or the required users' consents. |

TMT7 and integrate optimisation functions for synthesised reporting. Lastly, efforts will be considered to generalise the TARA 2.0 from CAV-specific environments to other CPS [39]. This generalisation can be accomplished by considering the SAE Lx as a domain specific parameter. For instance, in agricultural, the domain specifity parameter can be a crop-related factor, while in healthcare, it may represent a medical parameter.

## VII. CONCLUSION

A review of the state-of-the-art in the context of TARA within the automotive domain reveals gaps in terms of privacy, the SAE automation level, and expert objectivity, which are crucial dimensions often excluded or underestimated from TARA processes. The current work addresses the identified concerns and gaps by proposing an enhanced TARA 2.0.

The main contributions of this work are threefold: (i) propose five enhancements avenues, through TARA 2.0, to make the traditional TARA more privacy-centric and to address L4 and L5 CAVs' specific properties; (ii) provide guidelines in a step-by-step manner to conduct a TARA for L4 and L5 CAVs; and (iii) demonstrate a granular TARA per AP rather than conducting high level analysis at the level of DSs. The findings show that TARA 2.0 captures additional privacy threat classes, assesses fine-grained privacy impact and incorporates the SAE level as a metric influencing the attack likelihood. Additionally, with the consideration of the experts' objectivity index, TARA 2.0 pushes towards reliable risk analysis supporting risk decision makers and CAV's stakeholders in determining appropriate cybersecurity goals and claims. The applicability of TARA 2.0 was demonstrated through a concrete PoC performed in the context of the ULTIMO project in close collaboration with CAVs stakeholders. Moreover, a thorough comparative analysis between TARA 1.0 and TARA 2.0 was performed, followed by an outline of future efforts envisioned as part of ongoing research.

Despite its strong performance, TARA 2.0 has limitations that future work will address. While it transparently reflects expert involvement, it still relies on them for several steps. Future efforts aim to automate these expert-dependent processes, including the planned automation of LINDDUN and the AI-based threat modelling. The ultimate goal is to extend assessment inputs to be more granular and to generalise TARA 2.0 for application across all CPS and IoT systems.

TABLE XII
LIST OF NOTATIONS USED THROUGHOUT THE PAPER

| Symbol | Definition |
|---|---|
| $A.\#$ | Asset ID starting from 1 |
| $AP_{T_6}^{L_7}$ | Superscript indicates the threat class combined with the AP number. Subscript shows the threat scenario ID. |
| $c$ | Certainty parameter for the objectivity index |
| $D.\#$ | Damage scenario ID starting from 1 |
| $e$ | Specialised expertise |
| $F$ | Attack feasibility sum |
| $I$ | Impact sum |
| $i_f$ | Financial parameter for the impact sum |
| $i_o$ | Operational parameter for the impact sum |
| $i_p$ | Privacy parameter for the impact sum |
| $i_s$ | Safety parameter for the impact sum |
| $k$ | Knowledge of item or component |
| $l$ | Automation level (SAE Lx ($l$)) |
| $L_I$ | Impact level |
| $L_F$ | Feasibility level |
| $m$ | Measurable tools parameter for the objectivity index |
| $n$ | Total number of factors included in the objectivity index |
| $o$ | Windows of opportunity |
| $O$ | Objectivity index |
| $p$ | Impartiality parameter for the objectivity index |
| $q$ | Equipment |
| $r$ | Peer review parameter for the objectivity index |
| $t$ | Elapsed time |
| $T.\#$ | Threat scenario ID starting from 1 |
| $w_f$ | Financial weight for the financial impact |
| $w_o$ | Operational weight for the operational impact |
| $w_p$ | Privacy weight for the privacy impact |
| $w_s$ | Safety weight for the safety impact |

## APPENDIX A
## STRIDE & LINDDUN THREAT CLASSES

TableXI presents the STRIDE and LINDDUN threat classes used in Section V-D. The threat classes were determined based on ISO/IEC 27000 [28], Microsoft Threat Modelling tool documentation [55] and LINDDUN organisation [63]. Additional supplementary materials is available at https://isec.unige.ch/.

## APPENDIX B
## LIST OF NOTATIONS

Table XII presents the notations used in the paper.

## ACKNOWLEDGMENT

ULTIMO project https://doi.org/10.3030/101077587. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them. This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI).

## REFERENCES

[1] B. R. Kiran et al., "Deep reinforcement learning for autonomous driving: A survey," *IEEE Trans. Intell. Transport. Syst.*, vol. 23, no. 6, pp. 4909–4926, Jun. 2022.

[2] R. Roriz, J. Cabral, and T. Gomes, "Automotive LiDAR technology: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6282–6297, Jul. 2022.

[3] C. Creß, Z. Bing, and A. C. Knoll, "Intelligent transportation systems using roadside infrastructure: A literature survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 7, pp. 6309–6327, Jul. 2024.

[4] W. M. D. Chia, S. L. Keoh, C. Goh, and C. Johnson, "Risk assessment methodologies for autonomous driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 16923–16939, Oct. 2022.

[5] T. Gong, L. Zhu, F. R. Yu, and T. Tang, "Edge intelligence in intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 9, pp. 8919–8944, Sep. 2023.

[6] S. Shirvani, Y. Baseri, and A. Ghorbani, "Evaluation framework for electric vehicle security risk assessment," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 1, pp. 33–56, Jan. 2024.

[7] X. Sun, F. R. Yu, and P. Zhang, "A survey on cyber-security of connected and autonomous vehicles (CAVs)," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6240–6259, Jul. 2022.

[8] S. Z. Khan, M. Mohsin, and W. Iqbal, "On GPS spoofing of aerial platforms: A review of threats, challenges, methodologies, and future research directions," *PeerJ Comput. Sci.*, vol. 7, p. e507, May 2021.

[9] M. Girdhar, J. Hong, and J. Moore, "Cybersecurity of autonomous vehicles: A systematic literature review of adversarial attacks and defense models," *IEEE Open J. Veh. Technol.*, vol. 4, pp. 417–437, 2023.

[10] M. Benyahya, A. Collen, S. Kechagia, and N. A. Nijdam, "Automated city shuttles: Mapping the key challenges in cybersecurity, privacy and standards to future developments," *Comput. Secur.*, vol. 122, Nov. 2022, Art. no. 102904.

[11] B. Chah, A. Lombard, A. Bkakria, R. Yaich, A. Abbas-Turki, and S. Galland, "Privacy threat analysis for connected and autonomous vehicles," *Proc. Comput. Sci.*, vol. 210, pp. 36–44, Jan. 2022.

[12] N. Azam, L. Michala, S. Ansari, and N. B. Truong, "Data privacy threat modelling for autonomous systems: A survey from the GDPR's perspective," *IEEE Trans. Big Data*, vol. 9, no. 2, pp. 388–414, Apr. 2023.

[13] *R155*, UNECE, Geneva, Switzerland, 2020.

[14] *R156*, UNECE, Geneva, Switzerland, Apr. 2020.

[15] *ISO/SAE 21434 Road Vehicles-Cybersecurity Engineering*, ISO, Geneva, Switzerland, 2021.

[16] European Commission. *E-Safety Vehicle Intrusion ProTected Applications EVITA Project*. May 1, 2024. [Online]. Available: https://cordis.europa.eu/project/id/224275

[17] M. M. Islam, A. Lautenbach, C. Sandberg, and T. Olovsson, "A risk assessment framework for automotive embedded systems," in *Proc. 2nd ACM Int. Workshop Cyber-Phys. Syst. Secur.*, New York, NY, USA, May 2016, pp. 3–14.

[18] *ETSI TS 102 165 Method and Pro Forma for Threat, Vulnerability, Risk Analysis (TVRA)*, ETSI, Sophia Antipolis, France, 2017.

[19] J.-P. Monteuuis, A. Boudguiga, J. Zhang, H. Labiod, A. Servel, and P. Urien, "SARA: Security automotive risk analysis method," in *Proc. 4th ACM Workshop Cyber-Phys. Syst. Secur.*, New York, NY, USA, May 2018, pp. 3–14.

[20] S. Park and H. Park, "PIER: Cyber-resilient risk assessment model for connected and autonomous vehicles," *Wireless Netw.*, vol. 30, no. 5, pp. 4591–4605, Aug. 2022.

[21] M. Benyahya, T. Lenard, A. Collen, and N. A. Nijdam, "A systematic review of threat analysis and risk assessment methodologies for connected and automated vehicles," in *Proc. 18th Int. Conf. Availability, Rel. Secur.*, New York, NY, USA, Aug. 2023, pp. 1–10.

[22] A. Abdo, H. Chen, X. Zhao, G. Wu, and Y. Feng, "Cybersecurity on connected and automated transportation systems: A survey," *IEEE Trans. Intell. Vehicles*, vol. 9, no. 1, pp. 1382–1401, Jan. 2024.

[23] *SAE J3016 Taxonomy and Definitions for Terms Related to Driving Automation Systems*, SAE Int., Warrendale, PA, USA, 2021.

[24] C. Nobles, D. N. Burrell, S. L. Burton, and T. Waller, "Driving into cybersecurity trouble with autonomous vehicles," in *Handbook of Research on Cybersecurity Risk in Contemporary Business Systems*. Hershey, PA, USA: IGI Global, 2023, pp. 255–273.

[25] F. Luo et al., "Cybersecurity testing for automotive domain: A survey," *Sensors*, vol. 22, no. 23, p. 9211, Nov. 2022.

[26] *ISO 26262 Road Vehicles—Functional Safety*, ISO, Geneva, Switzerland, 2018.

[27] D. Omeiza, H. Webb, M. Jirotka, and L. Kunze, "Explanations in autonomous driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10142–10162, Aug. 2022.

[28] *Information Security Management Systems*, ISO/IEC, Geneva, Switzerland, 2018.

[29] D. Van Landuyt and W. Joosen, "A descriptive study of assumptions made in LINDDUN privacy threat elicitation," in *Proc. 35th Annu. ACM Symp. Appl. Comput.*, Mar. 2020, pp. 1280–1287.

[30] A. Lautenbach, M. Almgren, and T. Olovsson, "Proposing HEAVENS 2.0—An automotive risk assessment model," in *Proc. Comput. Sci. Cars Symp.* New York, NY, USA: ACM, Nov. 2021, pp. 1–12.

[31] S. Khastgir, S. Birrell, G. Dhadyalla, H. Sivencrona, and P. Jennings, "Towards increased reliability by objectification of hazard analysis and risk assessment (HARA) of automated automotive systems," *Saf. Sci.*, vol. 99, pp. 166–177, Nov. 2017.

[32] A. Bolovinou, U.-I. Atmaca, A. T. Sheik, O. Ur-Rehman, G. Wallraf, and A. Amditis, "TARA+: Controllability-aware threat analysis and risk assessment for L3 automated driving systems," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 8–13.

[33] Q. He, X. Meng, and R. Qu, "Towards a severity assessment method for potential cyber attacks to connected and autonomous vehicles," *J. Adv. Transp.*, vol. 2020, pp. 1–15, Sep. 2020.

[34] T.-C. Wen, H.-Y. Chung, K.-H. Chang, and Z.-S. Li, "A flexible risk assessment approach integrating subjective and objective weights under uncertainty," *Eng. Appl. Artif. Intell.*, vol. 103, Aug. 2021, Art. no. 104310.

[35] V. Agrawal, B. Achuthan, A. Ansari, V. Tiwari, and V. Pandey, "Threat/hazard analysis and risk assessment: A framework to align the functional safety and security process in automotive domain," *SAE Int. J. Transp. Cyber. Privacy*, vol. 4, pp. 83–96, Dec. 2021.

[36] T. Vogt et al., "A comprehensive risk management approach to information security in ITS," *SAE Int. J. Transp. Cybersecurity Privacy*, vol. 4, no. 1, pp. 4–11, 2021.

[37] J. Dobaj, G. Macher, D. Ekert, A. Riel, and R. Messnarz, "Towards a security-driven automotive development lifecycle," *J. Softw., Evol. Process*, vol. 35, no. 8, pp. 1–22, Aug. 2023.

[38] C. Schmittner, B. Schrammel, and S. König, "Asset driven ISO/SAE 21434 compliant automotive cybersecurity analysis with ThreatGet," in *Proc. Eur. Conf. Softw. Process Improvement*. Cham, Switzerland: Springer, Jan. 2021, pp. 548–563.

[39] C. Plappert, D. Zelle, H. Gadacz, R. Rieke, D. Scheuermann, and C. Krauß, "Attack surface assessment for cybersecurity engineering in the automotive domain," in *Proc. 29th Euromicro Int. Conf. Parallel, Distrib. Netw.-Based Process. (PDP)*, Mar. 2021, pp. 266–275.

[40] M. Ebrahimi, C. Striessnig, J. C. Triginer, and C. Schmittner, "Identification and verification of attack-tree threat models in connected vehicles," *SAE Threat Model.*, SAE Technical Paper 2022-01-7087, 2022, doi: 10.4271/2022-01-7087.

[41] D. Zelle, C. Plappert, R. Rieke, D. Scheuermann, and C. Krauß, "ThreatSurf: A method for automated threat surface assessment in automotive cybersecurity engineering," *Microprocessors Microsyst.*, vol. 90, Apr. 2022, Art. no. 104461.

[42] S. Ghosh, A. Zaboli, J. Hong, and J. Kwon, "An integrated approach of threat analysis for autonomous vehicles perception system," *IEEE Access*, vol. 11, pp. 14752–14777, 2023.

[43] Z. Abuabed, A. Alsadeh, and A. Taweel, "STRIDE threat model-based framework for assessing the vulnerabilities of modern vehicles," *Comput. Secur.*, vol. 133, Oct. 2023, Art. no. 103391.

[44] I. Loskin, "TARA+AD threat analysis and risk assessment for automated driving," Ph.D. dissertation, Fac. Inf. Technol., Univ. Jyvaskyla, Jyvaskyla, Finland, 2023.
[45] D. Dominic, S. Chhawri, R. M. Eustice, D. Ma, and A. Weimerskirch, "Risk assessment for cooperative automated driving," in *Proc. 2nd ACM Workshop Cyber-Phys. Syst. Secur. Privacy*, New York, NY, USA, Oct. 2016, pp. 47–58.
[46] A. O'Hagan, "Expert knowledge elicitation: Subjective but scientific," *Amer. Statistician*, vol. 73, no. 1, pp. 69–81, Mar. 2019.
[47] K.-H. Chang, "Integrating subjective–objective weights consideration and a combined compromise solution method for handling supplier selection issues," *Systems*, vol. 11, no. 2, p. 74, Feb. 2023.
[48] G. Xie, W. Wu, G. Zeng, R. Li, and S. Hu, "Risk assessment and development cost optimization in software defined vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3675–3686, Jun. 2021.
[49] N. Moselhy and A. A. Mahmoud, "Standardization of cybersecurity concepts in automotive process models: An assessment tool proposal," in *Advances in Information and Communication*. Cham, Switzerland: Springer, 2023, pp. 635–655.
[50] L. O. Mailloux, M. Span, R. F. Mills, and W. Young, "A top down approach for eliciting systems security requirements for a notional autonomous space system," in *Proc. IEEE Int. Syst. Conf. (SysCon)*, Apr. 2019, pp. 1–7.
[51] M. Benyahya, P. Bergerat, A. Collen, and N. A. Nijdam, "Symbiotic analysis of security assessment and penetration tests guiding real L4 automated city shuttles," *Telecom*, vol. 4, no. 1, pp. 198–218, Mar. 2023.
[52] *Regulation (EU) 2016/679 of 27 April 2016 on the Protection of Natural Persons With Regard to the Processing of Personal Data and on the Free Movement of Such Data*, Eur. Commission, Brussels, Belgium, 2016.
[53] *Data Protection Engineering*, Eur. Union Agency for Cybersecurity, Athens, Greece, 2022.
[54] C. Treacy, J. Loane, and F. McCaffery, "A developer driven framework for security and privacy in the Internet of Medical Things," in *Communications in Computer and Information Science*, vol. 1251. Cham, Switzerland: Springer, 2020, pp. 107–119.
[55] Microsoft. (2023). *Microsoft Threat Modeling Tool*. [Online]. Available: https://learn.microsoft.com/en-us/azure/security/develop/threat-modeling-tool-threats#stride-model
[56] *Security Techniques—Privacy Framework*, Standard ISO/IEC 29100:2011, 2011.
[57] M. Benyahya, S. Kechagia, A. Collen, and N. A. Nijdam, "The interface of privacy and data security in automated city shuttles: The GDPR analysis," *Appl. Sci.*, vol. 12, no. 9, p. 4413, Apr. 2022.
[58] *Information Security, Cybersecurity and Privacy Protection—Evaluation Criteria for IT Security—Methodology for IT Security Evaluation*, Standard ISO/IEC 18045, 2022.
[59] *Conformity Assessment—Requirements for Bodies Certifying Products, Processes and Services*, Standard ISO/IEC 17065, 2012.
[60] *ULTIMO Advancing Sustainable User-Centric Mobility With Automated Vehicles*, Eur. Commission, Brussels, Belgium, 2023. [Online]. Available: https://cordis.europa.eu/project/id/101077587
[61] L. Sion, "Automated threat analysis for security and privacy," Ph.D. dissertation, Fac. Eng. Sci., KU Leuven, Leuven, Belgium, 2020.
[62] Matt Lewis. (2016). *The Automotive Threat Modeling Template*. [Online]. Available: https://research.nccgroup.com/2016/07/20/the-automotive-threat-modeling-template/
[63] K. Wuyts, L. Sion and W. Joosen, "LINDDUN GO: A lightweight approach to privacy threat modeling," *IEEE Eur. Symp. on Secur. Privacy Workshops (EuroS&PW)*, Genoa, Italy, 2020, pp. 302–309, doi: 10.1109/EuroSPW51379.2020.00047.
[64] A.-M. Konsta, A. Lluch Lafuente, B. Spiga, and N. Dragoni, "Survey: Automatic generation of attack trees and attack graphs," *Comput. Secur.*, vol. 137, Feb. 2024, Art. no. 103602.
[65] NIST and U.S. Department of Commerce. *National Vulnerability Database*. Mar. 3, 2024. [Online]. Available: https://nvd.nist.gov/vuln/full-listing
[66] S. Park, S. Ryu, J. Lim, and Y.-S. Lee, "A real-time high-speed autonomous driving based on a low-cost RTK-GPS," *J. Real-Time Image Process.*, vol. 18, no. 4, pp. 1321–1330, Aug. 2021.
[67] K. Geeta and K. Gulshan, "Machine learning and deep learning methods for intrusion detection systems: Recent developments and challenges," *Soft Comput.*, vol. 25, no. 15, pp. 9731–9763, 2021.

**Meriem Benyahya** (Member, IEEE) received the Ph.D. degree in information systems from the University of Geneva. She is a Post-Doctoral Researcher at the University of Geneva. Currently, she works on cybersecurity and data privacy implications and risk assessment tasks on several Horizon Europe projects, such as ULTIMO. She successfully managed IT projects from different fields, including PCI-DSS certifications, business continuity plans, and disaster recovery.

**Anastasija Collen** (Member, IEEE) received the Ph.D. degree in information systems from the University of Geneva. She is a Senior Researcher and a Lecturer at the University of Geneva. She is an Experienced Research and Development Engineer in the fields of privacy and security. She contributes to multiple EU-funded projects, including ULTIMO, OPEVA, and AutoTRUST. Her current interests lie in the automotive cybersecurity with the focus on human factors.

**Teri Lenard** received the joint Ph.D. degree in information technology from the University of Geneva and the "G. E. Palade" University of Medicine, Pharmacy, Science and Technology of Târgu Mureş. His research interests include security protocols in automotive and IoT systems, applications of trusted computing, and trust modelling and management.

**Niels Alexander Nijdam** received the Ph.D. degree in computer science from the University of Geneva, where his topics included collaborative systems, ubiqutous computing and rendering and programmable graphics. He is a Computer Scientist and a Senior Lecturer at the University of Geneva and is leading the Information Security Laboratory (I-Sec). He has been coordinating the scientific efforts in several EU-funded projects, including ULTIMO, OPEVA, AutoTRUST, and AI4SWEng.