CHALMERS
UNIVERSITY OF TECHNOLOGY

A Systematic Literature Review on Automotive Digital Forensics: Challenges, Technical Solutions and Data Collection

Citation for the original published paper (version of record):
Strandberg, K., Nowdehi, N., Olovsson, T. (2023). A Systematic Literature Review on Automotive Digital Forensics: Challenges, Technical Solutions
and Data Collection. IEEE Transactions on Intelligent Vehicles, 8(2): 1350-1367.
http://dx.doi.org/10.1109/TIV.2022.3188340

(article starts on next page)

A Systematic Literature Review on Automotive Digital Forensics: Challenges, Technical Solutions and Data Collection

Kim Strandberg, *Member, IEEE*, Nasser Nowdehi, and Tomas Olovsson, *Member, IEEE*

***Abstract*—A modern vehicle has a complex internal architecture and is wirelessly connected to the Internet, other vehicles, and the infrastructure. The risk of cyber attacks and other criminal incidents along with recent road accidents caused by autonomous vehicles calls for more research on automotive digital forensics. Failures in automated driving functions can be caused by hardware and software failures and cyber security issues. Thus, it is imperative to be able to determine and investigate the cause of these failures, something which requires trustable data. However, automotive digital forensics is a relatively new field for the automotive where most existing self-monitoring and diagnostic systems in vehicles only monitor safety-related events. To the best of our knowledge, our work is the first systematic literature review on the current research within this field. We identify and assess over 300 papers published between 2006–2021 and further map the relevant papers to different categories based on identified focus areas to give a comprehensive overview of the forensics field and the related research activities. Moreover, we identify forensically relevant data from the literature, link the data to categories, and further map them to required security properties and potential stakeholders. Our categorization makes it easy for practitioners and researchers to quickly find relevant work within a particular sub-field of digital forensics. We believe our contributions can guide digital forensic investigations in automotive and similar areas, such as cyber-physical systems and smart cities, facilitate further research, and serve as a guideline for engineers implementing forensics mechanisms.**

***Index Terms*—Automotive forensics, car forensics, cyber attacks, cyber security, forensic investigations, forensics guidelines, forensics mechanisms, forensic solutions, in-vehicle network, V2X communication, vehicle architecture, vehicle forensics.**

I. Introduction

THE complexity of vehicles is increasing at a high pace. A modern vehicle can contain more than 150 ECUs (Electronic Control Units) and over 100 M lines of code. Moreover, current vehicles have various connection interfaces and a large amount of forensically interesting data exchange between a multitude of entities such as sensors, actuators, ECUs, the Internet, and infrastructure. To enable a proper forensics investigation, data must be collected, stored, and processed in a forensically sound and secure manner.

In the remainder of this section, we explain the complexity of the vehicle architecture, its relationship to other similar areas, and the field of automotive digital forensics. We define the goal with our paper, problem, approach, and our main contributions.

A. The Interconnected Vehicle

A vehicle uses various *hardware*, *software* and *storage* components, and different *communication* technologies.

Hardware can be broken down into ECUs, sensors, and actuators. The complexity of an ECU varies depending on its task, which can range from simple processing of sensor signals to an infotainment system with a multitude of applications. *Sensors* can give information about speed, temperature, distance, and identification of obstacles (e.g., pedestrians and animals). The *actuators* turn input from these *sensors* (via an ECU) into actions, such as braking, steering, and engine control. *Software* can be either *in transit*, *at rest*, or *running*, e.g., software provisioning systems, such as over-the-air or workshop updates, *transmits* software, and software can be installed (*at rest*) or be *running* in ECUs. *Data storage* includes storage of, e.g., cryptographic keys, forensics logs, system information, and reports about the vehicle and the driver.

Fig. 1 shows two examples of In-Vehicle Networks (IVNs) with various nodes belonging to different bus technologies, where the basic topology in the top currently is most common. Transmission can occur over, e.g., CAN, FlexRay, MOST, LIN and Ethernet. A primary gateway connects to sub-gateways responsible for translating and relaying traffic to the correct network segment. The vehicle is connected to the outside world via various connection interfaces giving rise to Vehicle-to-Everything (V2X) communication. Wireless connections occur via, e.g., 3 G/4 G/5 G, WiFi, and Bluetooth, and physical connections via, e.g., OBD-II, USB, and debug-ports. The communication is extensive considering the amount of data generated in the vehicle and the increasing

Manuscript received 28 April 2022; revised 15 June 2022; accepted 29 June 2022. Date of publication 4 July 2022; date of current version 20 March 2023. This work was supported by VINNOVA, the Swedish Governmental Agency for Innovation Systems through the CyReV project under Grant 2019-03071. *(Corresponding author: Kim Strandberg.)*

Kim Strandberg is with the Department of Research and Development, Volvo Cars, 405 31 Gothenburg, Sweden, and also with the Department of Computer Science and Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: kim.strandberg@volvocars.com).

Nasser Nowdehi is with the Volvo AB, 417 15 Gothenburg, Sweden (e-mail: nasser.nowdehi@volvo.com).

Tomas Olovsson is with the Department of Computer Science and Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: tomas.olovsson@chalmers.se).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIV.2022.3188340.

Digital Object Identifier 10.1109/TIV.2022.3188340

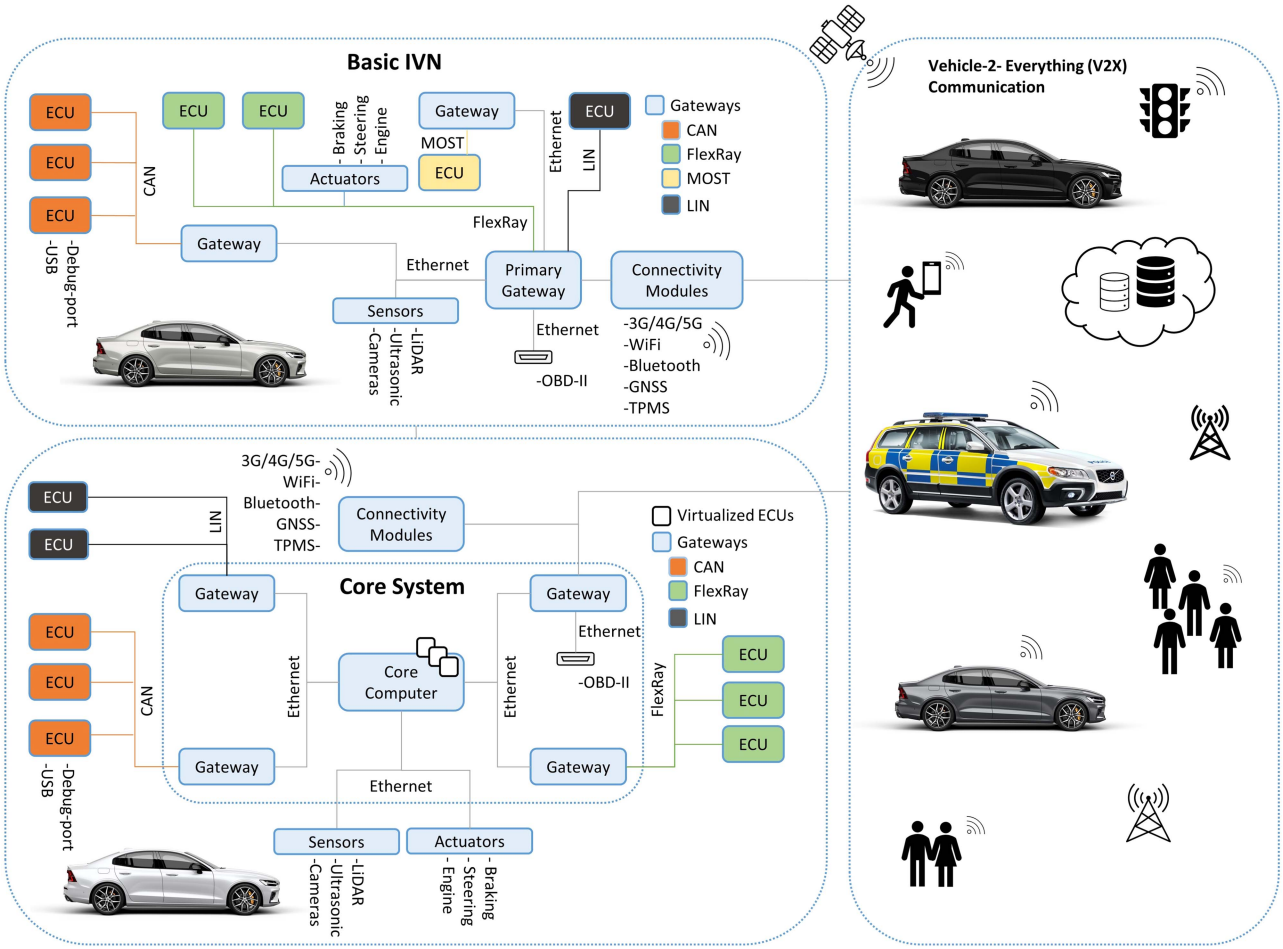

Fig. 1. The In-Vehicle Network and V2X Communication.

communication with the outside world, such as with other vehicles, roadside units (RSUs), and with cloud-based services.

The other example in the lower part of Fig. 1, shows that there is a trend in the automotive industry to move towards a more centralized architecture with fewer yet more powerful ECUs due to increased requirements for system performance in self-driving vehicles, where high-performance computers can virtualize hardware [1]. Automotive Ethernet communication becomes more prevalent due to its higher bandwidth and more straightforward adaptation to security features (e.g., authentication and encryption of messages) compared to other automotive communication technologies. Thus, the potential for better use of technical solutions for automotive forensics come to light.

ECUs have different functionalities and responsibilities, e.g., braking, steering, and engine control are part of the safety critical system. Advanced driver assistance systems (ADAS) technology handles automatic emergency braking, lane assistance, and driver monitoring.

B. Related Areas

Internet of Things (IoT) relates to the automotive domain because of the increasing use of IoT devices in vehicles [2]. IoT refers to all devices with Internet capabilities such as cellular phones, medical devices, home appliances (e.g., TVs, washing machines, printers, etc.), including sensors and actuators. The concept of the Internet of Vehicles (IoV) strives to connect Vehicular Adhoc Networks (VANETs) to the IoT domain [3]. However, vehicles are safety-critical systems, and failures in any part of a safety-critical system have the risk of severe lethal outcomes for, e.g., drivers, passengers, and everyone around it.

For comparison, both IoT and the automotive domain suffer from previous designs not developed with security and forensics in mind. Instead, emphasis has been on cost and usability. However, safety has been imperative within automotive, but related security properties are often deficient. Moreover, there are no standardized processes for the forensic process within IoT or automotive. In IoT, data can be found on devices, networks, or the cloud [4]. The same applies to automotive. However, within automotive forensics, the main emphasis is on vehicles as safety-critical systems and the extensive communication during operation (cf. Table III). Cooperative Intelligent Transport Systems (C-ITS) development has facilitated V2X communication and prepared to integrate vehicles into an upcoming smart infrastructure. ITS can refer to services for all modes of transport but relates mainly to transport on roads [5]. Future vehicles will become more autonomous and part of a much larger system with various interactions between a virtual and

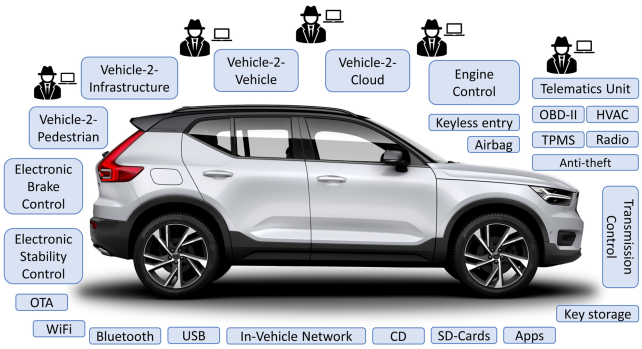

Fig. 2. Example of potential attack vectors.

physical world. Services within ITS allow vehicles to share information, such as traffic conditions. Vehicles braking, speed, and position can be communicated, travel routes optimized, and traffic congestion limited. Virtual entities can replace physical entities, e.g., physical traffic lights and road signs can instead communicate traffic information wirelessly. Although there are similarities and interactions with other areas, the automotive domain is still very distinct and has its specific challenges. From an automotive digital forensic perspective, data communicated within the ITS is imperative to securely log and store to enable post-incident investigations.

C. Automotive Digital Forensics

Digital forensic investigations include the *identification*, *preservation*, *acquisition*, *verification*, *analysis*, and *reporting* of data [6]. The definition of digital forensics has expanded from involving just computers to include all digital devices that can store, process, or transmit data [7], a shift that leads to increased complexity. For example, when it comes to vehicles, file formats and operating systems differ in ECUs making it challenging to create unified standards and tools; and vehicle IVNs consist of many interconnected devices communicating using different communication protocols. Fig. 2 shows some examples of potential entry points that can be of interest to cyber criminals.

It is well known that increased complexity increases the risk of vulnerabilities and, thus, potential attack vectors [8]. Moreover, increased connectivity broadens the attack surface with a higher potential to expose vulnerabilities. Since a modern vehicle can contain over 100 M lines of code, aligned with a rough estimation of at least one bug per 1000 lines of code, it indicates more than 100 k bugs in a modern vehicle. Moreover, in [9], T. Llanso and M. McNeil estimate that at least 1% of software vulnerabilities can be exploited, further indicating around 1 k potential ways to compromise the vehicle software. Thus, due to technological advancements such as increased connectivity and the introduction of autonomous driving, incidents that require digital forensic investigations will inevitably rise to become more prevalent in the future.

Incidents can be intentionally caused by targeted cyber attacks and non-intentionally due to, e.g., a distracted driver or a software or hardware failure. Moreover, as shown in [10]–[13], cyber attacks can be associated with life-threatening hazards due to their potential to affect safety-critical systems such as braking, steering, and engine control. For instance, in 2015, Charlie Miller and Chris Valasek hacked a Jeep Cherokee remotely over the Internet [10]. Although the safety-critical systems were isolated, a control unit had access to the communication bus and was vulnerable to reprogramming. Thus, the hackers managed to add code to the control unit and use it to send arbitrary CAN signals over the Internet to control, e.g., brakes and steering. In [12], Karl Koscher et al. demonstrated the potential to extract and reverse-engineer firmware to understand hardware features, which enabled them to add new functionalities and malicious code to a telematics unit. Moreover, the malicious code automatically erased any evidence of its existence after a crash. Thus, there was no post-incident available data for an investigation.

In 2018, a woman was killed by an autonomous vehicle [14]. The software did not correctly identify the individual as a pedestrian. In [15], a driver was using the autopilot and crashed into a vehicle in front. Afterward, vehicle data, such as, information from sensors, was used in digital forensics investigations. Establishing incident traceability regarding the driver, the autopilot, and potential threat actors is imperative.

In [16], six threat actors are identified, namely, the Financial Actor (FA), the Foreign Country (FC), the Cyber Terrorist (CT), the Insider (IN), the Hacktivist (HA), and the Script Kiddie (SK). For instance, the FC, CT, and HA can use one or many vehicles as moving weapons targeting humans or buildings. The FA might install ransomware that disables the ignition until a ransom is paid. The IN might add backdoors in vehicle software, while the SK can execute others developed exploits, usually with an unclear agenda. Beyond cyber attacks, the owner might manipulate the vehicle to gain more functionality, such as chip tuning, avoiding route tracing, and changing the odometer. Cases and incidents, as previously described, are highly relevant to identify and trace in automotive digital forensic investigations.

Current regulations, such as the UN R.155, require that vehicles provide data forensic capability for analyzing attempted or successful cyber attacks [17]. However, no details are provided on how to fulfill these legal requirements, and current vehicles have limited capabilities for enabling digital forensic investigations in cybersecurity-relevant incidents.

Failures in hardware, software, and cyber security issues must be possible to detect and investigate. Thus, a forensic-enabled vehicle needs to support mechanisms that generate, store, and secure forensically relevant data and differentiate between different types of malfunction.

D. Goal

1) Problem: We believe that the lack of digital forensics guidelines and digital forensics mechanisms within the automotive industry is a valid concern. To our best knowledge, no previous Systematic Literature Review (SLR) has been done within this field, that identifies the current work, and identify, categorize and map forensically relevant information to security properties and data users. The paper aims to answer the following questions:

TABLE I
SELECTED PAPERS CONCERNING TECHNICAL SOLUTIONS

Ref.	Author	Publ. Year	Details	1a. **Data:** Data Collection	1b. **Data:** Extract. techniques	2a. **Challenges:** General	2b. **Challenges:** Req./Guidlines	3a. **Com.:** Cloud/Fog/Edge	3b. **Com.:** VANETs	4a. **Software:** In-vehicle	4b. **Software:** Tools	5a. **Hardware:** Architecture	5b. **Hardware:** Sensor	5c. **Hardware:** EDR/Blackbox	6a. **Alg.:** Machine learning	6b. **Alg.:** Other alg.	7a. **Cryptography:** Blockchain	7b. **Cryptography:** Other crypt.	8. **Framework and Processes**	9. **Practical Experiments**	10. **Infrastruct./Smart Cities**	11. **TEE/Virtualization**
[35]	C. Oham et al.	2021	Jo. I.A.N.P.									●					●	●	●	●		
[36]	C. Yoon et al.	2021	Jo. I.P.					●					●				●		●			
[37]	C. Alexakos et al.	2021	Jo. I.	●		●	●				●	●							●			
[38]	M. Waltereit et al.	2021	Co.	●									●	●					●	●		
[39]	P.A. Abhay et al.*	2021	Jo. C.I.							●			●	●			●		●			
[40]	M.A. Hoque et al.*	2021	Co. C.I.			●												●	●			
[41]	J. Daily et al.	2020	Jo. I.	●	●						●			●					●	●		
[42]	P. Sharma et al.	2020	Co.									●	●		●				●			
[43]	H. Guo et al.	2020	Jo. I.	●										●		●	●					
[44]	A. Philip et al.	2020	T.Jo. I.P.												●		●		●			
[45]	M. Li et al.	2020	Jo. C.I.A.N.P.	●			●									●	●	●	●			
[46]	Z. Ma et al.	2020	Jo. C.I.P.						●				●			●		●				
[19]	N. Vinzenz et al.	2020	Jo. C.I.A.N.P.	●			●					●		●				●				
[47]	A. Mehrish et al.	2020	Jo.							●			●		●	●						
[48]	M. Waltereit et al.	2019	Co.			●										●						
[49]	K. Bahirat et al.	2019	Co.	●			●						●		●							
[50]	L. Cintron et al.	2019	Co. I.N.P.	●												●	●		●		●	●
[51]	S. Lee et al.	2019	Jo. I.N.				●					●		●		●		●				●
[52]	L. Davi et al.	2019	Jo. I.N.									●					●		●			
[53]	D. Billard et al.	2019	Co. N.P.										●				●			●		
[18]	X. Feng et al.	2019	Co. C.I.A.P.	●		●	●											●		●	●	
[33]	X. Wang et al.	2019	Co.					●								●				●	●	
[54]	M. Ugwu et al.	2018	Jo. I.N.				●							●			●	●	●	●		
[55]	M. Marchetti et al.	2018	Jo.	●		●										●						
[56]	H. Guo et al.	2018	Co. I.N.	●										●			●	●				
[57]	R. Hussain et al.	2018	Jo. C.I.N.P.					●	●				●					●	●			
[24]	M. Cebe et al.	2018	Jo. I.N.P.	●		●	●										●	●	●			
[32]	M. Hossain et al.	2017	Co. C.I.	●		●						●				●	●		●	●		●
[58]	A. Mehrish et al.	2017	Co.										●			●						
[59]	X. Feng et al.*	2017	Co. C.I.															●	●		●	
[25]	H. Mansor et al.	2016	Co. C.I.A.P.	●			●	●		●												
[60]	A.D. Sathe et al.	2016	Co.	●						●			●									
[61]	N.Watthanawisuth et al.	2012	Co.	●									●	●						●		
[62]	D. Nilsson et al.	2008	Jo. I.N.	●			●												●			

Retrieved from Snowballing

(Co)Conference
(Jo)Journal

(C)Confidentiality
(I)Integrity
(A)Availability
(N)NonRepudiation
(P)Privacy

– *What research exists within the field of automotive digital forensics?*
– *What is the coverage and specificity in different databases for automotive forensics search queries?*
– *What technical solutions exist with regard to automotive digital evidence, and how do these solutions uphold security properties?*
– *What forensically relevant data can be derived from existing literature and who are the stakeholders for this data?*

2) Approach: We have performed an SLR over work published between 2006 - 2021 within the field of automotive digital forensics and grouped them into two core categories, namely *technical solutions* and *surveys*. Another 11 categories, some with sub-categories, were identified from the selected work based on focus areas (cf. Table I and II). We have identified gaps and discussed issues and challenges with respect to these categories. Moreover, we stated if any security properties were considered in the proposed *technical solutions*.

From the result of the SLR, we identified additional categories (cf. Table III), this time specifically concerning forensic data, which were further mapped to required security properties and potential stakeholders. The aim was to identify data to be considered for automotive digital forensic investigations and ensure that data is reliable and secured.

An SLR is a recognized and standardized approach to provide broad coverage over publications concerning a particular field of interest. By following well-established processes, we provide confidence that other practitioners and researchers in the area do not need to repeat this work for the same period of time. Still, we give enough details to enable replicating the approach for a future time span to follow the progress within the field.

3) Contributions: Our main contributions are:
– We have performed an SLR within the field of automotive digital forensics.
– We performed database searches in four of the largest databases with the aim to get broader coverage and to investigate the individual coverage and specificity for each database. Backward and forward snowballing was performed on the selected work to increase the coverage even more.

TABLE II
SELECTED PAPERS CONCERNING SURVEYS

Ref.	Author	Publ. Year	*Retrieved from Snowballing (Co)Conference (Jo)Journal	1a. **Data:** Data Collection	1b. **Data:** Extract. techniques	2a. **Challenges:** General	2b. **Challenges:** Req./Guidlines	3a. **Com.:** Cloud/Fog/Edge	3b. **Com.:** VANETs	4a. **Software:** In-vehicle	4b. **Software:** Tools	5a. **Hardware:** Architecture	5b. **Hardware:** Sensor	5c. **Hardware:** EDR/Blackbox	6a. **Alg.:** Machine learning	6b. **Alg.:** Other alg.	7a. **Cryptography:** Blockchain	7b. **Cryptography:** Other crypt.	8. **Framework and Processes**	9. **Practical Experiments**	10. **Infrastruct./Smart Cities**	11. **TEE/Virtualization**
[26]	K. Buquerin et al.*	2021	Jo.	●		●					●								●	●		
[63]	A. Attenberger et al.	2020	Jo.	●		●									●							
[64]	R. Rak et al.	2020	Jo.	●		●								●								
[65]	D. Kopenkova et al.	2020	Co.	●		●								●								
[66]	H.S. Lallie	2020	Jo.	●		●					●		●							●		
[67]	K. Dološ et al.	2020	Jo.												●	●			●	●		
[68]	N. Le-Khac et al.	2020	Jo.	●	●	●					●									●		
[69]	D. Steiner et al.	2019	Co.	●	●															●		
[70]	D. Sladović et al.	2019	Co.	●	●	●					●								●	●		
[71]	N. Vinzenz et al.	2019	Co.	●										●						●		
[72]	M. Hussain et al.	2019	Jo.			●		●												●	●	
[73]	C. Urquhart et al.	2019	Jo.	●		●														●		
[74]	C.J. Whelan et al.	2018	Jo.	●							●									●		
[75]	A. Koch et al.	2018	Co.	●		●	●													●		
[76]	S. Tatjana et al.	2018	Co.	●							●		●	●						●		
[77]	F. Leuzzi et al.	2018	Co.	●		●	●							●	●				●			
[78]	I. Cvitić et al.*	2018	Jo.	●	●																	
[20]	C. Huang et al.	2017	Jo.			●		●				●								●		
[79]	Z.A. Baig et al.	2017	Jo.	●		●		●												●	●	
[31]	D. Jacobs et al.	2017	Co.	●	●	●														●		
[80]	R. Altschaffel et al.*	2017	Co.			●	●				●			●					●			
[81]	W. Bortles et al.*	2017	Co.	●							●		●	●					●	●		
[82]	J. Lacroix et al.	2016	Jo.	●	●				●		●	●								●		
[83]	J.S. Ogden et al.	2016	Jo.	●	●						●			●					●	●		
[84]	N.Krishnamurthy et al.	2014	Jo.										●							●		
[85]	D.W. Park et al.	2014	Jo.	●										●						●		
[86]	K. Lim et al.	2014	Jo.	●									●									
[87]	J. Johnson et al.*	2014	Co.		●		●				●			●				●	●			
[88]	T. Hoppe et al.*	2012	Jo.								●		●						●	●		
[89]	S. Al-Kuwari et al.*	2010	Co.									●	●									
[90]	D. Nilsson et al.	2008	Co.				●												●			
[91]	J. Daily et al.*	2008	Co.	●							●			●					●			
[92]	D. Nilsson et al.*	2008	Co.	●															●			

– We have identified categories based on focus areas in the selected work and mapped the technical solutions in this work to the security properties which are considered.
– We identified and categorized forensically relevant data and mapped this data to potential stakeholders.
– We have also identified and discussed challenges, issues and research gaps within the area of automotive digital forensics.

We believe our contributions can be used as the basis for incorporating forensics into vehicle design for stakeholders such as automakers and law enforcement agencies. We also believe that our contributions can guide both performing automotive digital forensic investigations and encourage more research within this area.

The remainder of this paper is organized as follows: In Section II, we list and explain requirements and the security properties used, Section III lists stakeholders for automotive digital forensic, and Section IV presents related work. Section V details our approach and presents a comprehensive list of further categorized and mapped elements based on automotive forensic relevance. Section VI presents the identified data categories mapped to security properties and data users, followed by a discussion of the result in Section VII. We end the paper with the conclusion in Section VIII.

II. REQUIREMENTS AND SECURITY PROPERTIES

A forensically enabled vehicle must fulfill basic security requirements and support techniques such as secure data logging and secure data storage [16]. A forensic investigation requires trust in the chain of events, such as the logical order of braking, acceleration, and steering. Moreover, digital forensics has strong dependencies on information security to ensure trustable data. As mentioned in Section I-C, a forensics investigation includes the following basic steps, here further elaborated with relevant questions.

Identification. What is the reason for the incident? What data is relevant, and where is the data stored? What resources, e.g., tools and subject matter experts, are needed?

TABLE III
A MAPPING FROM DATA TYPE TO SECURITY PROPERTIES AND DATA USERS

Data Category and Reference	*Security Properties	**Data Users	Example of data and source
External devices. [24]–[26], [31], [32], [36], [63], [65], [68]–[70], [73], [74], [77], [81], [82], [112]	C.I.A.N.P.	LE.VM.VD.	*Cellular phones* may contain data from calendars, call logs, text messages, email communication, images, documents and location data. *USB memory* may contain documents and media files. *Remote keyless entry systems (RKE)* has information about VIN, time and date for use. *OBD-II dongles* has internal memory that may contain, e.g., logs with information about use.
Sensors. [18]–[20], [24], [31], [32], [38]–[40], [42]–[44], [47], [49]–[54], [57]–[60], [63], [65]–[71], [73]–[77], [79]–[82], [84], [88], [89], [112]	C.I.A.N.P.	LE.VM.VD.IC.	Information about speed, position (gps), temperature, airbag and object detection (from cameras, LIDAR and lasers).
Actuators. [36], [42], [75], [80], [82]	I.A.N.P.	LE.VM.IC.	Signals sent to initiate braking, steering and engine control (e.g., throttle).
ECUs. [18], [19], [25], [26], [31], [35], [44], [51], [52], [62], [63], [65], [71], [73], [75], [77], [79], [80], [82], [83], [90], [92]	C.I.A.N.P.	LE.VM.VD.IC.	Data storage, may contain information about operating mode(s), internal state and decisions recently made.
Software update events. [62], [73], [88], [90], [92]	C.I.A.N.	LE.VM.VD.	Software update logs, e.g., information about failed and successful installations, software version numbers and authorization attempts. Detected events with regard to software manipulation, e.g., caused by cyber attacks or the vehicle owner that tries to extend functionalities.
Security events: Diagnostics. [19], [24]–[26], [31], [41], [62], [66], [68], [71], [73], [77], [79], [83], [88], [90], [92]	C.I.A.N.	LE.VM.VD.	Events such as attempts to activate/deactivate firewalls and run privileged diagnostics commands.
Security events: Resilience tech. [54]	C.I.A.N.	LE.VM.	Executed resilience mechanisms (events) as a response to, e.g., a malfunction or cyber security issue. Examples of such events can be a system state change, reconfiguration, and migration [16].
Security events: Anomalies. [25], [26], [55], [79], [80], [90]	C.I.A.	LE.VM.	IDS software. Detected anomalies (events) in communication traffic, such as during a cyber attack, e.g., port scans, brute force and DoS. Events in forbidden situations, e.g., software updates during driving and maximum velocity during parking mode. Events connected to malfunctions, erroneous results and hardware failures.
Safety events. [18], [19], [24]–[26], [36], [38], [41], [50], [52], [55], [57], [60], [62], [63], [66], [69]–[71], [73]–[79], [81], [90], [113]	C.I.A.N.	LE.VM.VD.IC.	EDR crash data. Information about safety-critical events such as braking, acceleration, steering, engine control, airbag release, and seat belt traction/on/off. Indirect safety-related data such as warning messages about distracted driving and tired driver.
Normal events. [24], [44], [62], [63], [70], [74], [77], [81], [90]	I.A.N.P.	LE.VM.VD.IC.	Events such as opening/closing of doors and trunk. Turing on/off the engine and events with regard to activation/deactivation of alarms and locking/unlocking the vehicle. Information about fuel levels and consumption, and oil level and temperature.
Biometrics. [88]	C.I.A.N.P.	LE.VM.VD.IC.	Events from biometric driver identification such as face recognition, fingerprints, and voice.
Settings. [88]	C.I.A.N.P.	LE.VM.VD.IC.	Information about individual driver settings with regard to position of the seat, mirrors and driving mode (e.g., economic, sport and normal).
Vehicle-2-Vehicle. [18], [24], [32], [44], [46], [50], [57], [59], [62], [63], [68], [70], [74], [82], [114]	C.I.A.N.P.	LE.VM.VD.IC.	Communication within VANETs such as information about accidents, location and traffic conditions.
Vehicle-2-Infrastructure. [18], [24], [26], [31]–[33], [44], [46], [50], [57], [59], [62], [63], [68], [72], [74], [82], [90], [114]	C.I.A.N.P.	LE.VM.VD.IC.	Information from traffic lights and traffic regulations, e.g., speed, as well as detected violation of such. Exchanged traffic information, e.g., information about weather conditions, accidents and traffic jam, and route recommendations.
Vehicle-2-Pedestrian. [32], [44], [50], [68]	C.I.A.N.P.	LE.VM.VD.IC.	Information about location data (e.g., gps and gyrometer in relation to collision and rollover detection). Communication from vehicle to the relevant profession concerning accidents, e.g., vehicle to a physician or police.
Vehicle-2-Cloud. [18], [25], [32], [57], [63], [65], [67], [68], [72]	C.I.A.N.P	LE.VM.VD.IC.	Events from IDS/IPS, software updates and traffic information. Communication using applications (e.g., onStar [105] and VolvoOnCall [104]).

*(C)Confidentiality, (I)Integrity, (A)Availability, (N)Non-Repudiation and (P)Privacy.
**(LE)Law enforcement, (VM)Vehicle Manufacture, (VD)Vehicle Driver and (IC)Insurance Company.

Preservation. How can we preserve integrity during data collection? For instance, for running devices and devices with remote access capabilities. Can the devices be turned off without losing data? Can data be remotely changed or erased?

Acquisition and verification. How can we extract the data (e.g., creating images and performing live acquisition)? How can we validate the authenticity of the data (e.g., with signatures and hashes)?

Analysis. What type of information is relevant to assess?

Reporting. How can we document all parts of the forensic investigation process?

Requirements to ensure admissibility in legal proceedings for vehicle forensic data regarding the fulfillment of security properties are stated in works, such as [18]–[20]. Thus, in line with these requirements, and shown in Fig. 3, we adopt the well-known *CIA* security triad extended with two other properties and consider the first four as prerequisites for securing vehicle forensic data and the fifth for personal data, namely:

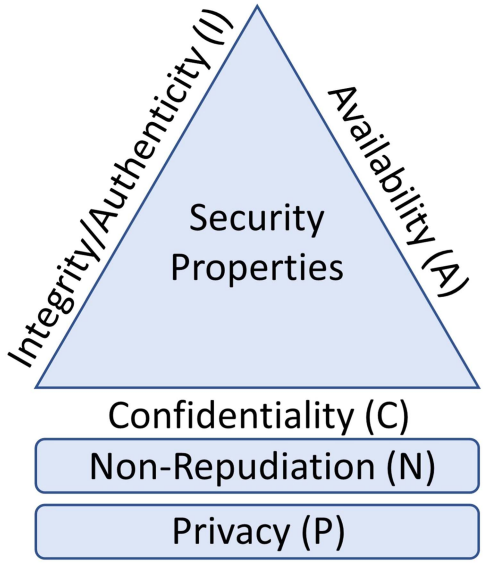

Fig. 3. Visualization of the considered security properties.

- *Confidentiality (C)*
- *Integrity/Authenticity (I)*
- *Availability (A)*
- *Non-Repudiation (N)*
- *Privacy (P)*

Confidentiality ensures that only authorized entities can access and disclose data. *Privacy* is related to personal data, such as traffic violations, location data, and synced data from external devices (e.g., text messages, calender's and phone records). Thus, there is a need to protect such data according to local laws and regulations [21]–[23]. *Authenticity* is a form of *integrity* that ensures data origin and is of particular interest for forensic investigations. *Availability* of data should be ensured, e.g., in the event of a crash and secure and tamper-proof storage guaranteed. *Non-Repudiation* ensures that the occurrence of an event and its origin can not be denied. Thus both *authenticity* and *integrity* are prerequisites for *Non-Repudiation*.

III. Stakeholders

We have chosen the four most common data users referred to in the literature [18], [19], [24]–[26].

- *Law Enforcement (LE)*
- *Vehicle Manufacturers (VM)*
- *Vehicle Drivers (VD)*
- *Insurance Companies (IC)*

LE such as the police and the related legal system requires reliable data to make a case and for the data to be admissible in a court of law. *VMs* need to have fault tracing data to distinguish between software and hardware failures and cyber security issues, e.g., fixing bugs and releasing software update patches. *VD* may manipulate forensic data, e.g., to hide, remove or manipulate digital evidence. *ICs* are interested in insurance cases and accidents and cost/risk policies for driver behavior (e.g., driver profiling).

IV. Related Work

In 2004, the National Institute of Standards and Technology (NIST) published SP 800-72 for Personal Digital Assistant (PDA) forensics [27] of PDAs such as Pocket PC, Palm OS, and Linux based PDAs. Currently, PDAs have to a large extent, been replaced by other technologies such as smartphones and apps. In 2006, NIST released SP 800-86, a document for practical guidance on performing computer and network forensics [28]. SP 800-86 defines digital forensics as applying science to the identification, collection, examination, and analysis of data while preserving the integrity of the information and maintaining a strict chain of custody for the data. Another standard for digital forensics, the ISO 27037, was established in 2012 further reviewed and confirmed in 2018 [29]. The ISO 27037 provides guidelines for identifying, collecting, acquiring, and preserving digital evidence. In 2014 SP 800-101r1 was released for mobile device forensics providing guidelines for tool usage and procedures [30].

As mentioned in the introduction (cf. Section I-B), another related area is IoT forensics which focuses on devices with Internet capabilities, including smartphones. For example, in [4], Stoyanova et al. list challenges, approaches, and open issues within IoT forensics. Although we can find similarities between IoT/mobile and automotive digital forensics (e.g., embedded devices with limited performance), automotive forensics is a comparatively different area, given the complexity of IVNs (cf. Section I-A) and safety-critical requirements.

Previous work within automotive digital forensics (cf. Table I and II) briefly mention the field of automotive digital forensics, such as issues and challenges. However, to our best knowledge, there has so far not been any systematic literature review that offers the comprehensiveness of present work. We have assessed over 300 papers and contribute with a comprehensive categorization and overview specific to the digital automotive forensics landscape based on focus areas, forensic data, security properties, and stakeholders. In the absence of standardized methods or guidelines for automotive digital forensic investigations, guidelines that consider the complexity of the vehicle architecture are of vital importance.

V. A Systematic Literature Review

A. Approach

We have performed a systematic literature review based on the approach visualized in Fig. 4.

In total, 327 papers were acquired from the searches. Each individual database search was assessed further in a screening process by first reading the title followed by the abstract, conclusion, and for potentially relevant papers also skimming through the whole text in the article. Papers were included based on:

1) *relevance to the automotive domain.*
2) *papers published between 2006 - 2021.*
3) *papers published in journals and conferences.*

We have excluded articles:

1) *not specific to automotive digital forensics.*
2) *not written in English.*

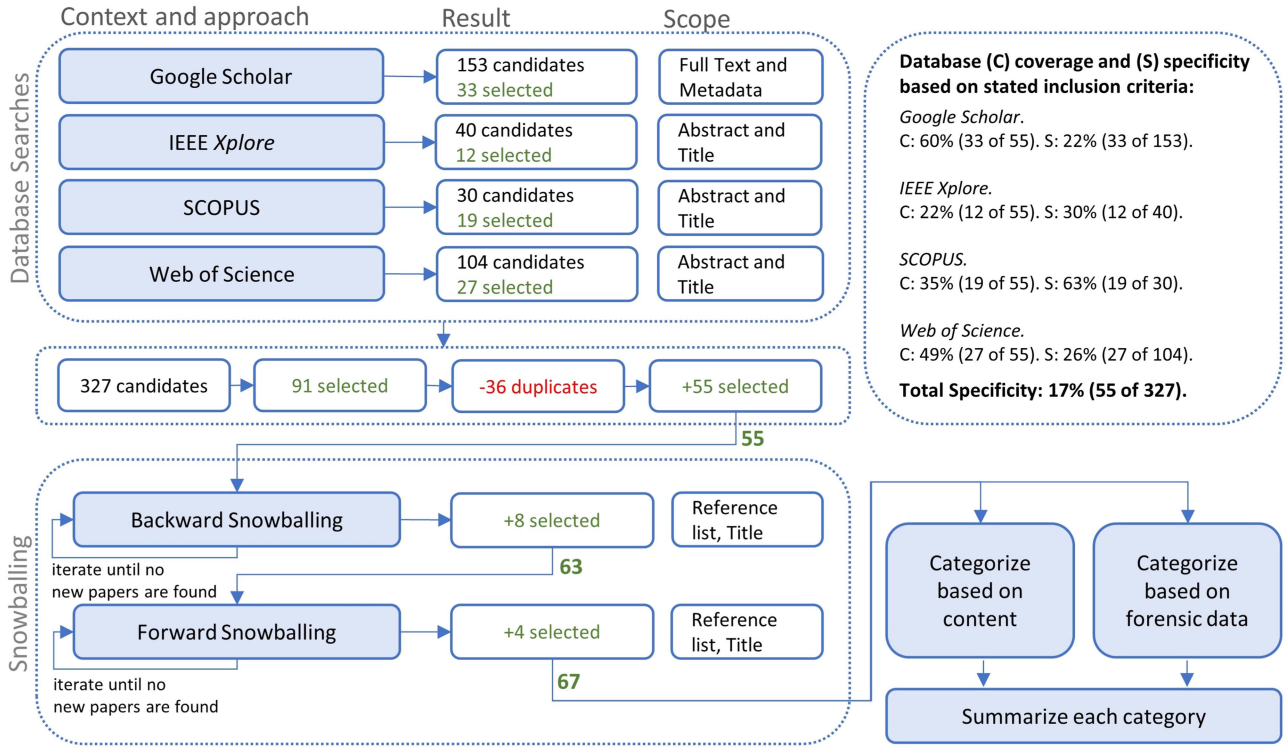

Fig. 4. The process of our approach, result and scope.

We have used the following search terms: *forensics* in conjunction with *vehicle*, *car*, or *automotive*. As shown in Fig. 4, *Google scholar*[1] resulted in 153 candidates, where 33 papers were selected. *IEEE Xplore*[2] resulted in 40 candidates, where 12 papers were selected. *Scopus*[3] resulted in 30 candidates, where 19 were selected. *Web of Science*[4] resulted in 104 candidates, where 27 papers were selected. Fig. 4 shows that the screening process resulted in 91 papers. After removing 36 duplicate papers, 55 papers remained. Fig. 5 shows that some of the selected papers were present in more than one database where, e.g., *SCOPUS*, *IEEE Xplore*, and *Web of Science* have three of the selected papers in common [25], [31], [32]. Only two selected papers were found in all four databases [25], [31]. The selected papers that are unique and can only be found in one of the four databases are also shown where, e.g., *Google Scholar* has 14 unique papers and *IEEE Xplore* only one [33]. Figs. 4 and 5 are further discussed in Section VII.

Snowballing. We then performed a snowballing approach [34]. As shown in Fig. 4, we first performed *Backward Snowballing* where we extracted all references from the previous result and used the same search criteria as before with regard to the title and venue. Duplicates were removed, and another eight papers were found. We then performed *Forward Snowballing* with *Google Scholar*, based on those papers which cite any one

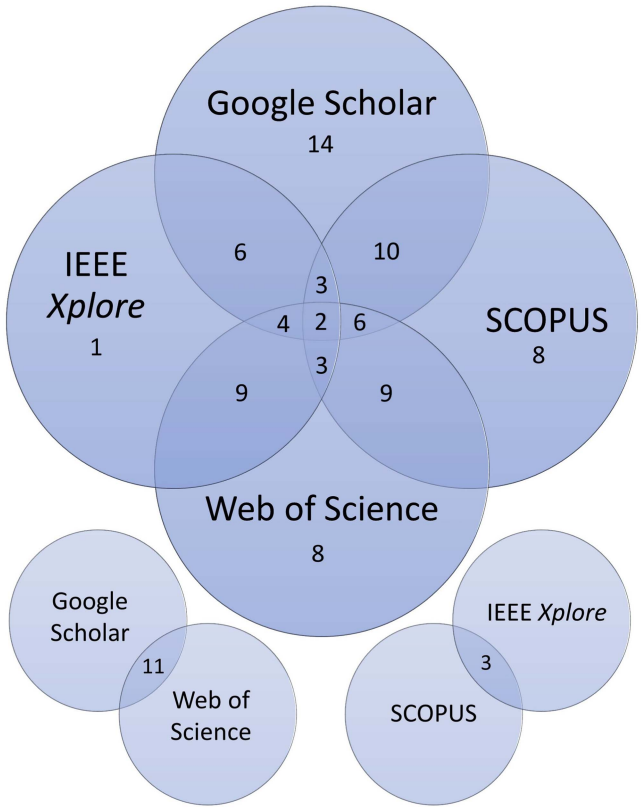

Fig. 5. Overlaps and uniqueness for the selected papers.

[1]https://scholar.google.com/ search date: 2021-02-16
[2]https://ieeexplore.ieee.org/ search date: 2021-03-22
[3]https://www.scopus.com/ search date: 2021-03-24
[4]https://www.webofscience.com/ search date: 2021-03-29

of the previous 63 papers. The same search criteria were used once more, and another four papers were found; and as shown in Fig. 4 and listed in Table I and II, 67 papers were selected in total.

B. Categorization of Papers

We have divided the result from the SLR into two main categories: *technical solutions* and *surveys*. Articles that propose technical solutions are included in the former category, and all others are included in the latter. Furthermore, all papers are mapped to one or more of the below 11 categories that were identified during the SLR based on focus areas we could identify when reading the papers. We specify whether the papers are published in a conference or journal and what security properties are considered concerning the proposed technical solutions.

In the remainder of this section, we have picked the most representative studies for each category and refer to Table I and II for the complete list.

1a. Data: Data Collection.

Papers in this category discuss the different types of forensic data and retrieval of such.

A. Attenberger [63] presents considerations for data generated in vehicles and categorizes these into two groups, namely, *front end* and *back end*. The former are vehicle electronics inside the vehicle, such as the infotainment module, and the latter, outside the vehicle, such as the cloud. A significant challenge for the *front end* is that there are no standardized interfaces for information extraction and no standardized format for storage. Additionally, in some cases, debug ports are lacking; thus, it is necessary to remove storage circuits for further data handling. Moreover, considering the steady increase of relevant data, manual approaches become infeasible.

Vinzenz et al. [71] investigate data storage in vehicles and analyze its significance. Four data types were identified, airbag Event Data Recorder (EDR), Electronic Control Unit (ECU), Telematic Platform, and Infotainment System, where the EDR data was emphasised.

Utilized by a case study on a 2016 Mercedes Benz E-Class model, D. Steiner et al. investigate the communication in a modern vehicle and the forensically relevant data artifacts which can be retrieved [69]. Essential data and communication endpoints were identified.

L. Cintron et al. [50] model a transportation event data collection system as a *Hyperledger Fabric blockchain Network* which is simulated in a virtual transportation environment. Available accident data is collected from other vehicles and roadside units. An open-source framework tested in a production environment with community and commercial support are highlighted as beneficial in their solution. However, their solution was simulated as a limited environment with stated assumptions and not tested in a natural setting that most likely does not adhere to these expectations. Several benefits, drawbacks, challenges, security, and privacy considerations for their solution are mentioned, e.g., it is beneficial that all identities in a distributed ledger network are authenticated and thus accountable for their action. At the same time, this raises privacy concerns. Storage is another issue since the ledger increases in size quickly. There are also reliability concerns because used components still are immature, as well as requirements for high-throughput that might not be fulfilled.

1b. Data: Extraction Techniques. *A category that looks into techniques for extraction of automotive digital forensic data from IVN memory storage.*

In [41], J. Daily et al. present a data extraction technique for non-volatile memory when standard extraction via vehicle diagnostic tools is not possible due to, e.g., damaged electronic control modules. The in-circuit debug port (JTAG) is suggested as a non-destructive data extraction method. However, manufacturers sometimes close these ports for production vehicles to protect against cyber attacks; thus, the assumption, in this case, is that the port remains open when vehicles reach the market. The first step is to extract the complete image as raw binary data or decoded data to another workable destination, such as a surrogate that is not damaged. Two different tools for data extraction are tested, i.e., the Alientech KTag [93] and the PEmicro Cyclone [94]. The extracted images from the two tools were compared by calculating an SHA-256 hash, and the result was a match with the conclusion that any one of the two tools is acceptable to use. The next step was to extract specific events from decoded data, e.g., sudden deceleration. The performed data extraction is valid for the actual hardware used and might not apply to all hardware.

In [70], D. Sladović et al. describe the digital forensic stages performed within an investigation, the type of information extracted from the vehicle, and the extraction process. Three connection points are mentioned: the OBD-II port, directly to an ECU, and directly to the EPROM (requires disassembly). Finally, the infotainment system is discussed as containing the most useful information. The Berla iVe [95] is the recommended tool for data extraction from infotainment units. However, something to consider before purchasing tools is to validate brand compatibility.

2a. Challenges: General Challenges.

A category that considers automotive digital challenges of a more general nature.

In [64], [65], Rak et al. provide an introduction to automotive digital forensics and provide examples of a few data sources, e.g., EDR, telematics, keyfobs, ECUs, and cameras. A few issues are mentioned, e.g., the high development pace aligned with the strive to gain an advantage over competitors results in a lack of security measures. Thus, vehicles become more vulnerable to hacker attacks. Moreover, the lack of standardization in, e.g., data interfaces, recording units, data storage, and lack of unified approaches for the digital forensics process makes different manufacturers use their own developed strategies. Thus, making forensic investigation challenging due to variations between brands and vehicle models.

In [31], D. Jacobs et al. focus on that vehicles have various standalone computing devices, lack of data security, difficulties in extracting data, lack of guidelines and tools for vehicle forensics, and problems obtaining proprietary information from manufacturers. Simerly, in [68], Le-Khac et al. highlight that data in vehicles is spread in a distributed system in various

locations, which requires extensive manual work to find and extract relevant data for automotive digital forensics investigations. Additionally, data needed to, e.g., connect a driver to a crime is in many cases not sufficient for a majority of existing vehicles. In contrast, a modern smart car contains a vast amount of valuable data, but there is still no available process or framework to guide automotive digital investigations in this case. Moreover, very few automotive forensic data extraction tools exist, e.g., Berla iVe [95] for infotainment and telematics systems limited to a few brands. Aligned with [31], [64], [65], the lack of security and forensic mechanisms and a framework to guide the forensic process are identified as challenges.

In [75], A. Koch et al. emphasize on data collection and its privacy-related challenges, such as performing privacy evaluation of all data and adhering to the existing laws and regulations. The importance of user transparency regarding data stored, transmitted, and processed is highlighted. Three data streams for data collections are identified, namely main memory, mass storage, and communication. The management of the increasing amount of data and how to retrieve data from these data streams are mentioned as challenges.

In [80], R. Altschaffel et al. discuss challenges such as the absence of openly discussed automotive forensic processes within the scientific community, lack of standardized components, inaccessible memory due to security measures to guard intellectual property, low storage capacity for in-vehicle devices, and no authentication of messages for in-vehicle communication. They mention that the reconstruction of previous events must follow scientific and well-proven principles to preserve the authenticity and integrity of the data and highlight the importance of the investigator not being affected too much by a starting hypothesis, something which can lead to bias. Existing solutions are rare, often isolated, and limited by secrecy and intellectual properties. In [82], J. Lacroix et al. mention challenges for digital forensics for VANETs, such as the continuous changes in network topologies, unreliable communication channels, and source tracing difficulties. They suggest using GPS data, vehicle cameras, and analysis of data remnants from infotainment system applications to trace and find perpetrators better.

2b. Challenges: Requirements, Guidelines.

A category that proposes solutions to make automotive forensic data admissible for digital forensic investigations.

N. Vinzent el al. [19] state requirements for securing vehicle forensic data, and propose to adapt the telematics unit to accommodate storage for a limited dataset by using a circular buffer (i.e., a *Last In First Out* approach). Furthermore, they state that their implementation only requires minor software changes and no additional hardware because they consider the latter an unrealistic requirement due to its associated costs. However, current vehicle hardware usually consists of the least required memory for the task at hand; thus, additional memory still has to be added (i.e., hardware changes) since it would otherwise consider a too small time span for a realistic amount of required forensic data. Moreover, the telematics unit, most likely by design, only has access to a fraction of the data since the telematics unit can belong to a separate domain for security reasons with limited data access. In our opinion, IVNs have to be adapted both in hardware and software to enable a realistic approach for storing available forensic data.

In [62] and [90], D. Nilsson and U. Larsson present a list of requirements for data collection and event reconstruction in three categories. First, requirements for detection and storage of security violations. Second, requirements to address the five forensics W questions: who (traceability for event), what (type of event), where (sender/receiver ID), when (time for start, duration, and end), and why (data/value content). Third, a list of hashes for all ECU firmware should be securely stored and accessible for comparison when extracting in-vehicle firmware to detect manipulation. We consider the first category as an imperative prerequisite for automotive digital forensics, which potentially can be solved by Intrusion Detection System (IDS) mechanisms. For the second category, security mechanisms such as cryptographic primitives for validating the authenticity of events can be considered. The last requirement relates to firmware update where mechanisms such as signed software and secure boot can mitigate undetected manipulation of software.

F. Leuzzi et al. provide an organizational framework of requirements for the traffic police to guide future research efforts [77]. The emphasis is on road events, such as traffic congestion, accidents, crimes, or natural disasters. Data from events can be found inside the vehicle, e.g., logs, and outside the vehicle, e.g., traffic data. Aligned with [63], machine learning is mentioned as important in future crime investigation. Automatic approaches to managing large volumes of data, alerting when data matches previous crimes, and predicting and finding preparations patterns for potential future crimes are identified as important in future research. Data from, e.g., number plate detection systems and locations from cellular devices' can create a database over the traffic flow to be used in data mining. The goal can vary, but a few examples are mentioned, such as linking a particular vehicle and individual to a specific location and time, aligned with opening and closing doors to determine the time window when an individual left and got back to the vehicle. Synced data between individuals' cellular devices and vehicle communication is an important source, e.g., phone calls, text messages, and calendar data. However, a significant challenge is privacy, where several organizations are still not compliant with the law. The lack of standardization for vehicles regarding data quantity, quality, and formats makes data retrieval problematic.

3a. Communication: Cloud/Fog/Edge.

Cloud, fog, and edge node communication concerning digital automotive forensic data is the focus of this category.

In [25], Mansor et al. suggest a mechanism that enables data collection and transfer to the cloud via smartphones. The phone is proposed to be connected to the OBD-II port via a Bluetooth or WiFi interface. However, exposing a Bluetooth or WiFi interface to the OBD-II port to communicate with the phone can potentially create a bridge between the vehicle and the Internet, thus being considered a cyber security risk. The OBD-II facilitates the execution and tracking of sensitive diagnostic commands. Various vulnerabilities and exploits related to the OBD-II connection have been found in the past, e.g., [96]–[103]. For instance, a user's smartphone can be compromised (e.g., via malicious applications), creating remote access for hackers.

Trust in external devices should be kept to a minimum, and storage and transfer to the cloud are better handled by internal mechanisms controlled by the vehicle manufacturer. Already existing phone applications developed by the vehicle manufacturers (e.g., VolvoOnCall [104], OnStar [105]) are better suited due to increased control and, thus, less risk. The vehicle and the application can use secure connections and communicate via trusted cloud sources, and application privileges can be separated and controlled according to architectural design decisions, such as by isolating safety-critical functions.

According to C. Huang et al., there is a trend to move away from cloud implementations in favor of fog/edge computing to save communication bandwidth [20]. Fog nodes can be situated near roads to collect, process, and store data. Thus, roadside units are a potential enabler for fog nodes as an extension of cloud servers with performance benefits in communication and data storage. Increased traffic control to improve safety and data for forensic investigations are potential use cases. However, fog computing is still in early development with many challenges, e.g., security and a large volume of data. C. Huang et al. elaborate that many millions of connected cars, with an average of 30 TB of produced data every day, clearly challenge storage, processing, and bandwidth capacity.

In [33], X. Wang et al. propose a scheme to speed up accident handling based on *Multi access Edge Computing (MEC)* to determine the liability in rear-end accidents. This scheme consists of a forensic model for data collection of driving information before and after a collision to establish a data chain for the vehicles involved in the accident. Vehicles are assumed to periodically upload data, such as position, speed, and acceleration, to an edge infrastructure that collect, process, and analyze data. A MEC infrastructure is suggested for accident-prone locations, such as intersections and parking lots. An evaluation metric based on the driver attention level is presented, and estimation of vehicles liabilities with regards to accidents. However, data regarding a few places is not enough; therefore, a broader approach covering various areas and traffic situations is necessary. Moreover, more work is needed to establish if the proposed liability scheme connected to a driver's attention accurately reflects the fact.

3b. Communication: VANETs.

A category that looks into automotive digital forensic solutions for Vehicular Ad hoc Networks.

In [46], Z. Ma et al. propose a digital watermark algorithm for VANETs, embedding location, timestamp, and forensic device data into a real-time vehicle accident photograph. Digital watermarking is a technique used to hide a data label in digital carriers such as images and multimedia to verify integrity and authentication. VANETs can be used to communicate traffic warning messages to surrounding vehicles, and photographs can add valuable information in addition to text messages. However, there are several challenges, such as ensuring trustworthiness and the privacy aspect of the communication, still the main goals in [46] are to ensure integrity, detect tampering and at the same time enhance user privacy in digital photographs. Additionally, an approach based on neural networks for vehicle license plate recognition and information gathering is proposed.

In [57], R. Hussain et al. introduce incentives-based vehicle witnesses, which utilizes moving vehicles and roadside units as witnesses to incidents. The emphasis is on security, privacy, and the adoption of the proposed service. Cameras in roadside units and vehicles are assumed to collaborate and take pictures of their environment. Forensic data should be sent anonymously to the cloud, thus preserving privacy. However, although privacy might be ensured for the data source (e.g., a vehicle), data may include potentially sensitive information, such as videos and pictures of individuals, an issue also highlighted by H.S. Lallie [66].

4a. Software: Applications and Software.

A category that focuses on specific applications and software used to manage and communicate automotive digital evidence.

In [60], A. Sathe et al. propose a *road safety and location monitoring system* by using a module in the vehicle to gather geographical coordinates and acceleration variations. The aim is to provide road condition updates. A three-axis accelerometer is used, and the authors suggest a correlation between sudden changes in these axes to confirm road disturbances. The location for these road disturbances is mapped to accident zones. A phone application is used to retrieve updates from the database. However, the work lacks the security and privacy aspects of their solution. Insecure HTTP communication is used for the communication, which leads to the potential for malicious actors to intercept privacy-sensitive data.

N. Watthanawisuth et al. [61] propose a similar approach using an accelerometer, GPS device, GSM module, and a microcontroller to detect and act on accidents and automatically send messages to appropriate recipients such as a family member or an emergency medical service. Although the title of their work indicates usability for vehicles, the practical tests performed, with a claimed high detection accuracy, were only done for bicycles, and the same deficit as in [60] applies; the security or privacy aspect is not discussed for their solution.

4b. Software: Forensic Tools.

A category that focuses on tools that might be applicable for automotive digital forensics.

J. Lacroix et al. [82] highlight that automotive forensics is an under-researched area. Privacy aspects have to be considered for IVN data due to its uniqueness concerning driver habits and actions. They discuss the type of data possible to extract, such as data dumps through OBD, USB, and JTAG ports. They examine a data dump from a truck infotainment system with tools such as *Forensics Toolkit (FTK)*, *Encase*, and *Autopsy* and present the extracted sensitive data. The data can be used to derive information about who previously has driven the vehicle, which can be helpful in forensic investigations.

In [76], S. Tatanja et al. investigates devices and tools for reading out data from vehicles. Many devices are mentioned, such as Bosch CDR500 (Crash Data Retrieval) for post-crash analysis and Bosch KTS540 for diagnostics and retrieving fault codes and reports. A case study of a Toyota Yaris previously involved in an accident was performed, where forensic data retrieval was performed with a Bosch CDR500 on the EDR in the vehicle.

In [68], Le-Khac et al. compare a few general forensic tools: *Encase*, *Accessdata Forensic ToolKit*, and *Xways Forensic*. The

tools are compared based on the compatibility of different filesystems and non-structured memory dumps with varied results. Notably, none of the tools support the QNX filesystem, which is common in the automotive industry. There are very few automotive digital forensics-specific tools available, and as mentioned in [76], the *Bosch CDR* is discussed, and the *Berla iVe*.

5a. Hardware: Architecture.

A category around vehicle architectural design and its alignment with automotive digital forensics.

In [52], L. Davi et al. propose a blockchain architecture for ECUs that can be used as a blackbox. Their approach utilizes consensus algorithms to allow only agreed transactions to be added to the blockchain. However, safety-critical systems' real-time requirements make this approach less practical. The cost of accommodating an utterly new architecture where all ECUs can sign and validate messages has to be considered.

J. Lacroix et al. [82] briefly introduce the vehicle architecture, such as explaining internal components and communication busses. Especially, the CAN bus is highlighted as an essential source for live vehicle forensic data as it can contain relevant error messages. Moreover, the infotainment system holds valuable data, e.g., media content from external devices, internal logging, and localization data.

5b. Hardware: Sensor.

Automotive digital forensic solutions and mechanisms concerning various sensors such as GPS, LIDAR, and cameras are the focus of this category.

In [58] and [47], A. Mehrish et al. discuss the increased use of cameras by, e.g., law enforcement and private individuals and its relation to forensics. However, for videos produced by vehicle dashboard-mounted cameras to be admissible in the court of law, the video's authenticity must be verified. The authors propose an algorithm to extract engine vibration characteristics from blur patterns in the video as a unique vehicle signature and claim an identification accuracy of 91.04 percent. The generated signature, together with other related forensic data, may help achieve a higher accuracy level. The captured video may contain information about unrelated individuals and vehicles; thus, privacy concerns shall be considered and addressed in such a case [21]–[23].

In [84], N. Krishnamurthy et al. presents a new area of research called audio-based vehicle verification, a field that can be useful within automotive digital forensics. Their work experiments on vehicle sound to verify whether a particular sound sample can be mapped to a specific vehicle. The sound from the engine and air-conditioner was shown practical and discriminative for these use cases. However, there are many challenges, such as that audio-based vehicle verification is sensitive to sound disturbances. Moreover, the privacy and security aspects are not mentioned; noise data collection (i.e., audio recordings) and storage of such data require these considerations.

5c. Hardware: EDR and Blackbox.

A category that discusses requirements for Event Data Recorders and Blackboxes concerning automotive digital forensics.

In [87], J. Johnson presents a review of the digital forensics concept, emphasizing on data recorders, such as EDRs, and cryptographic methods to ensure the integrity of digital evidence. The authors discuss the trustworthiness of extracted data using existing methods and propose recommendations to align with currently accepted standards for digital forensic soundness. Examples of such can be archiving network data and encrypting and hashing to conceal and detect tampering of data. Moreover, they mention that forensic soundness is defined by the methods of extracting, analyzing, and presenting digital evidence that must be performed in such a manner that the results can be used in legal proceedings with a high degree of confidence in their admissibility. However, today, data extraction concerning vehicles often uses tools unsuitable for digital forensics, thus lacking forensics soundness requirements. Current storage is often not resistant to tampering, and if encryption is used, it is still often too weak, and no integrity validation occurs.

6a. Algorithms: Machine Learning.

Papers in this category look into the automation of the automotive forensic process concerning data handling.

In [42], P. Sharma et al. propose a forensic investigation protocol utilizing a supervised deep neural network architecture for post-analysis of attacks targeting vehicle sensors. Anomalies in gathered data from IVN memory storage, together with an analysis of an accident, can reveal traces of attacks such as spoofing/jamming of sensors. However, it is stated that the investigator should use real-world sensor data for the analysis, but instead, a simulator is used for the data generation. Thus, it would have been interesting to see if their protocol could detect sensor manipulation within an actual vehicle. A more in-depth elaboration on the potential use cases for their approach is lacking, which would have justified the useability.

In [67], K. Dolos et al. use a freely available data set from the *Hacking and Countermeasure Research Lab* [106] to identify and classify drivers based on driving behavior. A hypothetical hit-and-run forensic scenario was used with three suspects, and a supervised learning algorithm was utilized to find the most likely suspects. However, due to the novelty of the area, additional work is needed before such driver identification methods can be used within a forensic context.

In [63], Attenberger et al. suggest Machine Learning to automate data handling and driver identification. Supervised classification can be used to train an algorithm into classes of data connected to individual drivers. These classes can then be used to determine who has been driving the vehicle. However, since no standardized exchangeable data format exists for the automotive, automated approaches such as these become challenging. The Cyber-investigation Analysis Standard Expression (CASE) is mentioned [107] as a possible standard for the automotive to use.

6b. Algorithms: Other Algorithms.

A category for proposal of specific algorithms for management of automotive digital forensic data.

In [55], M. Marchetti et al. highlight that there is no public specification for the proprietary data exchanged over IVNs; this limits researchers' potential to develop solutions that detect deviations. Thus, they propose *Reverse Engineering of Automotive Data Frames (READ)*, an algorithm to extract unknown CAN messages to identify and label CAN frames. The authors in [55] argue forensic usability due to the possibility of identifying

deviation in time series for periodic messages (e.g., acceleration and steering), which can indicate sharp turns, acceleration, and braking, known as indices before potential accidents.

In [38] and [48], M. Waltereit et al. propose an approach to calculate the probability that suspects were present at a crime scene when GPS data is unavailable. The algorithm executes in an automated manner and saves time compared to other manual approaches. Further, they consider a hit-and-run accident as stated by Hoppe et al. [88], where some suspects without alibi claim their innocence. A route reconstruction within proximity to the incident area is suggested for the involved vehicles to absolve innocent suspects. In-vehicle data concerning driving behavior such as braking, acceleration patterns, and wheel speed is used as input to an algorithm for reconstruction and likelihood for specific routes and the probability for suspects' location.

7a. Cryptography: Blockchain.

A category that looks into papers that propose solutions utilizing blockchain technology.

In [36], a high-level traffic investigation framework for sensor data is proposed based on a decentralized identity distribution on blockchain, derived from case studies of accidents utilizing digital data. Due to its high-level abstraction, an industrial application of the proposed method seems infeasible. In [35], C. Oham et.al. propose a blockchain-based framework for securing the IVN and argue that using this framework keeps track of authorized historical operations (state changes) executed by vehicle ECUs, thus enabling traceability and identification of compromised ECUs. However, the framework only considers actions from the vehicle manufacturer, service technicians, and communication between the vehicle and roadside units, thus failing to provide complete event traceability. A more comprehensive range of potential events caused by other actors is needed.

In [43], H. Guo et al. propose an event recording system enabled by blockchain with a *Proof of Event* mechanism for vehicle networks. The aim is to provide trustable information about events based on an election algorithm that selects a leader/verifier in a blockchain network consisting of three participants: *accident*, *witness*, and *verifier*. A lead verifier (a vehicle or a roadside unit) is selected with the help of an election algorithm, and other vehicles can either be witnesses or other verifiers. The involved participants in the proximity of the accident will agree on the order of events according to a voting scheme and save these events into a blockchain. The authors implement a proof of concept prototype to demonstrate the feasibility of the proposed method. Given the diversity and complexity of IVN and its data [20], we believe the approach proposed by Guo et al. might only be applicable for a small fraction of the forensic data.

7b. Cryptography: Other Cryptography.

A category that discusses and proposes cryptographic primitives to secure automotive digital evidence.

In [40], M.A. Hoque et al. propose a solution named AVGuard for data collection and storage of forensics logs with the ability to verify data integrity. The integrity of the logs is verified by using a hash chain and a Bloom filter [108], and logs are encrypted to ensure confidentiality. Moreover, they demonstrate a proof-of-concept implementation of their approach. Proofs of the logs are suggested to be published on the web. However, securing the web's communication and storage is not discussed, and although integrity and confidentiality are considered for the data collection, they lack the privacy aspects for their solution.

8. Framework and Processes.

A category that discusses and proposes frameworks for the management of automotive digital evidence and simplification of automotive digital forensic processes.

In [44], A. Philip et al. propose a framework for road accidents and traffic violations based on deep learning and blockchain. An accident warning system for vehicles is established by considering road and climate conditions and driving patterns as parameters. The best parameters for specific traffic segments can be predicted, and warnings can be issued to vehicles from roadside units.

In [24], an approach for a permission-based blockchain framework is proposed for data collection of various types of data such as health data (e.g., from wearable devices) and automotive diagnostic. Their approach integrates the vehicular public key infrastructure (VPKI) into the blockchain to provide membership and privacy. M. Hossain et al. propose a vehicle data collection framework for distributed, decentralized, and mobile entities with secure storage [32]. Mechanisms to collect and store digital evidence and a specific algorithm for data integrity verification are proposed for evidence verification.

In [26], K. Buquerin et al. presents a potential automotive digital forensic process according to four phases: *forensic readiness*, *data acquisition*, *data analysis*, and *documentation*. The OBD-II port was used as an example for a connection point for data collection with Wireshark. A Packet Capture (PCAP) file and a hash of the file were stored for analysis, and finally, a report was generated.

In [37], C. Alexakos states various challenges for integrating digital forensics into the Internet of Vehicles (IoV) context. Examples are that there are no standardized data formats and a dynamic network topology, i.e., nodes are added and removed. Thus, the topology continuously changes in different vehicles models. Moreover, there is no open access due to intellectual properties and privacy concerns. The authors in [37] propose a forensic readiness tool that follows the digital forensic process model from Valjarevic et al. [109]. The tool is implemented by software into the nIoVe framework [110]. The tool's purpose is to collect forensically sound data, which enables reconstruction of events to be presented in a court of law and to learn, mitigate and predict future anomalies such as cyber attacks.

In [80], R. Altschaffel et al. suggest using a Desktop IT forensic process model from S. Kiltz [111] additionally to EDRs for the automotive domain, which consists of various investigation steps: first strategic and operational preparations, which consist of measures taken before and after incidents, followed by data gathering, data investigation, data analysis, and finally, the actual documentation of the complete investigation. Moreover, various tools and potential use cases are mentioned, as well as live and static data acquisition and forensic data acquisition from ECUs, sensors, and actuators.

In [88], T. Hoppe et al. introduce an overview of the digital forensics process and put this into an automotive context. They suggest using a data recorder that securely logs existing

navigation data transmitted over the CAN bus. Data, such as route information (e.g., street names) aligned with other data (e.g., speed, start, or destination position), can be used to reconstruct vehicle routes for post-incident investigations and further used as indices to individuals' locations at the time of crimes. Information such as this can either connect or free someone from involvement in the incident.

9. Practical Experiments.

A category that focuses on different types of practical forensic experiments.

In [73], C. Urquhart et al. perform practical experiments to pinpoint vulnerabilities within a Scoda Octavia vRS and suggest vehicle components and data that can be considered for digital forensic investigations. Infotainment data (e.g., Bluetooth ID of paired devices, call logs, and pictures/thumbnails), GPS, ECU memory, and diagnostic messages are mentioned. C. Whelan et al. [74] present a study that investigates available forensics artifacts in two different infotainment systems: a *Uconnect* system and a *Toyota Extension Box*. The *iVe* tool from *Berla Corporation* was used for data acquisitions. *Uconnect* provided only location data and *Toyota Extension Box* extensive user-related information such as contacts, call logs, and location data.

Vinzenz et al. [71] have analyzed crash data from the NHTSA NASS CDS database retrieved in the proprietary EDRX-format from a Bosch EDR tool. Findings were that before the year 2000, only airbag deployment status was stored. After the year 2000, vehicle speed during crashes was included, and starting in the year 2005, additional data was added, such as engine throttle. Increasingly more data in the form of Diagnostic Trouble Code (DTC) could be found for upcoming years. In total, 28 different DTC types were identified. Privacy and security considerations are mentioned. However, it is not completely clear if the data is admissible in court due to privacy laws and the lack of security measures to guarantee trustable data.

Le-Khac et al. [68] perform two case studies on extracting vehicle data. For the first case, they extract data from an infotainment system. For the second case, they investigate communication traffic on GSM/3 G/4 G networks and discuss how potential sources of evidence can be intercepted (e.g., PCAP data, metadata, and call detail records (CDR)). In [31], D. Jacobs et al. present a case study on a Volkswagen Golf version 6, 2012 station wagon and discuss considerations to avoid potential data losses (e.g., do not start the vehicle).

10. Infrastructure/Smart Cities.

A category that look into infrastructure communication and smart cities with regard to automotive digital forensics.

X. Feng et al. [18] discuss Autonomous Vehicles (AVs), Smart Cities (SCs), and digital forensics with an emphasis on sensor data. They provide examples of forensic data handling issues such as no data integrity validation and often unprofessional data extraction compared to other forensic areas. The authors conclude that current automotive data forensic acquisition and analysis approaches are not acceptable concerning legal evidence. Moreover, they propose a mechanism for acquiring sensor data from AVs in SCs and securely uploading it to cloud storage.

In [79], Z.A. Baig et al. discuss the challenges of digital forensics in smart cars. The difficulty of proving the validity of and access to vehicle data and incorporating cars into the context of smart cities are examples of the challenges identified by the authors. A hypothetical use case of a reckless driver is presented with the interaction of other smart city entities followed by the forensic investigation process.

11. Trusted Execution Environments and Virtualization. A category that focuses on solutions that secure automotive digital evidence using Trusted Execution Environments (TEEs) and virtualization.

In [51], S. Lee et al. suggest an automotive data recording system named T-Box that executes inside a trusted execution environment to detect data manipulations such as deletion, replacement, replaying, and truncation. S. Lee et al. claim that the T-Box data can be used as digital forensic evidence. However, their approach is platform-dependent and fails to protect data against tampering before storage. Moreover, the T-box does not consider the confidentiality and privacy aspects of the stored data.

VI. Categorizing and Mapping Forensic Data to Security Properties and Data Users

Automotive digital forensics requires identifying, acquiring and analyzing data that potentially can be used as digital evidence. The relevancy of a particular data to a forensic investigation depends on the type of crime being investigated. In this work, we consider the following data types:

- *data at rest:* data stored in the memory of ECUs or other automotive modules.
- *data in transit:* data that flows over networks.
- *data executed:* by ECUs/modules leading to specific events, e.g., state changes.

As shown in Table III, we have identified forensically relevant data from the literature and placed them into categories, and further associated them to the required *security properties* and the potential *data users* (cf. section II and III).

We have elaborated on the required security properties concerning the various data categories. For instance, detected anomalies are most likely not related to any privacy (P.) sensitive information or requirement for the fulfillment of non-repudiation (N.). Detecting anomalies, e.g., by an anomaly-based intrusion detection system, usually do not include the need to prove source origin. However, the data stored related to anomaly events might contain sensitive information and must be trusted and available to forensic investigations, i.e., fulfill the C.I.A. properties.

Another example is the use of actuators, concerning, e.g., braking, steering, and throttle control, that can be privacy sensitive due to their connection to individual driving patterns. However, the signal that initiates the actuator response does not need the confidentiality (C.) property but must be authentic (I.) and secured against, e.g., replay attacks; thus, the signal source requires (N.). Therefore, we consider I.A.N.P. as enough for this data category. Although requirements need evaluation on a case-to-case basis, our mapping to security properties can indicate and guide the reasoning when securing forensic data.

The mapping of forensic data categories to data users is based on the potential value for stakeholders. For instance, resilience techniques, such as responses to malfunction and cyber security

issues, are most likely to interest the vehicle manufacturer (VM.) and law enforcement (LE.); for the former to solve potential vulnerabilities and for the latter to investigate possible crimes. In contrast, the value of data from such events is minimal for the vehicle driver (VD.) and the insurance company (IC.).

We do not provide a complete list of data, and we know that there can be overlaps. Nevertheless, the identified data in Table III, in conjunction with the categorization of the current work in Table I and II can help developers and architects to make informed decisions on data collection concerning automotive digital forensics.

VII. Discussion

An interesting observation from Figs. 4 and 5 is that research publications are distributed in various databases. Thus, combining many database searches with snowballing improves the coverage of the work. Google Scholar has the broadest coverage of the selected papers and, at the same time, the lowest specificity. Therefore, Google Scholar required the most effort compared to the other databases. SCOPUS required the least effort but at the cost of less coverage.

The low overlap was surprising and its potential consequence on coverage, e.g., without Google Scholar, the number of selected papers would have decreased by a quarter. IEEE Xplore seems redundant with only one unique publication. However, the impact of one paper can be significant both when it comes to contribution and later in the snowballing process; thus, we still consider the use of multiple databases critical for coverage.

Data collection is part of the digital forensic acquisition and, as shown in Table I and II, is logically included in a majority of the studies. Most technical solutions consider C.I.A., but only two consider C.I.A.N.P., while some do not consider security properties at all. Blockchain is relatively common in technical solutions in comparison to, e.g., solutions utilizing Trusted Execution Environments (TEE) and virtualization. The latter two are both promising approaches due to their potential to isolate the execution of sensitive processes such as cryptographic operations, e.g., hash generation, signing, and encryption/decryption. Few solutions propose forensics tools and extraction techniques, which align with challenges such as no standardized data formats and no guidelines or standards for the automotive domain concerning automotive digital forensics. Automotive digital forensics is a rather immature area, and we believe that security and privacy need to be emphasized more in the automotive literature.

By adding forensics mechanisms to vehicles, such as increased data collection concerning individual driving patterns, biometrics, and whereabouts, the potential to solve crimes increases. However, we face the risk that a vast amount of sensitive data will be stored in various locations. As mentioned in the introduction (cf. Section I-C), several potentially exploitable vulnerabilities in vehicles can be assumed, with the risk of compromising this data. Security mechanisms that protect such data need to be strengthened, and individual control of this data established. Still, there is always a risk that data can be misused even when the retrieval as such is authorized. For example, car rental or insurance companies can use data from, e.g., individual driving patterns to decide customers' prices for their services according to risk profiling. The ethical discussion of where to put the bar between data collection and the user's right to privacy is out of scope in this work and something for politicians and lawmakers to find common ground.

Lack of proposals for TEE and virtualization solutions, lack of usable tools, and lack of possibilities for data extraction require additional research effort. Essential observations from Table III are that data related to, e.g., *sensors*, *ECUs*, and *safety* are considered valuable forensic data. In contrast, data from *biometrics* and *settings* are only mentioned in one study. Security events from detected anomalies and executed resilience techniques are only considered in seven papers. Thus, more research on detecting, storing, and extracting forensically relevant events such as these are required.

Vehicle security mechanisms, e.g., IDS and firewalls, perform actions and can detect relevant events for digital forensic investigations. Examples of such events can be anomalies in traffic patterns during cyber attacks, successful and failed authentication attempts, and port scans. Such events must be securely stored together with information about the logical order w.r.t. to time. For instance, the airbag might be deployed due to a cyber attack causing a crash. Thus, establishing the order of events is imperative for the investigation to find the root cause of the crash. Moreover, abnormal activation of lane assist function, opening/closing doors, increased volume of music, and activated windscreen wiper before a fatal accident may provide digital evidence of a crime, e.g., an attack intended to distract the driver, potentially leading to an accident.

Data related to vehicle-to-everything (V2X) communication will continue to increase and, thus, be more prominent in future vehicles. Communication between vehicles and infrastructure, roadside units, other vehicles, and pedestrians can impact a specific accident. Therefore, location data and events between vehicles and pedestrians are relevant and important to synchronize to correlate and connect different incidents. Moreover, it is essential to differentiate between faults such as safety-related malfunctions (e.g., component failures) and cyber security issues.

VIII. Conclusion

We have performed a systematic literature review in the field of automotive digital forensics. We performed our searches in four major databases and performed backward and forward snowballing to maximize the coverage. Two core categories were identified for the selected work, namely *Technical Solutions* and *Surveys*. Additionally, 11 categories, some with sub-categories, were identified from the contents of the papers, categories to which all papers were then mapped. Moreover, technical solutions from the selected papers were linked to the security properties they cover. We have further identified and categorized relevant forensic data types derived from the selected papers and linked them to security properties and data users. We have identified and discussed challenges, issues, and research gaps within the area of automotive digital forensics.

The use of a well-known and standardized approach, SLR, gives confidence that the essential papers in the searched

databases are found. Thus, it should not be necessary to repeat this work by practitioners or researchers to find relevant publications from this time period. Still, the SLR approach makes it possible to repeat the work in the future to follow the development of the area. The categorization gives a comprehensive overview of the forensics field and related research activities and makes it easy to find relevant papers in a particular sub-field of digital forensics. The number of papers in the categories also indicates what research areas have been considered important and challenging during the studied time period. The comparison of the search results from four large databases is interesting since it shows how much additional value searches in multiple databases may give. It is also useful to know to what extent they overlap and which ones to focus on when searching for publications since it might also apply to other areas. Our contributions are helpful not only for automotive digital forensics but also for similar systems, such as cyber-physical systems and smart cities.

In summary, our performed analysis guides further work within the area and benefits both researchers and practitioners. The identified forensic data categories can be used to indicate relevant data types to look for within investigations. The data categories can be used in conjunction with the identified technical solutions to serve as a guideline for implementing forensic mechanisms into vehicular and similar systems. Thus, it provides both tools and techniques for the data collection aligned with the data types to consider.

REFERENCES

[1] R. Pallierer and B. Schmelz, "Combine AUTOSAR standards for high-performance in-car computers," Accessed: Sep. 20, 2021. [Online]. Available: https://innovation-destination.com/2017/12/13/combine-autosar-standards-high-performance-car-computers/

[2] M. A. Rahim, M. A. Rahman, M. Rahman, A. T. Asyhari, M. Z. A. Bhuiyan, and D. Ramasamy, "Evolution of IoT-enabled connectivity and applications in automotive industry: A review," *Veh. Commun.*, vol. 27, 2021, Art. no. 100285. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214209620300565

[3] S. Sharma and B. Kaushik, "A survey on Internet of Vehicles: Applications, security issues, & solutions," *Veh. Commun.*, vol. 20, 2019, Art. no. 100182.

[4] M. Stoyanova, Y. Nikoloudakis, S. Panagiotakis, E. Pallis, and E. K. Markakis, "A survey on the Internet of Things (IoT) forensics: Challenges, approaches, and open issues," *IEEE Commun. Surv. Tut.*, vol. 22, no. 2, pp. 1191–1221, Apr.–Jun. 2020.

[5] Official Journal of the European Union, "Directive 2010/40/EU of the european parliament and of the council," Accessed: Feb. 2, 2010. [Online]. Available: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:207:0001:0013:EN:PDF/

[6] B. Nelsons, A. Philips, and C. Steuart, *Guide to Computer Forensics and Investigations*. Boston, MA, USA: Cengage, 2018.

[7] A. MacDermott, T. Baker, P. Buck, F. Iqbal, and Q. Shi, "The Internet of Things: Challenges and considerations for cybercrime investigations and digital forensics," *Int. J. Digit. Crime Forensics*, vol. 12, pp. 1–13, Jun. 2019.

[8] ENISA, "Is software more vulnerable today?," Accessed: Feb. 3, 2018. [Online]. Available: https://www.enisa.europa.eu/publications/info-notes/is-software-more-vulnerable-today

[9] T. Llansó and M. McNeil, "Estimating software vulnerability counts in the context of cyber risk assessments," in *Proc. 51st Hawaii Int. Conf. Syst. Sci.*, 2018.

[10] C. Miller and C. Valasek, "Remote exploitation of an unaltered passenger vehicle," *Black Hat USA*, 2015. [Online]. Available: https://illmatics.com/Remote\%20Car\%20Hacking.pdf

[11] S. Checkoway et al., "Comprehensive experimental analyses of automotive attack surfaces," in *Proc. USENIX Secur. Symp.*, 2011, pp. 77–92.

[12] K. Koscher et al., "Experimental security analysis of a modern automobile," in *Proc. IEEE Symp. Secur. Privacy*, 2010, pp. 447–462.

[13] E. Khanapuri, V. V. T. K. Chintalapati, R. Sharma, and R. Gerdes, "Learning based longitudinal vehicle platooning threat detection, identification and mitigation," *IEEE Trans. Intell. Veh.*, 2021, doi: 10.1109/TIV.2021.3122144.

[14] Reuter, "Uber, distracted backup driver cited by NTSB in fatal self-driving crash," Accessed: Feb. 2, 2021. [Online]. Available: https://www.reuters.com/article/us-uber-crash/ntsb-cites-uber-distracted-backup-driver-in-fatal-self-driving-crash-idUSKBN1XT2IL

[15] Reuters, "Exclusive dutch forensic lab says it has decoded Tesla's driving data," Accessed: Jun. 10, 2022. [Online]. Available: https://www.reuters.com/business/autos-transportation/dutch-forensic-lab-says-it-has-decoded-teslas-driving-data-2021-10-21/

[16] K. Strandberg, T. Rosenstatter, R. Jolak, N. Nowdehi, and T. Olovsson, "Resilient shield: Reinforcing the resilience of vehicles against security threats," in *Proc. IEEE 93rd Veh. Technol. Conf.*, 2021, pp. 1–7.

[17] United Nations, "UN regulation no 155," Accessed: Jun. 8, 2022. [Online]. Available: https://unece.org/sites/default/files/2021-03/R155e.pdf

[18] X. Feng, E. S. Dawam, and D. Li, "Autonomous vehicles' forensics in smart cities," in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov.*, 2019, pp. 1688–1694.

[19] N. Vinzenz and T. Eggendorfer, "Proposal for a secure forensic data storage," *J. Cyber Secur. Mobility*, vol. 9, no. 3, pp. 469–88, 2020.

[20] C. Huang, R. Lu, and K. R. Choo, "Vehicular fog computing: Architecture, use case, and security and forensic challenges," *IEEE Commun. Mag.*, vol. 55, no. 11, pp. 105–111, Nov. 2017.

[21] Proton Technologies AG, "Complete guide to GDPR compliance," Accessed: Nov. 11, 2020. [Online]. Available: https://gdpr.eu/

[22] Electronic Privacy Information Center, "The drivers privacy protection act (DPPA) and the privacy of your state motor vehicle record," Accessed: Nov. 17, 2020. [Online]. Available: https://epic.org/privacy/drivers/

[23] Electronic Privacy Information Center, "Automobile event data recorders (black boxes) and privacy," Accessed: Nov. 17, 2021. [Online]. Available: https://epic.org/privacy/edrs/

[24] M. Cebe, E. Erdin, K. Akkaya, H. Aksu, and S. Uluagac, "Block4forensic: An integrated lightweight blockchain framework for forensics applications of connected vehicles," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 50–57, Oct. 2018.

[25] H. Mansor, K. Markantonakisy, R. Akramz, K. Mayesx, and I. Gurulian, "Log your car: The non-invasive vehicle forensics," in *Proc. IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun.*, 2016, pp. 974–982.

[26] K. K. Gomez Buquerin, C. Corbett, and H.-J. Hof, "A generalized approach to automotive forensics," *Forensic Sci. Int.: Digit. Investigation*, vol. 36, 2021, Art. no. 301111. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666281721000056

[27] N. I. of Standards and Technology, "Guidelines on mobile device forensics," Accessed: Feb. 2, 2021. [Online]. Available: https://csrc.nist.gov/publications/detail/sp/800-72/final

[28] National Institute of Standards and Technology, "Guide to integrating forensic techniques into incident response," Accessed: Oct. 6, 2020. [Online]. Available: https://csrc.nist.gov/publications/detail/sp/800-86/final

[29] International Organization for Standardization, "Information technology – security techniques – guidelines for identification, collection, acquisition and preservation of digital evidence," Accessed: Oct. 6, 2020. [Online]. Available: https://www.iso.org/standard/44381.html

[30] National Institute of Standards and Technology, "Guidelines on mobile device forensics," Accessed: Feb. 1, 2021. [Online]. Available: https://www.nist.gov/publications/guidelines-mobile-device-forensics

[31] D. Jacobs, K. R. Choo, M. Kechadi, and N. Le-Khac, "Volkswagen car entertainment system forensics," in *Proc. IEEE Trustcom/BigDataSE/ICESS*, Piscataway, NJ, USA: IEEE, 2017, pp. 699–705.

[32] M. Hossain, R. Hasan, and S. Zawoad, "Trust-IoV: A trustworthy forensic investigation framework for the Internet of Vehicles (IoV)," in *Proc. IEEE Int. Congr. Internet Things*, 2017, pp. 25–32.

[33] X. Wang, Y. Zhou, X. Ma, N. Lu, N. Cheng, and K. Zhang, "Smart cyber forensics of rear-end collision based on multi-access edge computing," in *Proc. IEEE/CIC Int. Conf. Commun. China*, 2019, pp. 1012–1017.

[34] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *Proc. 18th Int. Conf. Eval. Assessment Softw. Eng.*, 2014, pp. 1–10.

[35] C. Oham, R. A. Michelin, R. Jurdak, S. S. Kanhere, and S. Jha, "B-FERL: Blockchain based framework for securing smart vehicles," *Inf. Process. Manage.*, vol. 58, no. 1, 2021, Art. no. 102426.

[36] C. Yoon, J. Hwang, M. Cho, and B. G. Lee, "Study on DID application methods for blockchain-based traffic forensic data," *Appl. Sci.*, vol. 11, no. 3, 2021, Art. no. 1268.

[37] C. Alexakos, C. Katsini, K. Votis, A. Lalas, D. Tzovaras, and D. Serpanos, "Enabling digital forensics readiness for Internet of Vehicles," *Trans. Res. Procedia*, vol. 52, pp. 339–346, 2021.

[38] M. Waltereit, M. Uphoff, P. Zdankin, V. Matkovic, and T. Weis, "A digital forensic approach for optimizing the investigation of hit-and-run accidents," in *Digital Forensics and Cyber Crime*, S. Goel, P. Gladyshev, D. Johnson, M. Pourzandi, and S. Majumdar, Eds. Berlin, Germany: Springer, 2021, pp. 204–223.

[39] P. A. Abhay, N. V. Jishnu, K. T. Meenakshi, P. S. Yaswanth, and A. O. Philip, "Auto block IoT: A forensics framework for connected vehicles," *J. Phys.: Conf. Ser.*, vol. 1911, no. 1, May 2021, Art. no. 012002. [Online]. Available: https://doi.org/10.1088/1742-6596/1911/1/012002

[40] M. A. Hoque and R. Hasan, "Avguard: A forensic investigation framework for autonomous vehicles," in *Proc. IEEE Int. Conf. Commun.*, 2021, pp. 1–6.

[41] J. Daily, M. DiSogra, and D. Van, "Chip and board level digital forensics of cummins heavy vehicle event data recorders," *SAE Int. J. Adv. Curr. Pract. Mobility*, vol. 2, no. 4, pp. 2374–2388, 2020.

[42] P. Sharma, U. Siddanagaiah, and G. Kul, "Towards an AI-based after-collision forensic analysis protocol for autonomous vehicles," in *Proc. IEEE Secur. Privacy Workshops*, 2020, pp. 240–243.

[43] H. Guo, W. Li, M. Nejad, and C. C. Shen, "Proof-of-event recording system for autonomous vehicles: A blockchain-based solution," *IEEE Access*, vol. 8, pp. 182776–182786, 2020.

[44] A. Philip and R. Saravanaguru, "Secure incident & evidence management framework (SIEMF) for Internet of Vehicles using deep learning and blockchain," *Open Comput. Sci.*, vol. 10, Nov. 2020, Art. no. 408.

[45] M. Li, J. Weng, J.-N. Liu, X. Lin, and C. Obimbo, "Towards vehicular digital forensics from decentralized trust: An accountable, privacy-preservation, and secure realization," *IEEE Internet of Things J.*, vol. 9, no. 9, pp. 7009–7024, May 2022.

[46] Z. Ma, M. Jiang, and W. Huang, "Trusted forensics scheme based on digital watermark algorithm in intelligent VANET," *Neural Comput. Appl.*, vol. 32, pp. 1665–1678, Mar. 2020.

[47] A. Mehrish, P. Singh, P. Jain, A. V. Subramanyam, and M. Kankanhalli, "Egocentric analysis of dash-cam videos for vehicle forensics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 9, pp. 3000–3014, Sep. 2020.

[48] M. Waltereit and T. Weis, "An approach to exonerate innocent suspects in hit-and-run accidents via route reconstruction," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, 2019, pp. 447–448.

[49] K. Bahirat, N. Vaishnav, S. Sukumaran, and B. Prabhakaran, "Addfar: Attacked driving dataset for forensics analysis and research," in *Proc. 10th ACM Multimedia Syst. Conf.*, New York, NY, USA, 2019, pp. 243–248. [Online]. Available: https://doi.org/10.1145/3304109.3325817

[50] L. Cintron, S. Graham, D. Hodson, and B. Mullins, "Modeling liability data collection systems for intelligent transportation infrastructure using hyperledger fabric," in *Proc Crit. Infrastructure Protection XIII*, J. Staggs and S. Shenoi, Eds. Berlin, Germany: Springer, 2019, pp. 137–156.

[51] S. Lee, W. Cho, H. J. Jo, and D. H. Lee, "T-box: A forensics-enabled trusted automotive data recording method," *IEEE Access*, vol. 7, pp. 49738–49755, 2019.

[52] L. Davi, D. Hatebur, M. Heisel, and R. Wirtz, "Combining safety and security in autonomous cars using blockchain technologies," in *Computer Safety, Reliability, and Security*, A. Romanovsky, E. Troubitsyna, I. Gashi, E. Schoitsch, and F. Bitsch Eds. Cham, Switzerland: Springer, 2019, pp. 223–234.

[53] D. Billard and B. Bartolomei, "Digital forensics and privacy-by-design: Example in a blockchain-based dynamic navigation system," in *Proc. Privacy Technol. Policy*, M. Naldi, G. F. Italiano, K. Rannenberg, M. Medina, and A. Bourka Eds., 2019, pp. 151–160.

[54] M. Ugwu, C. Oham, I. Nwakanma, and O. Izunna, "A tiered blockchain framework for vehicular forensics," *Int. J. Netw. Secur. Appl.*, vol. 10, pp. 25–34, Sep. 2018.

[55] M. Marchetti and D. Stabili, "Read: Reverse engineering of automotive data frames," *IEEE Trans. Inf. Forensics Secur.*, vol. 14, no. 4, pp. 1083–1097, Apr. 2019.

[56] H. Guo, E. Meamari, and C. Shen, "Blockchain-inspired event recording system for autonomous vehicles," in *Proc. 1st IEEE Int. Conf. Hot Inf.-Centric Netw.*, 2018, pp. 218–222.

[57] R. Hussain et al., "Secure and privacy-aware incentives-based witness service in social Internet of Vehicles clouds," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2441–2448, Aug. 2018.

[58] A. Mehrish, A. V. Subramanyam, and M. Kankanhalli, "Multimedia signatures for vehicle forensics," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2017, pp. 685–690.

[59] X. Feng, E. S. Dawam, and S. Amin, "A new digital forensics model of smart city automated vehicles," in *Proc. IEEE Int. Conf. Internet Things*, 2017, pp. 274–279.

[60] A. D. Sathe and V. D. Deshmukh, "Advance vehicle-road interaction and vehicle monitoring system using smart phone applications," in *Proc. Online Int. Conf. Green Eng. Technol.*, 2016, pp. 1–6.

[61] N. Watthanawisuth, T. Lomas, and A. Tuantranont, "Wireless black box using MEMS accelerometer and GPS tracking for accidental monitoring of vehicles," in *Proc. IEEE-EMBS Int. Conf. Biomed. Health Informat.*, 2012, pp. 847–850.

[62] D. K. Nilsson and U. E. Larson, "Conducting forensic investigations of cyber attacks on automobile in-vehicle networks," in *Proc. Int. Conf. Forensics Inst. Comput. Sci., Soc.-Inform. Telecommun. Eng.*, 2008, pp. 1–6.

[63] A. Attenberger, "Data sources for information extraction in automotive forensics," in *Proc. Int. Conf. Comput. Aided Syst. Theory*, R. Moreno-Díaz, F. Pichler, and A. Quesada-Arencibia Eds., 2020, pp. 137–144.

[64] R. Rak and D. Kopencova, "Actual issues of modern digital vehicle forensic," *Internet Things Cloud Comput.*, vol. 8, pp. 12–16, Mar. 2020.

[65] D. Kopencova and R. Rak, "Issues of vehicle digital forensics," in *Proc. XII Int. Sci.- Tech. Conf. Automot. Saf.*, 2020, pp. 1–6.

[66] H. S. Lallie, "Dashcam forensics: A preliminary analysis of 7 dashcam devices," *Forensic Sci. Int.: Digit. Investigation*, vol. 33, 2020, Art. no. 200910.

[67] K. Dološ, C. Meyer, A. Attenberger, and J. Steinberger, "Driver identification using in-vehicle digital data in the forensic context of a hit and run accident," *Forensic Sci. Int.: Digit. Investigation*, vol. 35, 2020, Art. no. 301090.

[68] N.-A. Le-Khac, D. Jacobs, J. Nijhoff, K. Bertens, and K.-K. R. Choo, "Smart vehicle forensics: Challenges and case study," *Future Gener. Comput. Syst.*, vol. 109, pp. 500–510, 2020.

[69] D. Steiner, L. Chen, D. Hayes, and N.-A. Le-Khac, "Vehicle communication within networks -investigation and analysis approach: A case study," *Annu. ADFSL Conf. Digit. Forensics, Secur. Law* May 2019.

[70] D. Sladović, D. Topolčić, K. Hausknecht, and G. Sirovatka, "Investigating modern cars," in *Proc. 42nd Int. Conv. Inf. Commun. Technol., Electron. Microelectronics*, 2019, pp. 1159–1164.

[71] N. Vinzenz and T. Eggendorfer, "Forensic investigations in vehicle data stores," in *Proc 3rd Central Eur. Cybersecur. Conf.*, New York, NY, USA, 2019, pp. 1–6. [Online]. Available: https://doi.org/10.1145/3360664.3360665

[72] M. Hussain, M. Beg, M. Alam, and S. Laskar, "Big data analytics platforms for electric vehicle integration in transport oriented smart cities: Computing platforms for platforms for electric vehicle integration in smart cities," *Int. J. Digit. Crime Forensics*, vol. 11, pp. 23–42, 2019.

[73] C. Urquhart, X. Bellekens, C. Tachtatzis, R. Atkinson, H. Hindy, and A. Seeam, "Cyber-security internals of a Skoda octavia VRS: A hands on approach," *IEEE Access*, vol. 7, pp. 146057–146069, 2019.

[74] C. J. Whelan, J. Sammons, B. McManus, and T. Fenger, "Retrieval of infotainment system artifacts from vehicles using iVe," *J. Appl. Digit. Evidence*, vol. 1, no. 1, 2018, Art. no. 30.

[75] A. Koch, R. Altschaffel, S. Kiltz, M. Hildebrandt, and J. Dittmann, "Exploring the processing of personal data in modern vehicles - A proposal of a testbed for explorative research to achieve transparency for privacy and security," in *Proc. 11th Int. Conf. IT Secur. Incident Manage. IT Forensics*, 2018, pp. 15–26.

[76] S. Tatjana, B. Ištvan, T. Nena, and K. Biljana, "Application of digital forensics in traffic conditions," in *Proc. 23rd Int. Sci.-Professional Conf. Inf. Technol.*, 2018, pp. 1–4.

[77] F. Leuzzi, E. Del Signore, and R. Ferranti, "Towards a Pervasive and Predictive Traffic Police," in *Traffic Mining Applied to Police Activities*, F. Leuzzi and S. Ferilli Eds. Cham, Switzerland:Springer, 2018, pp. 19–35.

[78] C. Ivan, P. Dragan, P. Marko, and H. Sinisa, "Application possibilities of digital forensic procedures in vehicle telematics systems," *Zeszyty Naukowe Wyższej Szkoły Technicznej w Katowicach*, vol. 10, pp. 133–144, 2018. [Online]. Available: https://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-99e384a5-27b4-4c1a-9c61-f9b63a138a27/c/Art._10.pdf

[79] Z. A. Baig et al., "Future challenges for smart cities: Cyber-security and digital forensics," *Digit. Investigation*, vol. 22, pp. 3–13, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1742287617300579
[80] R. Altschaffel, K. Lamshöft, S. Kiltz, and J. Dittmann, "A survey on open automotive forensics," in *Proc. 11th Int. Conf. Emerg. Secur. Inf.*, 2017.
[81] W. Bortles, S. McDonough, C. Smith, and M. Stogsdill, "An introduction to the forensic acquisition of passenger vehicle infotainment and telematics systems data," SAE Tech. Paper 2017-01-1437, 2017, doi: 10.4271/2017-01-1437.
[82] J. Lacroix, K. El-Khatib, and R. Akalu, "Vehicular digital forensics: What does my vehicle know about me?," in *Proc. 6th ACM Symp. Develop. Anal. Intell. Veh. Netw. Appl.*, New York, NY, USA, 2016, pp. 59–66. [Online]. Available: https://doi.org/10.1145/2989275.2989282
[83] J. S. Ogden and M. Martonovich, "Forensic engineering tools and analysis of heavy vehicle event data recorders (HVEDRs)," *J. Nat. Acad. Forensic Eng.*, vol. 33, no. 2, Jan. pp. 33–49, 2016.
[84] N. Krishnamurthy and J. H. Hansen, "Car noise verification and applications," *Int. J. Speech Technol.*, vol. 17, no. 2, pp. 167–181, Jun. 2014.
[85] D.-W. Park, "Forensic analysis technique of car black box," *Int. J. Softw. Eng. Appl.*, vol. 8, no. 11, pp. 1–10, Jan. 2014.
[86] K.-S. Lim, C. Lee, J. Park, and S. Lee, "Test-driven forensic analysis of satellite automotive navigation systems," *J. Intell. Manuf.*, vol. 25, pp. 329–338, Jun. 2014.
[87] J. Johnson, J. Daily, and A. Kongs, "On the digital forensics of heavy truck electronic control modules," *SAE Int. J. Commercial Veh.*, vol. 7, no. 1, pp. 72–88, Apr. 2014. [Online]. Available: https://doi.org/10.4271/2014-01-0495
[88] T. Hoppe, S. Kuhlmann, S. Kiltz, and J. Dittmann, "IT-forensic automotive investigations on the example of route reconstruction on automotive system and communication data," in *Proc. Int. Conf. Comput. Saf., Rel., Secur.*, 2012, vol. 7612, pp. 125–136.
[89] S. Al-Kuwari and S. D. Wolthusen, "On the feasibility of carrying out live real-time forensics for modern intelligent vehicles," in *Proc Forensics Telecommun., Informat., Multimedia*, X. D. Lai Gu, B. Jin, Y. Wang, and H. Li Eds., 2011, pp. 207–223.
[90] D. K. Nilsson and U. E. Larson, "Conducting forensic investigations of cyber attacks on automobile in-vehicle networks," in *Proc. 1st Int. Conf. Forensic Appl. Techn. Telecommun., Inf., Multimedia Workshop*, 2008, pp. 1–6.
[91] J. S. Daily, N. Singleton, B. Downing, and G. W. Manes, "Light vehicle event data recorder forensics," in *Proc Adv. Comput. Informat. Sci. Eng.*, 2008, pp. 172–177.
[92] D. K. Nilsson and U. E. Larson, "Combining physical and digital evidence in vehicle environments," in *Proc. 3rd Int. Workshop Systematic Approaches Digit. Forensic Eng.*, 2008, pp. 10–14.
[93] Alientech, "Tools for ECU remapping," Accessed: Feb. 9, 2022. [Online]. Available: https://www.alientech-tools.com/
[94] P&E Microcomputer Systems Inc., "Easily manage i.MX RT secure boot for production programming," Accessed: Feb. 9, 2022. [Online]. Available: https://www.pemicro.com/
[95] Berla Corporation, "iVe 3.5," Accessed: Feb. 9, 2022. [Online]. Available: https://berla.co/
[96] CVE, "CVE-2018-11478," Accessed: Feb. 16, 2022. [Online]. Available: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2018-11478
[97] CVE List, "CVE-2019-12797," Accessed: Feb. 16, 2022. [Online]. Available: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2019-12797
[98] Argus Cyber Security, "A remote attack on the Bosch Drivelog connector dongle," Accessed: Feb. 16, 2022. [Online]. Available: https://argus-sec.com/remote-attack-bosch-drivelog-connector-dongle/
[99] S. Woo, H. J. Jo, and D. H. Lee, "A practical wireless attack on the connected car and security protocol for in-vehicle can," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 993–1006, Apr. 2015.
[100] CVE List, "CVE-2019-9493," Accessed: Feb. 2, 2022. [Online]. Available: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2019-9493
[101] CVE List, "CVE-2018-11477," Accessed: Feb. 16, 2021. [Online]. Available: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2018-11477
[102] A. Palanca1, E. Evenchick, F. Maggi, and S. Zanero, "A stealth, selective, link-layer denial-of-service attack against automotive networks," in *Proc. Int. Conf. Detection Intrusions Malware, Vulnerability Assessment*, 2017, pp. 185–206.
[103] CVE List, "CVE-2016-2354," Accessed: Feb. 16, 2021. [Online]. Available: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2016-2354
[104] Volvo Car Corporation, "Volvo on call," Accessed: Jul. 3, 2021. [Online]. Available: https://www.volvocars.com/intl/v/volvo-cars-app
[105] OnStar Corporation, "Welcome to onStar," Accessed: Jul. 3, 2021. [Online]. Available: https://www.onstar.com/us/en/home/
[106] Hacking and Countermeasure Research Lab, "Driving dataset," Accessed: May 18, 2021. [Online]. Available: https://ocslab.hksecurity.net/Datasets/driving-dataset
[107] CASE, "An international standard supporting automated combination, validation, and analysis of cyber-investigation information," Accessed: Feb. 7, 2021. [Online]. Available: https://caseontology.org/
[108] B. H. Bloom, "Space/time trade-offs in hash coding with allowable errors," *Commun. ACM*, vol. 13, no. 7, pp. 422–426, Jul. 1970. [Online]. Available: https://doi.org/10.1145/362686.362692
[109] A. Valjarevic and H. Venter, "A harmonized process model for digital forensic investigation readiness," in *Proc. Adv. Digit. Forensics IX*, G. Peterson and S. Shenoi, Eds., 2013, pp. 67–82.
[110] CORDIS, "A novel adaptive cybersecurity framework for the Internet-of-Vehicles," Accessed: Apr. 21, 2021. [Online]. Available: https://cordis.europa.eu/project/id/833742
[111] S. Kiltz, J. Dittmann, and C. Vielhauer, "Supporting forensic design - A course profile to teach forensics," in *Proc. 9th Int. Conf. IT Secur. Incident Manage. IT Forensics*, 2015, pp. 85–95.
[112] K. Chae, D. Kim, S. Jung, J. Choi, and S. Jung, "Evidence collecting system from car black boxes," in *Proc. 7th IEEE Consum. Commun. Netw. Conf.*, 2010, pp. 1–2.
[113] B. Canis and D. R. Peterman, "'Black boxes' in passenger vehicles: Policy Issues," Accessed: Nov. 17, 2020. [Online]. Available: https://fas.org/sgp/crs/misc/R43651.pdf,
[114] C. Patsakis and A. Solanas, "Privacy-aware event data recorders: Cryptography meets the automotive industry again," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 122–128, Dec. 2013.

Kim Strandberg (Member, IEEE) is a senior cyber security engineer at the Department of Research and Development at Volvo Cars and an industrial Ph.D. student in the Department of Computer Science and Engineering at the Chalmers University of Technology, Gothenburg, Sweden. He has a licentiate degree of engineering in automotive cyber security, two B.Sc. and two M.Sc. in computer science and engineering with specializations in software engineering, computer systems, networks, and cyber security. He has been working as an engineer within the IT area for 17 years, including automotive cyber security, for around seven years. His main research field is automotive cyber security, with an emphasis on secure and resilient automotive system design and development.

Nasser Nowdehi is an automotive cybersecurity expert and information security officer at Volvo AB, Gothenburg, Sweden. He has a Ph.D. degree in automotive cybersecurity from Chalmers University of Technology and a M.Sc. degree in computer systems and networks (specialized in cybersecurity) from the same university. His main research interests include intrusion detection systems, V2X security, and cyber-resilient systems.

Tomas Olovsson (Member, IEEE) is an Associate Professor at the Department of Computer Science and Engineering at Chalmers University of Technology, Gothenburg, Sweden. He has been working actively with computer security for more than 25 years, both in the industry and in the academia. His research interests are communications and security and he is currently focusing on security and privacy for Internet-connected vehicles. This work includes secure internal network architectures and secure V2X communications.