



LEEDS
BECKETT
UNIVERSITY

Citation:

Bakhshi, T and Ghita, B and Kuzminykh, I (2023) SAFE: a Standardized Automotive Forensic Engine for Law Enforcement Agencies. In: 2023 15th International Conference on Innovations in Information Technology (IIT), 14-15 Nov 2023, Al Ain, United Arab Emirates. DOI: <https://doi.org/10.1109/iit59782.2023.10366498>

Link to Leeds Beckett Repository record:

<https://eprints.leedsbeckett.ac.uk/id/eprint/10888/>

Document Version:

Conference or Workshop Item (Accepted Version)

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

The aim of the Leeds Beckett Repository is to provide open access to our research, as required by funder policies and permitted by publishers and copyright law.

The Leeds Beckett repository holds a wide range of publications, each of which has been checked for copyright and the relevant embargo period has been applied by the Research Services team.

We operate on a standard take-down policy. If you are the author or publisher of an output and you would like it removed from the repository, please [contact us](#) and we will investigate on a case-by-case basis.

Each thesis in the repository has been cleared where necessary by the author for third party copyright. If you would like a thesis to be removed from the repository or believe there is an issue with copyright, please contact us on openaccess@leedsbeckett.ac.uk and we will investigate on a case-by-case basis.

SAFE: a Standardized Automotive Forensic Engine for Law Enforcement Agencies

Taimur Bakhshi
Department of Computing
Leeds Beckett University,
Leeds, United Kingdom
t.bakhshi@ieee.org

Bogdan Ghita
Department of Computing
University of Plymouth
Plymouth, United Kingdom
bogdan.ghita@plymouth.ac.uk

Ievgeniia Kuzminykh
Department of Informatics
King's College London
London, United Kingdom
ievgeniia.kuzminykh@kcl.ac.uk

Abstract— Modern automotive infotainment systems offer a substantial source of evidence for digital forensic practitioners. However, due to lack of guidance and supporting validation tools, forensic analysts struggle with both data acquisition, analysis and reporting. There are general digital forensic frameworks and legislative acts that can be applied to automotive forensics. However, processing vehicles may prove challenging due to analysis of proprietary automotive systems and on-site crime scene dynamics, including cross-functional investigation with physical forensic teams. To gain an insight into emerging challenges, the present work surveyed current automotive forensics practices across law enforcement agencies (LEAs) in the EU, NA and AP region. The result of this survey enabled a qualitative evaluation, exposing an overall limited capability along with prevalence of invasive data retrieval methods and lack of standardized investigation trajectories. Based on this evaluation, a predominant set of recommendations were derived and streamlined in SAFE: a standardized automotive forensic engine. SAFE utilizes preliminary information from the crime-scene and presents a best-practice step-by-step investigation guide for front-line vehicle forensic analysts. The engine captures analyst rating on the validity of each investigation trajectory and KNN-based content filtering is employed to improve future recommendations. SAFE, therefore, aims to optimize vehicle forensic processing from initial crime scene to the courtroom.

Keywords—automotive forensics, digital forensics, vehicle forensics, machine learning, law enforcement.

I. INTRODUCTION

Traditional forensic science is associated with collection and analysis of physical crime-scene samples. Advancements in computing and networking has generated a new branch of forensic science – digital forensics – involved with the identification, collection, and analysis of digital evidence to assist criminal investigations as well as civil proceedings [1]. Recent technological progress has increased the capabilities of most digital devices in terms of storing and processing data, making digital forensic analysis an evolving challenge [2][3]. Automotive, particularly telematic and infotainment systems in vehicles provide an emerging source of useful forensic data including navigation history, connectivity with peripheral devices (such as mobile phones) and driving events. Current vehicle and infotainment systems may, however, employ proprietary and multiple data interfaces, resulting in challenges for the forensic analyst to logically retrieve data artefacts from the system [1][3][4]. At present, there is a lack of commercial software tools available for vehicle forensics due to several reasons such as lack of standardization, limited extraction options, lack of support from manufacturers, and the complexity and diversity of the systems [1][2][4]. The present work summarizes different approaches to vehicle forensics practiced across the globe and creates a best practice recommendation system to assist digital forensic practitioners

based on current capabilities. A survey-based approach is followed to ascertain the skills, resources, and challenges faced by LEAs in the UK, EU, US and AP with respect to vehicle forensics using qualitative measures appraising operational efficiency, evidence collection, chain of custody validation and presentation. Based on the survey results, a comprehensive set of best practice guidelines are analyzed and collated. The present work further realizes these recommendations in SAFE: a standardized automotive forensic engine to help streamline and add resilience to forensic processing of vehicles (and inherent) systems. SAFE captures initial input, the crime scene and vehicle-specific parameters from forensic analysts to recommend a step-by-step investigation trajectory detailing the process(es), tools and technologies to be followed during investigation. A KNN-based machine learning plugin is used to rate the appropriateness of each investigative step based on feedback from the forensic analyst. The recommendation engine, hence, self-improves keeping a score of previous practitioner-rating for each suggested investigation trajectory, and optimizes future recommendations to the forensic analyst. The rest of this paper is organized as follows. Section II provides background information on vehicular forensics, overviewing existing frameworks and related work. Section III discusses the methodology used for survey design, sample selection, data collection and recording. Section IV analyses the results from a qualitative perspective and presents SAFE. Section V provides final recommendations, summarizing the challenges and future work in this domain. Final conclusions are drawn in section VI.

II. BACKGROUND

A. Automotive Forensics

Automotive computing systems are challenging to analyze due to legislative and technical (artefact acquisition) limitations briefly overviewed as follows.

- **Legislative Inadequacies:** Legislative acts directly regulating vehicle forensics are relatively non-existent. The processing of a car, for example, from crime scene, protecting the integrity of data, transportation of vehicles, and overwriting data (by aggregation of non-driving related incidents) is not specifically addressed in existing models. General digital forensic frameworks are therefore, used including ACPO Good Practice Guide for Digital Evidence, ISO 17020 (a accreditation for crime scene), ISO 17025 (general requirements for the competence of testing and calibration laboratories), ISO 27037 (guidelines for identification, collection, acquisition and preservation of digital evidence), and the Police and Criminal Evidence (PACE) Act 1984 [5][6]. The Scientific Working Group on Digital Evidence (SWGDE) [7] published the *Best Practices for Vehicle Infotainment and Telematics Systems* however, the

document is not legally binding, provides limited insight needed for data acquisition, analysis and infrequent updates (despite increasing sophistication of infotainment systems).

Technical Limitations: Vehicle assets of forensic interest can be broadly classified into telematics and infotainment systems. Telematics collate type and serial number of the vehicle, navigation, connected devices, driving information and non-driving related events, as illustrated in Fig. 1. Telematics allow over-the-network communication and limit data deletion except for personal information. Telematic analysis is useful in establishing driving patterns and can be utilized for forensic investigations. Data on infotainment systems on the other hand can be easily accessed, and is relatively voluminous. Stored data may include device make and mode, IMEI, web history, call logs, text messages and media. QNX, Linux (autograde), Android (automotive) and Windows-based systems, are quite popular OS, along with several proprietary systems [1-4]. Non-invasive and invasive approaches can be used for data extraction, with the former employing data downloads from the upstream cloud, or connection of diagnostic software through OBD port, Wi-Fi, or Bluetooth. Invasive methods require unit removal from dashboard, which is relatively complicated. Only a limited number of commercial software platforms allow extraction of breadcrumb data using mainstream digital forensic tools in the US and EU markets [2][4][8].

B. Related Work

Recent publications discuss challenges in vehicle forensic examination [2][4]. Le-Khac *et al.* [1] demonstrated on-going issues pertaining to volatile data loss, data integrity, lack of forensic software and accuracy of data analysis. The authors analysed a VW Golf FAT32 file system experimenting with a selection of methods: a non-invasive extraction through OBD II port and an invasive JTAG and chip-off. Researchers recovered data by carving out over seven thousand files but were unable to recover any user-related data of value. Shin *et al.* [4] evaluated forensic analysis approaches for Android Auto and Apple CarPlay, summarizing existing studies primarily employing wired USB interface for data retrieval and developed a wireless infotainment system data extraction tool for Android and Apple chipsets. Bortles *et al.* [10] used Berla iVe and VBOX to investigate a Ford SYNC Gen 2 module and concluding limitations in event recording and recovery using these tools. Buquerin *et al.* [2] outlined vehicle examination methods, including gaps in forensic on-site readiness, data acquisition, analysis and documentation relevant to EU vehicles. Sladović *et al.* [11] successfully applied standard digital forensic processes to vehicle systems using a commercial vendor tool, summarising approaches in preparation, identification, seizure and acquisition, analysis, and reporting. On a separate strand, although researchers have used machine learning (ML) algorithms quite extensively in the context of digital forensics in emerging paradigms such as IoT and cloud technologies. AI applications in vehicle forensics are relatively scarce [2][12][13][14]. ML models may prove advantageous during vehicle forensic investigation and decision-making process and need to be considered as an enabling technology for the automotive forensic practitioner. Given, the limitations of existing approaches, it is imperative to document the challenges from a wider perspective by examining the state-of-the-art in UK-wide and global forensic units.



Fig. 1. Vehicle infotainment and telematics systems – Data Retrieval

III. SURVEY METHODOLOGY

As highlighted earlier, the approaches, tools, methods, techniques, and capabilities may differ across police forces/LEAs globally as well as nationally. To obtain a more comprehensive view of vehicle forensic challenges, a qualitative survey was distributed across widely accessible forensic professional forums and emailed directly to police forces and vehicle examiners in the UK, US, Japan, Australia, New Zealand and several EU LEAs. The survey was designed to gather information on the legislative, regulatory and operational processes being used for a holistic comparison of vehicle forensics capabilities. The survey questionnaire was broadly divided into four domains with open-ended questions summarized as follows.

- (1) **Legislative Acts and Frameworks:** included questions posed at level of standardisation, legislative acts and frameworks allowing comparison of practices between various countries and/or police forces: *legislative acts, standards and frameworks, as well as steps taken from seizure/crime scene until data analysis.*
- (2) **Technical Expertise – Tools & Skills:** questions focused on *tools and methods* being used for vehicle forensics, *methods of extraction*, and *data analysis* and (any) *requirement of additional vendor-specific tools.*
- (3) **Vehicle (Data) Usability:** included queries on successful vehicular data recovery and subsequent usefulness in courts: *common file systems/OS, usefulness of artifacts, browsing data recovery and value* (complete, partial, irrelevant), and *age-factor of vehicles being a limiting factor* in forensics.
- (4) **Open Challenges:** questions and comments focused on deriving commonly faced challenges faced by LEAs requiring further standardization: *possible improvements to process(es), primary challenges (systems/techniques/knowledgebase), legal frameworks, and any open-ended comments.*

The survey was created using Microsoft Forms and distributed across widely accessible media (*Forensic Focus, Reddit, The Vehicle Network podcast*) and emailed directly to police forces and vehicle examiners. To assure anonymity of the respondents, interviews and focus groups were not included in this survey. Consideration was given to qualitative analysis enabling more in-depth analysis (especially) of smaller samples, over statistical analysis. A limitation of this method includes possible low response rate due to relatively time-consuming responses to descriptive questions. Results from the survey are discussed in the next section.

IV. RESULTS & DISCUSSION

The survey was completed by seventy-nine LEA and police forensic analysts from Australia, Japan, New Zealand, UK, United States, and EU. Survey was distributed in two separate instances by researchers at NUCES, Pakistan and

UoP between September '22 and February '23. Responder distribution from each region, and respective role profile is depicted in Fig. 2. A summarization of survey insights pertaining each of the respective domains is depicted in Fig. 3. The responses from each area of the survey are analyzed in the following sub-sections.

A. Legislation and Guidelines

The respondents, specifically from UK mentioned compliance with ACPO guidelines, and PACE Act 1984. Respondents from US, EU, and AP specified digital forensics laws applied in their respective countries as well as using ISO 17020 and, ISO 17025. A few of EU-based respondents (<2%) mentioned knowledge of the SWGDE best practice guide but stated absence of wider adoptability among front-line staff. It was readily observable from responses that, there is **limited legislation** directly regulating vehicle forensics across all regions and an **absence of a universal framework** detailing initial crime-scene processing to courtroom presentation steps.

B. Tools & Skills

Common processes included **vehicle triaging** including VIN ID, vehicle specifics, etc.; **tool assessment** to check data requested by officer-in-charge (OIC) and data retrievable; **preliminary processing** including photographing, manual examination, recording data on screens; and **artifact acquisition** using invasive/ non-invasive techniques (with differential applicability even in the same country).

- **Vehicle triaging:** During triaging, only a small percentage of examiners (~28-30%) received precise information in advance regarding vehicle's infotainment, otherwise they had to wait till physically at the crime scene. The most commented challenges included **limited support from vendors** and **time involved** in researching a vehicle system.

- **Tool assessment:** a (non-exhaustive) list used for data analysis advised by respondents is presented in Fig. 3. All of responses from UK, US, and AP based professionals mentioned Berla [8] as being most popular but having limited support in EU market. Berla is effective in data acquisition, but subsequent interpretation is difficult sometimes resulting in low accuracy. Commonly used applications included FTK Imager, Autopsy and Cellebrite, not being vehicle-specific but able to provide a hex view of stored data and allow data carving. BT Fleet Telematics, VCDS Ross-Tech and Bosch CDR were observed to be non-forensic but occasionally enabling access to valuable forensic data [15][16].

- During **preliminary processing** around ~65% examiners stated they conduct an initial manual examination and run diagnostics prior to forensic work, to identify previous faults and prevent claims of a faulty system later. For network-connected vehicles, data regarding the location(s) of the vehicle could be potentially accessible after the vehicle has been returned, jeopardizing location of forensic/police laboratories or other classified facilities. According to ~85% of responses, proper data sanitization is a viable option, protecting personal data stored on vehicles.

- **Deficient automotive skills** among forensic front-line, was highlighted numerous times. Thatcham training [17] was mentioned by a few, having practical elements on removing and refitting bumpers, SRS airbag and seats – as well as holding a UK level 2/ASE US/Cert III AP auto electrician qualification or equivalent deemed to be advantageous for forensic examiners.

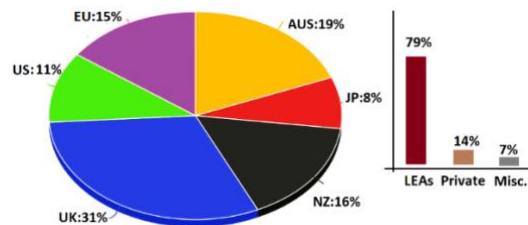


Fig. 2. Survey Responses – Regional and Profile Distribution

C. Data Usability

Most used OS and file systems found in vehicles are provided in Fig. 3. The most useful artifacts to recover are **GPS and location data, connected/paired devices, driving-related events, call logs, text messages, timestamps, and database files**. As outlined earlier, there is minimum support available for these operating systems specifically for vehicles. Web-browser data recovery was limited since only ~16% investigators indulged in this activity. A quarter of respondents stated **age-factor is irrelevant to rule out examination** as aftermarket unit could be installed in the vehicle. However, it was noted by ~11% of participants that vehicles manufactured post-2015 offered best forensic results. Recent ranges from BMW, Vauxhall, Mercedes, Ford, Bugatti, Mitsubishi were investigated by respondents, primarily using Berla for GPS data, logs and messages along with timestamps capturing successfully.

D. Open Challenges and Recommendations

Several suggestions for improvement included: *greater vendor support, decryption tools, and non-invasive imaging techniques*. Lack of *best practice guidelines* was highlighted as a major concern among ~79% of respondents. Comments on creating a *separate legal vehicle forensic framework* were mixed, with the majority negating the need for it, while almost all responses foresaw the need for having a *shared knowledgebase* for automotive forensics. A similar system is being followed by Europol and opening the KB to private practitioners would help the forensic community. Some miscellaneous concerns included retrieval of *vehicle cloud data*, expanding *vehicular forensic tools*, *incentives to enable result validation, pricing, and research facilitation*.

V. RECOMMENDATION ENGINE: SAFE

To cater for an overwhelming requirement of standardized automotive forensic guidelines by the practitioner community, a best practice recommendation engine (RE) is realized in SAFE. SAFE comprises of different knowledge-modules (KMs) employing a series of steps to be followed during investigations. Using ML-enabled content-filtering along with user-rating, the RE extracts an investigation trajectory from the KMs, specific to input parameters provided by the forensic analyst. The step-by-step process guideline aims to facilitate the forensic examiner during **crime scene handling, data acquisition** to subsequent **analysis and reporting**. The working principle is depicted in Fig. 4 and overviewed as follows.

- During initial crime scene processing forensic examiner provides initial information about the vehicle under investigation. Preliminary information (PI) from the crime scene, and vehicle triaging is input into the RE.
- Based on PI, an initial sequence of steps (investigation trajectory) to be followed is extracted from individual KMs (crime scene strategy, evidence transportation, data

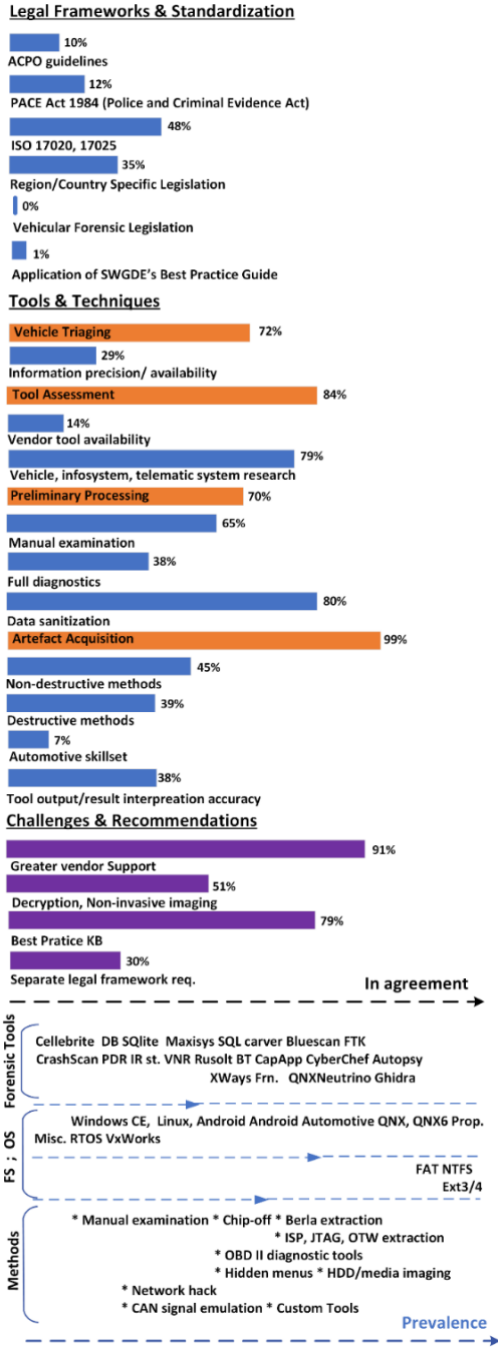


Fig. 3. Survey Insights: legal frameworks, tools & techniques, challenges

- extraction, data interpretation, reporting/ post-investigation conduct) and provided to the practitioner.
- The RE allows forensic practitioner the ability to score the subsequent investigation trajectory at every step of investigation on three equally weighted rating attributes: **(1) validity, (2) viability, (3) overall effectiveness**. The system also solicits and records open-ended comments on any alternate approach not presented for future inclusion in RE (as applicable).
- User-rating information is associated with initial input (via collaborative filtering) and used to train a K-Nearest Neighbor (KNN) classifier [18] at every successive investigation step.

- Subsequent, iterations/utilization, allow scoring of several investigation trajectories, stored in a DB, with the optimal steps advertised to the analyst based on initial user input.

SAFE uses scenario-based collaborative filtering to predict investigation steps that are most appropriate based on crime scene PI's similarity to earlier trajectories and their respective (practitioner) ratings. PI comprises of one-hot encoding of vehicle triaging information, VIN, infotainment and telematic system specifics, and network connectivity, fed to RE. Similarity between two trajectories T , and T' is given by (1).

$$Sim(T, T') = \frac{\sum(r_{Tp} - f_T)(r_{T'p} - f_{T'})}{\sqrt{(\sum(r_{Tp}) - (r_{Ta})^2) \sqrt{(\sum(r_{T'p}) - (r_{T'a})^2)}} \quad (1)$$

where r_{up} , is rating of user u against input item p , and p is all input items (PI) to earlier trajectory (T). T , a function of p is given by (2).

$$T(p) = P\{(VT), (IS), (C)\} \quad (2)$$

VT , IS , and C vector encoding is provided in (3)-(5) below.

$$Veh. Triag. VT = \{(VIN), (CSC)\} \quad (3)$$

$$Info. Telem. Sys ITS = \{(Make), (Model)\} \quad (4)$$

$$Connectivity C = \{(full|partial|unknown|none)\} \quad (5)$$

VIN can take on a range of pre-determined seventeen-character encodings (e.g., a Dodge Durango 2004 assigns a VIN of 1D8HB58D04F177301). These are represented as 24 signed integers in the RE, each integer having a 32-bit datum. Crime scene contamination CSC is accounted using two bits at high, medium, low, undetermined contamination levels. Infotainment system make and model is again translated to a ten-bit binary from pre-set values. Addition of any new model is automatically assigned a value and added in SAFE DB. Connectivity (C) can take on a two-bit binary value representing fully connected, intermittent, undetermined and offline systems. A sample, $T(p)$ for a Dodge Durango (with post-market installations), in a highly contaminated scene, a QNX infotainment system (Neutrino RTOS 7.1), and full Internet connectivity can, therefore, be given by (6).

$$T(p) = P\{(1D8HB58D04F177301, 11) (1011101101), (11)\} \quad (6)$$

KNN calculates the Euclidean distance between target trajectory $T(p)$, and every other in the DB, raking distances and returning the top K nearest neighbour case which is presented to the analyst. SAFE, therefore, continuously improves presented recommendations using KNN plugin trained on practitioners' inputs. This allows community-based KB generation that can lead to a robust investigation process. In present case KNN feature was implemented using Python package (scikit-learn) [19]. Primary modules of the recommendation framework are discussed as follows.

A. Initial Crime Scene Engagement

1) Vehicle Triaging: To allow better preparation, it is appropriate that digital forensic practitioner is contacted by OIC and provided subject vehicle information subject to examination. False plate fitting. Vehicle age, and failed/obsolete system installations need to be considered during triaging.

2) Preliminary Forensics: Priority of digital forensics vs. physical (wet) forensics needs to be ascertained. Wet forensics involves trace(s) preservation, DNA, fingerprinting, and non-digital forensic sources. Matters including compliance (of the

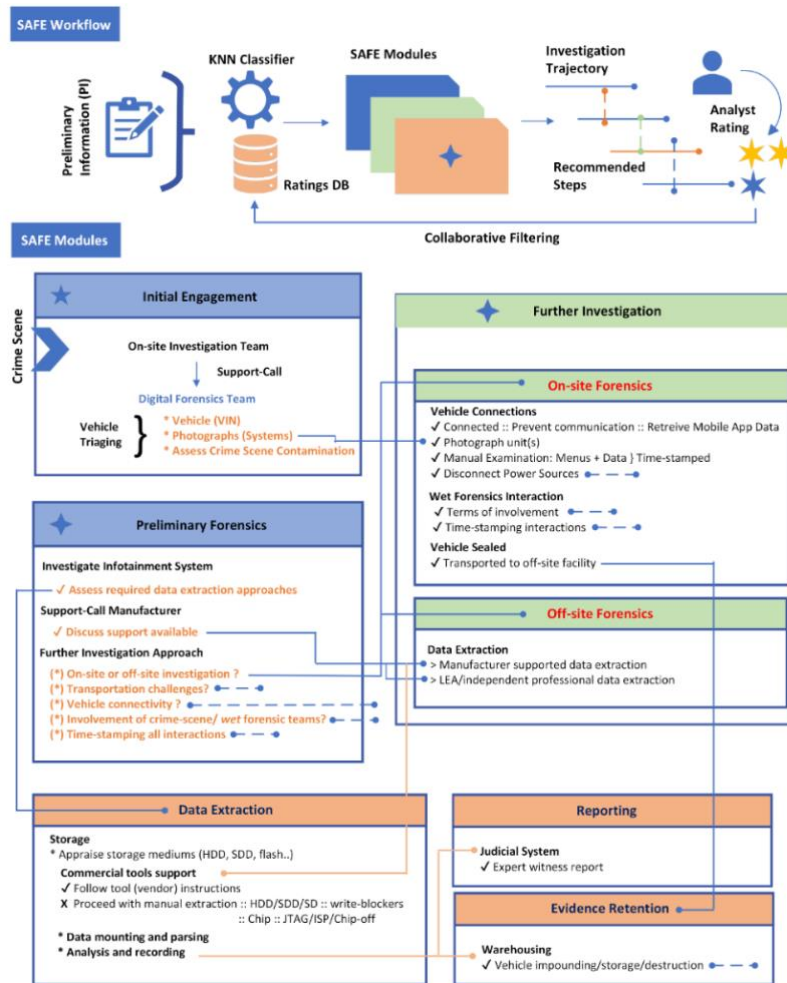


Fig. 4. SAFE Modules: primary and inherent investigation steps

applicable standards), **PPE usage**, **volatile data collection** from screen and on-line/mobile systems, **capturing timestamps**, understanding **availability of OBD ports**, **vehicle drivability**, **prioritization of wet forensics** considering evidence convoluting and **type of extraction** (logical or physical) are recorded and serve as input to subsequent investigative steps. For logical extractions forensics *requires on-scene extraction* – exploring *the use of OBD2*, *autodoctor* [20] and hardware tools. For physical extractions, *disconnecting power source*, *jamming communication*, *running diagnostics*, *deciding on/off-site processing* and *soliciting manufacture support* is best practice.

B. Further Investigation – On-site/Off-Site

Post-initial processing, vehicle must be properly bagged and tagged, including extracting and securing the infotainment unit. Chain of custody must be determined, eliminating any system tampering/updates. If feasible vehicle could be sealed and placed in long term storage facility or private recovery center (to eliminate data integrity issues third party interaction with vehicle systems needs to be curtailed). Post-investigation and before returning vehicles sanitizing or anonymizing system information (GPS, network, etc.) is necessary to prevent storage/LEA/police processing location discovery.

C. Data Extraction and Imaging

Data extraction should start with non-invasive followed by invasive extraction, and manufacturer support as available.

(1) **Manufacturer-assisted Extraction:** Vehicle manufacturers having a trained team member appointed to assist LEAs, with police oversight could be contacted if a forensic practitioner is unable to extract data from system (noting third-party access effects of chain of custody context) and then transported to a secure police facility.

(2) **Autoexaminer/Independent Extraction:** Qualified independent examiners can be used if vendor assistance is unavailable. Again using least invasive and most effective method is important. Usually, hard disk drive or SD cards can be imaged using a validated write-blocker, with additional required for ATA lock or file encryption. In case of chip technology, options include JTAG, in-system programming (ISP), or chip-off (where approved).

(3) **System Mounting:** A few mainstream file systems, such as FAT, NTFS, Ext3 and Ext4, can be easily mounted with minimal support for systems like QNX. After mounting the next step is parsing, using specific tool(s) determined on a file-by-file basis. Evolution of forensic plug-ins for real-time OS (such as X-Ways and Cellebrite), has realized vital vendor support [21][22].

D. Data Interpretation and Analysis

Interpretation of data criticality depends on tool(s) providing validity at binary level. Recommendations involve validating GPS, time stamps, analyzing it, presenting evidential value. Data can aid in developing investigative leads resulting in multiple (evidential) sources. Data to be

used as digital evidence, should always be fully validated, and explainable to the satisfaction of countries' justice system. For data analysis, (in addition to SD cards and HDDs, vehicle systems may also use eMMC chips and incorporate locking systems, such as the CMD42 lock [23]. These eMMC chips comprise of two chipsets: controller and raw NAND flash chip that holds data. Lock is present in the controller of the chip so bypassing the controller can lead to directly accessing data stored on NAND flash memory. To successfully obtain data from an encrypted device, SAFE recommends investigators to explore **physical image extraction** where NAND memory can be useful, **paging layout marking and allocation** setting the borders of the data area, service area and error correction code [25], **inversion** to structure data, and recognize scrambling applied via XOR, inverted, or big/little-endian byte ordering, **block and data management** to place logical blocks in-order, and **mounting and parse file system from a vehicle** to enables data extraction of evidential value.

E. Reporting

If relevant standards and country specific legal acts are adhered to, minimal issues (if any) will arise in admitting expert witness report for civil or criminal litigation. Regarding vehicle forensics, once extraction has been successfully completed, analysis and interpretation can be treated similar to other types of digital evidence (vehicles are usually not presented as evidence in court).

F. Data Sanitization and Evidence Retention

Some fundamental best practices for data sanitization and post-examination/investigation conduct included in SAFE are summarized as follows.

(1) Data sanitisation: Despite careful handling, it may remain unclear if there has been any impact on the unit/system/car due to digital forensic investigation, and tool usage. It is therefore, always a best-practice that when the vehicle is set to be returned, there should be policies in place regulating removal of any such data.

(2) Post-incident conduct: Destructive methods are usually not approved is due to high cost of replacing system after forensic examination. Based, on survey responses, there remains ambiguity on long-term processing of vehicle (e.g. storage, return, impounding, etc.). It would be recommendable to understand/improve local norms pertaining post-incident code on vehicle processing. Using forensic community derived best practices realized in SAFE allows for a first step in standardization and streamlining of automotive forensics applicable in different legislations.

VI. CONCLUSION & FUTURE WORK

The present paper investigated the current capabilities of vehicle forensics processes by surveying and collating responses from forensic practitioners working in LEAs and police forces across different regions. Primary challenges reported an overall lack of standardization, data acquisition challenges, limited vendor support and relatively non-existent shared knowledgebase of best practices. The feedback analysis resulted in formulation of an ML-based modular recommendation engine - SAFE. Using initial crime-scene triaging information, SAFE uses KNN associations to recommend an investigative trajectory detailing step-by-step process(es), tools and techniques to be employed during case preparation. Using practitioner-rated feedback, the investigation steps are continuously optimized to present the best course of action for automotive examiners encompassing

tool selection, data extraction, evidence retention and reporting. Future work will aim to make SAFE publicly available as an online web-based application. This will aid in empirical validation through different case-studies, help analyze the effectiveness of KNN, as well as streamline ethical policies to be considered in using SAFE across different legislative domains and justice systems.

REFERENCES

- [1] Le-Khac et al., "Smart vehicle forensics: Challenges and case study", *Future Generation Computer Systems*, vol 109, 2020, pp 500-510, ISSN 0167-739X, <https://doi.org/10.1016/j.future.2018.05.081>.
- [2] Buquerin et al., "A generalized approach to automotive forensics", *For. Sci. Int., Digital Investigation*, vol 36, Supplement, 2021, 301111, ISSN 2666-2817, <https://doi.org/10.1016/j.fsidi.2021.301111>.
- [3] T. Bakhshi, "Forensic of Things: Revisiting Digital Forensic Investigations in Internet of Things," 2019 4th Int. Conf. on Emerging Trends in Engg., Sci. & Tech, Pakistan, 2019, pp. 1-8, <https://10.1109/ICEEST48626.2019.8981675>.
- [4] Shin Y, Kim S, Jo W, Shon T. Digital Forensic Case Studies for In-Vehicle Infotainment Systems Using Android Auto and Apple CarPlay. *Sensors (Basel)*. 2022 Sep 22;22(19):7196. <https://10.3390/s22197196>.
- [5] ACPO (2012) ACPO Good Practice Guide for Digital Evidence. URL: https://www.digital-detective.net/digital-forensics-documents/ACPO_Good_Practice_Guide_for_Digital_Evidence_v5.pdf
- [6] UK Public General Acts (1994) Police and Criminal Evidence Act. URL: <https://legislation.gov.uk/ukpga/1984/60>
- [7] SWGDE, "SWGDE Best Practices for Vehicle Infotainment and Telematics Systems", version 3.0, 2022. URL: <https://www.swgde.org/documents/published>
- [8] Berla iVe Ecosystem. URL: <https://berla.co/>
- [9] N. Le-Khac, D. Jacobs, J. Nijhoff, K. Bertens, and K.R. Choo, "Smart vehicle forensics: Challenges and case study," *Future Gener. Comput. Syst.*, vol. 109, pp. 500-510, 2020, <https://10.1016/j.future.2018.05.081>
- [10] W. Bortles, S. McDonough, C. Smith, and M. Stogsdill, "An Introduction to the Forensic Acq. of Passenger Veh. Info. and Telem. Sys. Data," *SAE Tech.*,2017-01-1437, 2017, <https://10.4271/2017-01-1437>
- [11] D. Sladović, D. Topolčić, K. Hausknecht, and G. Sirovatka, G. "Investigating Modern Cars," in 2019 42nd Int. Conv. on Info. and Comm. Tech., Electronics and Microelectronics (MIPRO), pp. 1159-1164, 2019, <https://10.23919/MIPRO.2019.8756732>
- [12] Philip et al., "Multisource traffic incident reporting and evidence management in IoV using ML and blockchain", *Engg. App. of AI*, vol 117, 2023, ISSN 0952-1976, <https://doi.org/10.1016/j.engappai.2022.105630>.
- [13] Arumugam C et al., "Digital forensics: essential competencies of cyber-forensics practitioners". In: *Advances in machine learning and computational intelligence*. Springer,2021, Singapore, pp 843–851
- [14] Kemande et al., "Quantifying the need for supervised ML in conducting live forensic analysis of emer. conf. in IoT environments", *Fore. Sci. Int. Reports*, vol 2, 2020, 100122, ISSN 2665-9107.
- [15] Ross-Tech Diagnostic Software for VW-Audi Group Cars, URL: <https://www.ross-tech.com/vag-com/>
- [16] Bosch Diagnostics. URL: <https://cdr.boschdiagnostics.com/cdr/>
- [17] Training with Thatcham research. Available: <https://www.thatcham.org/what-we-do/automotive-academy/>
- [18] Cover, Thomas M.; Hart, Peter E.. "Nearest neighbor pattern classification" (PDF). *IEEE Transactions on Information Theory*. 1967, 13 (1): 21–27. CiteSeerX 10.1.1.68.2616. <https://10.4271/TIT.1967.1053964>.
- [19] Nearest Neighbours, Scikit-learn sklearn.URL: <https://scikit-learn.org/stable/modules/neighbors.html>
- [20] OBD2 Auto Doctor, Access your car's On-Board Diagnostic system. URL: <https://www.obdautodoctor.com/>
- [21] X-Ways Forensics X-Tensions, URL: <http://www.x-ways.net/forensics/x-tensions/>
- [21] Cellebrite Physical Analyzer, URL: <http://cdn5.cellebrite.org/>
- [23] Enabling SD/uSD Card Lock/Unlock Feature in Linux, Technical Note, Micron. URL: <https://media-www.micron.com/-/media/client/global/documents/products/technical-note/sd-cards/>
- [24] Aya Fukami, Radina Stoykova, Zeno Geradts, A new model for forensic data extraction from encrypted mobile devices, *Foren. Sci. Int.: Digital Investigation*, vol 38, 2021, 301169, ISSN 2666-2817.
- [25] Myoungsoo Jung et al., 2012. An evaluation of different page allocation strategies on high-speed SSDs. In *Proc. of 4th USENIX conf. on Hot Topics in Stor. & FS (HotStorage'12)*. USENIX US, 9.