IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 6 March 2024, accepted 31 March 2024, date of publication 5 April 2024, date of current version 16 April 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3385646*

SURVEY

# Categorizing Data-Driven Methods for Test Scenario Generation to Assess Automated Driving Systems

**MAXIMILIAN BÄUMLER, FELIX LINKE, AND GÜNTHER PROKOP**
Chair of Automobile Engineering, Technische Universität Dresden, 01062 Dresden, Germany

Corresponding author: Maximilian Bäumler (mb@mbaeumler.de)

This work was supported by the joint publication funds of Technische Universität (TU) Dresden for the Article Processing Charge (APC), including Carl Gustav Carus Faculty of Medicine, SLUB Dresden, and the Open Access Publication Funding of DFG.

**ABSTRACT** This survey aims to provide an overview of various methods for generating data-driven test scenarios for assessing automated driving systems (ADSs). The survey updates the overall process of scenario generation and categorizes the current methods using a systematic literature review of 64 studies identified between 2017 and 01/2023. Overall, we demonstrate that the data-driven scenario generation process should be updated by another process step, scenario fusion, leading to seven process steps: *1) scope definition*, *2) primary data source selection*, *3) primary data collection*, *4) scenario identification*, *5) scenario fusion*, *6) scenario generation*, and *7) scenario evaluation*. "Scenario fusion" aims to fuse scenarios identified in different data sources for a better coverage of the ADSs' operational design domains (ODDs) and a more comprehensive scenario description. Moreover, we introduce an improved definition for the representativity of test scenario catalogs, which helps improve the collection of traffic data using sampling plans. Also, we show that real driving and police accident data are the most commonly used data input sources. Besides, we illustrate that the ODD is often not defined. Finally, we discuss that the standardization of test scenario generation is difficult because most methods do not address specific ADSs and test environments, and do not provide standardized interfaces. Overall, we recommend comparing existing approaches using the same input data and researching the mutual supplementation of the existing methods. Finally, pre-defined case studies, further standardized terminology, and standards for test execution and evaluation can help speed up the standardization process.

**INDEX TERMS** Advanced driver assistance systems, autonomous driving, system validation, vehicle safety.

## I. INTRODUCTION

The trend in driver assistance systems is increasingly developing towards automated driving systems (ADSs) [1], as the Mercedes-Benz DRIVE PILOT [2] is capable of fully controlling the vehicle in traffic jams up to 60 km/h. In real traffic, the ADS can encounter an infinite number of traffic situations, which complicates the safety validation with every increasing level of automation [3]. Scenario-based testing is a promising method that has attracted increasing attention in recent years [4]. Real traffic situations are taken from the traffic and parameterized into test scenarios, which are applied using real driving or simulative test procedures. These test scenarios can be transferred into a scenario database, which, in the best case, depicts the traffic as representative as possible. The advantages of the scenario-based testing approach are cost and time efficiency as well as high repeatability and flexibility [5], [6]. Although the scenario-based testing approach is often addressed in the literature and established as a test approach in the automotive sector for development and validation [7], a standardized data-driven method for the generation of test scenarios has yet to be established. Various approaches exist for processing the available road traffic accidents and traffic

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

data. Determining a general approach would accelerate the scenario generation process and improve the compatibility of test platforms. To solve this problem, this survey provides an overview of state-of-the-art approaches for data-driven generation of test scenarios from accident and traffic data. The methods reviewed herein use scenarios generated for both development and validation processes. In addition to reviewing the literature on scenario generation methods, the methods are categorized according to five research questions (RQs):

(**RQ1**) How is the term representativity used and defined in the context of scenario generation?
(**RQ2**) Which types of data input sources are used for data-driven scenario generation?
(**RQ3**) Which type of operational design domain (ODD), describing the operating range of one specific automated driving system, is currently addressed by the generated scenarios?
(**RQ4**) What is the most widely discussed scenario generation method?
(**RQ5**) Can data-driven generation of test scenarios from accident and traffic data be standardized?

The remainder of this paper is organized as follows. In Section II, we introduce important definitions and explain the concept of scenario-based testing. In Section III, we explain the systematic literature review process and define the criteria for categorizing the literature. In Section IV, we evaluate and categorize the identified literature. Before concluding the review in Section VI, we discuss the principal results, RQs, limitations, and future research in Section V.

## II. BACKGROUND

The following section provides definitions and descriptions of scenario generation. Moreover, the ASAM OpenX file format, which is used in some methods, is presented.

### A. SCENARIO DEFINITION

A scenario describes "*the temporal relationship between several scenes in a sequence of scenes, with goals and values within a specified situation, influenced by actions and events*" [77, 3.26]. While a scene corresponds to a snapshot of the environment, such as dynamic elements and actors (e.g., persons, objects, and systems), a scenario illustrates the temporal development of the scenes [77]. An action (e.g., turning on car lights) can be a single act/behavior of an actor (e.g., ego-car), whereas an event occurs at a point in time (e.g., traffic light switches to green) [77]. The goals and values depend on the system under test (SuT). A goal of the SuT could, e.g., be to keep the current lane [77], while a value would be to minimize personal injury versus property damage [77].

To differentiate at the abstraction level, Menzel et al. defined functional, logical, and concrete scenarios [78]. The most abstract scenario description is the functional scenario expressed in natural language. Functional scenarios

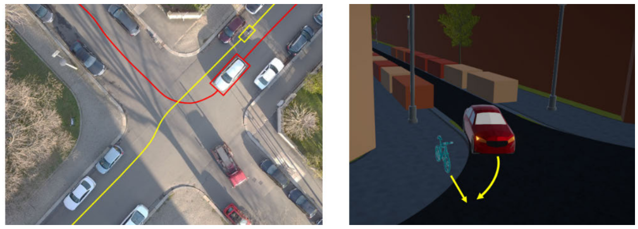

**FIGURE 1.** **Example of a car-bicycle scenario, observed using a camera drone (left) and applied for testing in a stochastic traffic simulation (right). Images by [19] and [79].**

are primarily used in the conceptual phase of the ADS development.

A functional scenario could be a car turning right, while a bicycle next to it wants to go straight at a four-way junction without traffic lights and priority signs (see Figure 1). Less abstract scenario descriptions are logical scenarios that use parameter ranges for scenario entities and their relationships. Logical scenarios are applied primarily during the development phase. Using the functional scenario described previously, parameter ranges for the variables are implemented; for instance, velocity ranges of both road users, for example, 10 to 30 km/h for the car and 5 to 15 km/h for the bicycle, or a road width range from 2.5 to 4 m. Each parameter range must be assigned a specific value to convert logical to concrete scenarios. This description is the least abstract, and represents a uniquely defined test scenario that can be used in real-world tests. The previously described logical scenario would be complemented with specific parameters from the ranges; for example, the velocity of the car is set to 20 km/h and that of the bicycle is set to 13 km/h. The number of scenarios increases considerably with respect to each practical value in the parameter range.

For this purpose, test scenarios worth evaluating must be identified from a large number of concrete scenarios for ADS validation. The test scenarios can be described according to the 6-layer model (6LM) [80], which is an extension of the 5-layer model defined in the Pegasus Project [81]. Layer 1 describes the road network and traffic guidance objects; layer 2 describes the roadside structures; and layer 3 describes the temporal modifications of the first two layers. Layer 4 deals with dynamic objects, such as vehicles and pedestrians, and layer 5 introduces the environmental conditions (weather and wind). Finally, layer 6 illustrates digital information, such as the state of traffic lights and vehicle2X messages.

### B. SCENARIO REPRESENTATIVITY

The current literature does not provide a specific definition of representativity in relation to test scenarios. According to [4], real-world data describing scenes that occur on the road are required to generate representative test scenarios. This ensures a high degree of realism. Based on [82] and [83], a scenario is acclaimed as representative, insofar as it represents the natural population as realistically as possible, thus being area-, time-, and objective-dependent. Reference [84] added that, according to the system's use case, each representative scenario catalog is valid only for a specific safety function or system. Therefore, it is necessary to investigate their corresponding ODD. Furthermore, [5] explains that a more extensive data foundation automatically leads to a higher representativity for the entire scenario dataset, owing to more situations that the SuT may encounter. As each method in this review uses data as a foundation, a fundamental level of representativity exists, depending on the population addressed by the data sampled and used. In addition to the use of representativity in relation to a natural population, another approach is to select the most representative test scenarios or to reduce the number of scenarios from the generated scenario set [84] (see Section IV).

### C. SCENARIO FORMATS (OPENX)

A short overview of the Association for Standardization of Automation and Measuring Systems (ASAM) OpenX file formats (XML) is provided below: OpenSCENARIO [85] defines the entities and their behavior on the road, and models the dynamic part of the scenario. Different actions are triggered by conditions, such as speed, distance, time parameters, and other pre-defined values. Whereas OpenSCENARIO describes traffic participants and their activities in the scenario, the road structure is provided in another OpenX format, called OpenDRIVE [86]. In OpenDRIVE, the road is shaped based on a reference line, with lanes on each side. Features, such as road marks, speed limits, and signals, can also be added. OpenDRIVE is commonly used in various simulation programs, and can be generated virtually or through terrain measurements and mapping. In addition to these two rudimentary, static, and dynamic content-defining formats, ASAM has designed other OpenX simulation standards [87]. Although they are not commonly used for scenario descriptions, a brief explanation follows to ensure completeness. OpenCRG specifies the road structure by assigning values to each cell in a grid on the road. The OpenLABEL format is used for labelling objects and data. Finally, OpenODD, a format that represents the ODD of a safety function, is still in the conceptual phase and needs to be finalized.

### D. SCENARIO-BASED TESTING

To consider the most relevant and valid ADS critical situations, a scenario-based approach is applied. In the testing process, conventional test drives are replaced by scenarios in which irrelevant driving is eliminated, and safety-critical events are considered [82]. This method efficiently optimizes the evaluation process and, thus, the approval of functional safety [88]. Various projects have dealt with scenario-based testing, including Pegasus [89], SePIA [90], Enable S3 [91], SAKURA [92], Headstart [93], SET Level [94], StreetWise [95], MOOVE [96], and L3Pilot [97]. Currently, running projects are V&V Methoden [98], HI-DRIVE [99], V4Safety [100] and SivaS [101]. Moreover, initiatives exist to standardize the testing process and exchange knowledge and scenarios, such as P.E.A.R.S [102], IAMTS [103], and Safety Pool™ [104]. In particular, Safety Pool™ offers a global scenario database containing curated functional and logical scenarios derived from accident and real driving data, as well as expert knowledge, which is accessible on request [104].

The scenario generation methods follow two main approaches [82].

The **knowledge-driven approach** is primarily based on expert opinion, using knowledge and logical linking to define scenarios. Engineers construct scenarios based on their expertise or guidelines related to traffic rules and physical laws; see the example of an urban cut-in scenario parameterized by expert knowledge in [105]. This human factor also contributes to a more comprehensive selection of accidents, from a human perspective [106].

In contrast, the **data-driven approach** derives scenarios primarily from accident, real driving, or simulated/synthetic data sources. Other data sources, such as weather databases, are often linked to enrich the primary data sources. Expert knowledge can, but must not, support a data-driven approach to classify and parametrize scenarios [107].

The general framework for generating scenarios from data is structured as follows. The collected data are stored in a database. Initially, the data is structured and irrelevant parameters are excluded. Different scenario types are identified and then parameterized, often within parameter ranges. An optional step is the clustering and exclusion of scenarios to reduce their number. The choice of data source depends on the evaluation objective or approach. For example, logical scenarios are often required for virtual validation when simulations are used as the assessment tools. Consequently, unprocessed police accident data that only provide functional scenarios are insufficient. This is because the number of variables recorded is usually much smaller in police accident databases than in in-depth accident databases [108]. Moreover, owing to the lack of reconstruction data, the dynamic information (i.e., trajectories with speed and collision courses) of those involved in the accident cannot be automatically retrieved.

The data sources can be divided into different categories. One main group comprises **in-depth accident databases** in addition to accident reports from police data, such as the German In-Depth Accident Study (GIDAS) [108], the China In-Depth Accident Study (CIDAS) [109], the Chinese National Automobile Accident In-depth Investigation System (NAIS), the Shanghai United Road Traffic Safety Scientific Research Center (SHUFO) [110], the Korean In-Depth Accident Study (KIDAS) [111], sub-databases of the American National Automotive Sampling System (NASS) [112], and the Initiative for the Global Harmonisation of Accident Data (IGLAD) [113]. In particular, IGLAD is attempting to harmonize in-depth data from more than 14 partners worldwide and make it available in an in-depth database [114]. **Real**

**driving data** can be used to consider accidents and critical situations, and are often collected using naturalistic driving studies or proving ground data. The most common sensors for collecting real driving data are camera systems, such as drones (see Figure 1), stationary or on-board cameras (see Figure 4), and sensor systems, such as GPS or radar. NDS data, such as the SHRP 2 study [115], often contain large amounts of information and parameters, resulting in large datasets. [107] provides an overview and comparison of different naturalistic driving datasets from 2004 to 2019. A recently published drone dataset, not incorporated in the list in [107], is the ListDB dataset covering three different intersections in Dresden, Germany [116]. Another rarely used data source is **simulated/synthetic data** obtained using simulation tools or driving simulators, which virtually generate information [26], [27], [28], [41], [45], [55].

Overall, the data sources differ in the content and level of information they provide (see Table 1) to describe the layers of the 6LM presented in [80]. The largest differences appear in layer 4, which describes dynamic objects. Although drone-based traffic observations, such as ListDB [116], can deliver the dynamic information of all objects visible in the video in principle, they cannot deliver any information about objects that are not visible to the sensor/camera (e.g., the driver in the car or visibility restrictions for the driver). Thus, information regarding drivers and road traffic participants (e.g., age, sex, driving experience, and vehicle model) remains unknown. In contrast, the police in Germany collect information about the participants involved in a road traffic accident and the conflict situation [117] but do not perform a reconstruction for research. Thus, German police accident data do not contain information on the temporal development of dynamic objects. In contrast, in-depth accident databases such as GIDAS provide information about participants as well as dynamic information via reconstruction [108], [118]. However, GIDAS, for example, only collects road traffic accidents with personal injuries, and not road traffic accidents with property damage [108], [118]. Moreover, NDSs are often limited to the ego-vehicle view [115] and cannot capture the temporal development of all objects surrounding the ego vehicle, as drone-based traffic observations can. Table 1 shows that there is no single data source that can perfectly describe all the layers of the 6LM; however, different data sources lead each other depending on the information required.

## III. METHOD

This review was based on a systematic literature review using the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses" (PRISMA) guidelines [8].

### A. LITERATURE DATABASES

The SCOPUS and IEEE Xplore databases were used in this study. IEEE Xplore covers studies in computer science, electrical engineering, electronics, and other related fields. SCOPUS provides peer-reviewed journals in physical, life, social, and health sciences. The initial search was conducted on 10/24/2022, with an additional search alert for later published articles. The inclusion of articles was stopped on 01/11/2023. Furthermore, some studies were included through side searches, based on recommendations or websites.

### B. SEARCH STRATEGY

First, a search string (see Table 2) was determined according to a pre-search to detect important keywords concerning ADSs and test scenario generation. Often occurring words were noted and categorized into subject areas, such as test concepts, different ADSs, or scenario-related terms. The words that appeared most frequently in each subject area were chosen as keywords for the database search. The components in the first column (see Table 2) represent the focus areas of this study and define keyword blocks (second column). The keywords in each row are combined by the Boolean expression "OR", and each row is assembled by an "AND". The asterisk (*) wildcard represents all continuations of a word, e.g., "generate", "generation", or "generating". Different terms were used to describe ADSs, as shown in the third row of Table 2. Other commonly used synonyms include highly automated driving (HAD), advanced driver assistance systems (ADAS), connected automated vehicles (CAV), and self-driving cars. The combined search string was used in both databases under consideration of their input syntax, and the search was conducted in all metadata, except for the term "scenario", which had to be part of the title.

### C. SEARCH RESULTS

Using the search string derived from Table 2, 781 studies were found in both databases, of which 181 were duplicates (see Figure 2). Another 110 papers were removed because their publication dates were earlier than 2017. The year 2017 was set as the lower limit, with the publication and introduction of the terms functional, logical, and concrete scenario [9] as a major step towards a uniform understanding of scenarios. The titles of the remaining 492 records were screened, resulting in 340 excluded articles, as shown in the "Screening" section of the flowchart in Figure 2. The criteria for inclusion, applied by two of the authors independently and then consolidated, were the direct designation of "scenario" or "generation" in the title and the thematic reference to ADS. Titles theming the evaluation or validation of ADS without naming scenario generation were further investigated in the abstract screening. Data usage was considered in the abstract screening: scenarios should be generated from accident or traffic data. Traffic data can be collected, for example, on roads by performing naturalistic driving studies (NDSs) or by using driving simulators. None of the other generation methods, such as those that primarily rely on ontologies, was investigated in this review. With these two criteria, both the main aspects of the research problem (scenario

**TABLE 1.** Qualitative comparision of four different types of data sources with respect to the first five layers of the 6LM [80], [108], [115], [116], [117], [118] (without synthetic/simulated data sources, which also rely on real world data to some extent). x = information is determinable directly or via post-processing using also additional data sources; X = delivers the best information.

| | **Accident data** | | **Real driving data** | |
|---|---|---|---|---|
| **Example** | **Police accident data** German data [117] | **In-depth accident data** GIDAS [108] | **NDS data** SHRP 2 [115] | **Drone data** ListDB [116] |
| **Main focus** | all road traffic accidents | road traffic accidents with personal injury only | normal-driving behavior | |
| **Period** | continuous collection | | collection in fixed period | |
| **Area / Coverage** | nationwide | regional / specified locations | | |
| **Study design** | total survey | sample | | |
| **Point of view** | reconstructed | | road user based | location based (birds' eye view) |
| **Layer 1:** Road network and traffic guidance objects | x | x | x | x |
| **Layer 2:** Roadside structures | x | x | X | X |
| **Layer 3:** Temporary modifications of L1 and L2 | x | x | x | X |
| **Layer 4:** Temporal development of dynamic objects | none | reconstructed | for all objects covered by ego-vehicle sensors | for all visible objects |
| **Layer 5:** Environmental conditions | x | x | x | x |

**TABLE 2.** Keywords of the search string.

| **Component** (AND) | **Keywords** (OR) |
|---|---|
| Generation | method*, generat*, framework, identif*, recogni* |
| Scenario | [Document title] scenario* |
| Automated Driving | "automated driv*", "highly automated driving", adas, "advanced driver assistance system", ads, "automated driving system", cav, "connected automated vehicle", "autonomous driv*", "self-driving", adf, "automated driving function" |
| Data-based | traffic, accident, data*, nds, "naturalistic driving study", collision, crash, fot, "field operational test", observation, risk, "road traffic accident" |

generation under data reference) are treated. After evaluating the abstracts, 66 studies were removed, and for the following full-text review, 86 titles were sought for retrieval, of which six were not retrieved due to missing access [10], [11], [12], [13], [14], [15]. The exclusion criteria for a more detailed review of the literature were "no scenario generation", "no data reference", "wrong language", and "different topics". The number of excluded papers is shown in Figure 2.

Four additional papers were added to the previously described database search via a side-search. The exclusion process is illustrated in Figure 2. A total of 64 papers were included in this review, of which 61 [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76] original research papers were evaluated in tabular form. The remaining three review-papers are part of the discussion in Section V. As shown in Figure 3, most studies were published in 2021 (18x) and 2022 (19x).

### D. CATEGORIZATION APPROACH

Before categorizing the different scenario generation methods, six categories were deductively determined from the 64 included studies based on the five RQs. The first category comprises the **data used** for the primary dynamics in the scenario described in layer 4 of the 6LM [80]. The data foundation is divided into the subcategories "accident data", "real driving data", and "simulated/ synthetically generated data". Each subcategory is assigned a specific data reference as an abbreviation; for example, Police Accident Data (P) or Dash Cams (DC). For unclear or not detailed descriptions of the recorded real driving data in the respective literature sources, "NDS" is stated, including all types of sensor data.

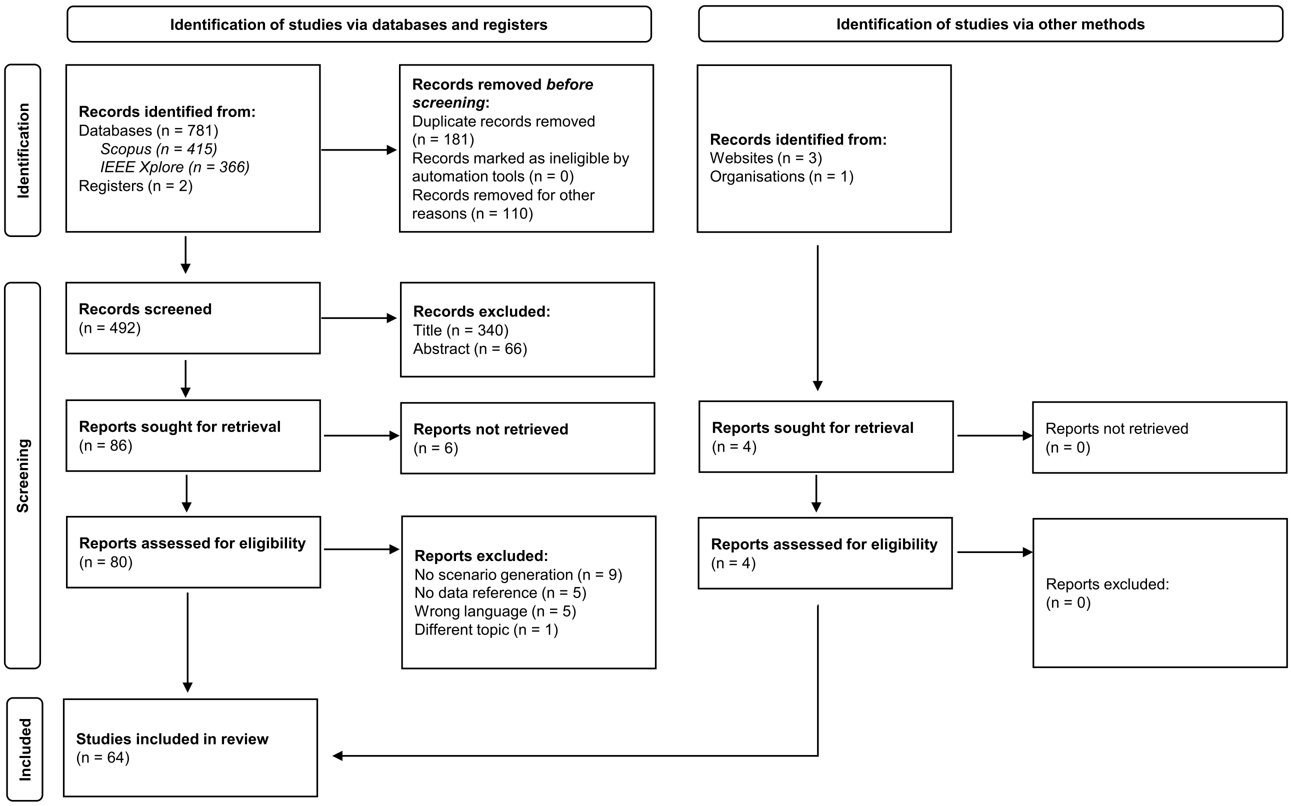


**FIGURE 2.** **Process of the literature search documented using the PRISMA flowchart [8].**

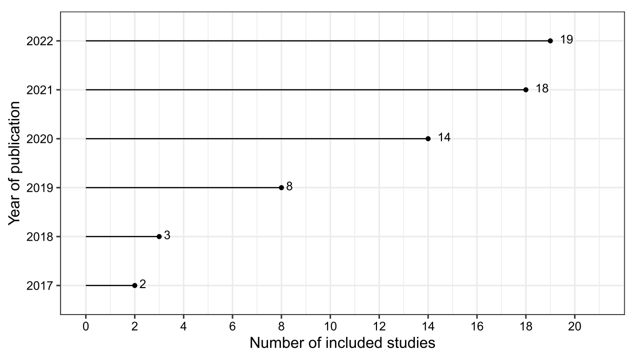


**FIGURE 3.** **Annual distribution of the 64 included studies.**

In contrast to the input data, the next category is the **output** of the scenario generation method. The output is divided into the subcategories of "scenario structure", "trajectories and maneuvers", and "values and conditions". Referring to [78], the first subcategory scenario structure is split into functional, logical, or concrete scenarios; specific file formats; and content-related descriptions or unspecific, not further detailed scenario output. The latter is generated using methods that create a concrete scenario that can be performed on a proving ground or in a simulation. The second subcategory includes trajectories or maneuvers as the output, mostly given as vehicle movement from the top view or with time steps. The last output subcategory includes values and conditions in the form of parameters or pre-crash conditions.

The third category represents the **representativity** of the generated scenarios. Even though a general definition of representativity or its measurement does not exist, representativity is indicated either as RA – representativity regarding the scenario set itself or as RB – representativity related to a natural population.

The fourth category indicates the **point of view** of the data used for scenario generation. The point of view is divided into "road user based" (e.g., on-board camera, see Figure 4), "location based" (e.g., stationary camera, see Figure 1), or "global view" (e.g., traffic flow), and "reconstructed", which is always the case for accident databases as data sources (see Table 1). Multiple viewpoints can be obtained when more than one data input source is used. An example of multiple points of view is the use of stationary, and thus location based cameras, in addition to dynamic parameters (GPS and speed values) recorded in a test vehicle (road user based point of view).

The next category evaluates the **ODD** of the generated scenarios, which is separated into "spatial" and "objective"-related subcategories. In general, the ODD describes the domain and conditions for which the ADS is developed to operate securely, such as the road layout, speed ranges, and environmental conditions [119]. The spatial subcategory encompasses different road sections addressed by the

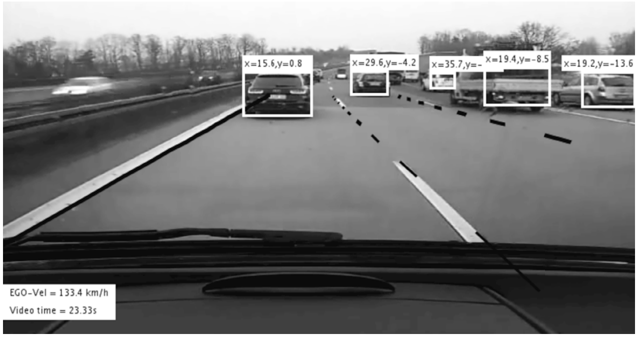


**FIGURE 4.** Example for a road user based view. In addition, other road users are marked with bounding boxes. Image by [83].

generated scenarios, such as "highway only" scenarios. The objective-related subcategory deals mainly with the driving maneuvers addressed by the generation method, such as "cut-in" or "deceleration" maneuvers. For more detailed accident or traffic parameter categories that can be created from clustering, the subcategory "both" (spatial and objective) is available. The term "unspecific" describes when no ODD is indicated.

The last category contains a description of the **scenario space**. The scenario space refers to the search method used for scenario identification. An open scenario space refers to the search for possible patterns in the data foundation. In contrast, a pre-defined scenario space refers to methods that rely on identifying existing patterns, such as accident types or rule-based approaches.

## IV. RESULTS

Before we classify the individual scenario generation approaches into the six categories introduced, we provide an overview of the data-driven scenario generation process based on existing research and literature review results. Finally, we introduce the studies by [53] and [74], building on each other and part of the SAKURA project [92].

### A. PROCESS OF SCENARIO GENERATION

In the following, we classify the scenario generation methods used in the 64 papers into the overall scenario generation process, which is discussed in [31], [120], [121], and [122]. Thereby, we extended the scenario generation process proposed by [31], [120], and [121] by emphasizing the data selection (2) and collection steps (3) and adding the optional fifth step of scenario fusion (see Figure 6). In addition to Figure 6, Table 3 (see the Appendix) shows which study addresses which of the following sub-steps of the scenario generation process:

**The first step** is to define the scope of validation. This includes the definition of the SuTs and their corresponding ODDs. Eleven out of 64 studies specified a SuT, whereby all of them can be assigned to either SAE Level 2 (partial driving automation) or Level 3 (conditional driving automation) [124].

Furthermore, the validation scope encompasses the types of scenarios for which the SuT should be tested – for example: Is the focus on accident, critical, or non-critical scenarios? Scenarios are critical when they contain a potential risk of harm, defined by the scenario likelihood of occurrence and severity of the potential harm [121], [125].

**The second step** is to define the primary data sources covering the defined ODDs and scenario types. Depending on the evaluation scope, different types of data sources, such as accident and real driving databases, may simultaneously be necessary. In the reviewed studies, five methods relied on two different types of primary data sources: police accident and real driving data [19], [46], [75], [76].

**The third step** consists of data collection, which attempts to draw a random sample of the desired scenario types (accident, critical, non-critical, etc.) from the defined ODDs. However, most methods are currently developed on either existing publicly available datasets (35x) or self-collected experimental datasets, without a survey methodology with a sampling design (22x).

**The fourth step** is to identify and extract the scenarios in the collected data using pre-defined scenarios (7x) or scenarios extracted by rule-based (28x), unsupervised (7x) or supervised approaches (5x) – or any combination of them (12x). Pre-defined scenarios [33], [37], [43], [64], [70], [71] are often extracted from accident data because every recorded accident can be considered a scenario. Rule-based approaches (see Table 3) rely on clearly defined parameter ranges or thresholds to identify scenarios and often require extracted road user specific trajectories from, e.g., video material using Convolutional Neural Networks (CNNs) [127]. A CNN is a supervised learning algorithm that requires a training dataset containing classified data. The algorithm learns from existing classes and then assigns them to the unseen data [128]. Applications of CNNs include object detection, scene labelling, and classification [127]. An example of a rule-based approach is to extract scenarios using vectors describing time-dependent traffic scenes, their specific actors (e.g., road users), and their actions [36]. Unsupervised approaches using agglomerative hierarchical [29], k-medoids [55] or entropy-based [30] clustering seek to identify scenarios by identifying patterns in, e.g., past accident data [29], [30], [55], [69]. Unsupervised approaches that leverage Generative Adversarial Neural Networks (GANs) attempt to generate new scenarios derived from a self-supervised learning task [62], [63]. In contrast, supervised approaches attempt to identify scenarios by predicting existing patterns, i.e. given scenario training classes. Examples of supervised prediction tasks used include modified Random Forests [17], Logistic Regressions [28], and Recurrent Neural Networks (RNNs) [40]. Furthermore, combinations of several approaches exist, such as the combination of supervised and unsupervised learning, to identify scenarios [18], [21], [60].

**The fifth step** is the optional step of fusing the scenarios identified in different data sources to maximize their informational content. Thus, the fusion of scenarios

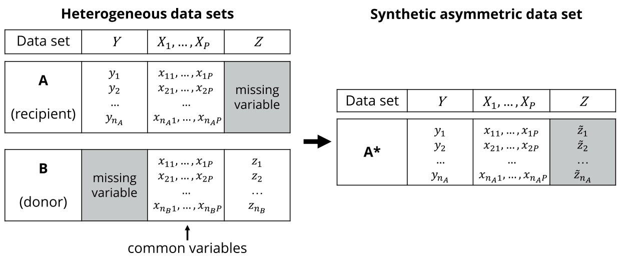


**FIGURE 5.** Concept of asymmetric statistical matching, as shown in [76].

extracted from police accident data with those extracted from video-based traffic observations can make it easier to describe all six scenario layers of the 6LM [80] in detail [19]. For example, ODD-representative test scenarios can be generated by enriching police accident data from an ODD with dynamic information, such as starting speeds and accelerations, determined from the ODD covering video-based traffic observations [19]. This type of fusion allows police accident data to be "parameterized", which in turn helps to determine all relevant accident scenarios in an ODD. While [75] used Record Linkage to fuse scenarios belonging to the same entity across different data sources, [76] tested Statistical Matching, a method of data fusion [126], to fuse scenarios that belong to one superordinate population only, but do not share the same entity in the data sources considered. Thus, with the help of Statistical Matching, ODD-specific scenario data sets **A** with variables $Y$ and $X$, and **B** with variables $X$ and $Z$, extracted from different accident and traffic data sources, which must not have observed the same road users in the ODD, can be fused into a common scenario data set **A***, which consequently contains more information (variables $X$, $Y$, and $Z$) than the respective individual scenario data sets (see Figure 5) [19], [76]. A real application, published after the systematic literature review, can be found in [122].

**The sixth step** involves transferring the identified scenarios into a format in which they can be applied to a test environment (simulation, proving ground, etc.). Thus, if necessary, this step includes the conversion of logical scenarios into concrete and executable scenarios. The most frequently used method is to sample, estimate, or combine parameters (6x) using, e.g., the Metropolis Hastings Algorithm belonging to the Markov Chain Monte Carlo (MCMC) approach [16], [23]. The second most are customized methods (5x) used to derive critical scenarios, using, e.g., the Improved Intelligent Driver Model (IIDM) [45]. Another option is to use reinforcement learning-based methods [31], [32], [72] (3x), or search-based methods (2x), such as genetic algorithms [41].

**The seventh step** is to evaluate the generated scenarios and how they correspond to the desired testing scope (see step 1). Eight methods were used to evaluate the scenarios generated in terms of their criticality using, e.g., risk metrics [16], [23], [31], [32]. Other criticality metrics rely on collision detections/systems that fail [34], maximum vehicle yaw rate [43], or are probability- [53] or expert-based [47]. The methods used to assess the achieved coverage of the targeted scenario space can be explorative using histograms [34]. Moreover, metric-based methods using a specific coverage metric based on scenario classes [66] or search-based methods [48] are used to assess coverage. Other methods also investigate the diversity/similarity (2x) [36], [58], the exposure to the real world [66], and the possibility of testing a cooperative ADS [45] using customized metrics.

Finally, Figure 7 maps the occurrence of steps four to seven of the scenario generation process in the 64 included studies to the publication year. While step four "scenario generation" has already been represented in 2017, steps "scenario generation" and "scenario evaluation" began to be discussed in 2019 and seem to be of increasing interest. The "scenario fusion" method was introduced in 2020 and has rarely been discussed.

### *B. CATEGORIZATION*

After considering the data-driven scenario generation process in general, we mapped the scenario generation approaches [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76] to the six previously defined categories in Section III-D (see Table 4 in the Appendix). A cross in the corresponding cell indicates if the category "output" or "ODD" is not specified.

The data were divided into three subcategories in the **data source** category: accident databases, real driving data, and simulated/synthetic data (Figure 8). Approximately 68% of the data used in the reviewed methods were different types of real driving data that were used as inputs for the generation process. The most common sensors for real driving data are various camera videos or images from drones [17], [18], [19], [20], [45], [48], [50], [65], [74], stationary cameras [19], [20], [42], [53], [62], [63], [71], [76], or dash cams/on-board cameras [26], [36], [43], [46], [51], [62], [63], [75]. Dynamic parameters, such as GPS or velocity, represent a large proportion of the real driving data (NDS and BUS data [16], [20], [21], [22], [24], [25], [27], [31], [32], [39], [44], [49], [51], [52], [53], [57], [58], [59], [66], [67], [72]).

A rarely used data source is a dataset of pre-crafted scenarios or trajectories [20], [23], [68], [73]. Accident databases, as data sources, share approximately 26% of the identified methods, which, in most cases, originate from a set of police accident reports [19], [24], [28], [29], [30], [33], [35], [43], [46], [47], [55], [61], [67], [69], [70], [75], [76]. In some cases, only accidents involving vehicles with an ADS were considered and filtered from existing accident databases [34], [37], [38], [64]. The least used data source was simulated or synthetically created data at approximately 6%. Simulated/synthetic data [40], [41], [42], [54], [67] are

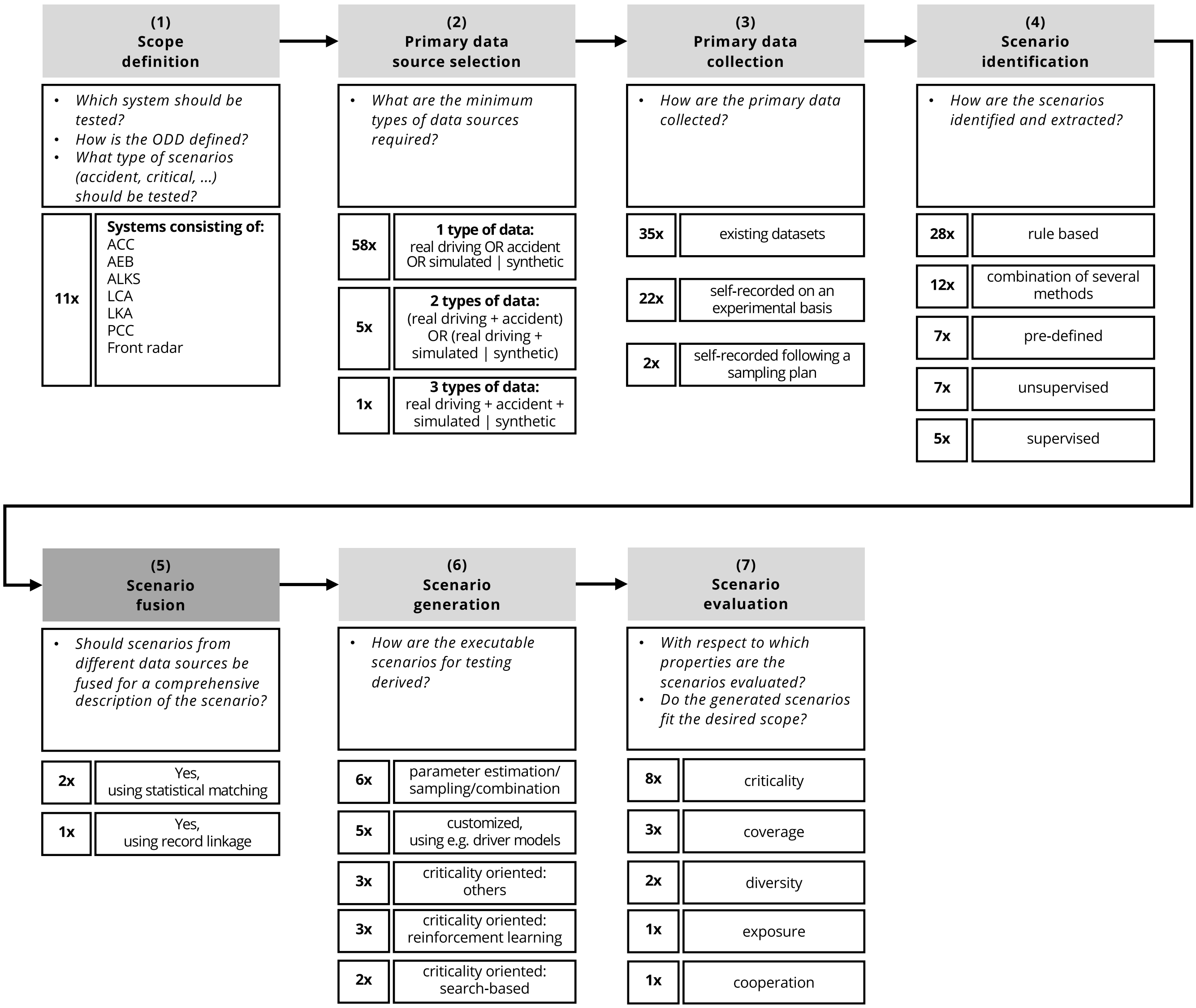


**FIGURE 6.** **Data-driven scenario generation process based on [31], [120], [121], and [122] and adapted by the results of the analysis of the 64 included studies. Each process step lists the number of occurrence in the 64 included studies.**

primarily based on virtual driving simulators or simulation programs.

The categorization of **representativity** is performed between RA and RB, with RA depicting representativity within each scenario set or space, which is the case in 25 generation approaches [16], [23], [24], [29], [30], [31], [32], [40], [41], [42], [43], [44], [45], [47], [48], [53], [55], [60], [65], [66], [67], [68], [69], [72], [74]. To achieve this goal, filtering of the generated scenario set to reduce the number of scenarios is frequently performed. Another possibility represents criticality measures to perform risk analysis or calculate dynamic parameters, such as time-to-collision (TTC) or time headway (THW). A unique example of the RA approach is the minimization of the performance outcome to identify worst-case scenarios and select them for the testing and evaluation of the ADS [72]. The representativity in the natural population is described by RB, which occurs ten times [19], [22], [28], [35], [38], [46], [54], [62], [63], [76].

RB is realized by dividing scenarios into groups based on global or environmental properties, such as different scenario types for different regions or countries, or environmental values, such as weather effects. Overall, the representativity of each scenario set (RA) predominates its counterpart RB, although in several generation methods the term representativity either deviates from RA/RB (4x) or is not specified more precisely (25x, see Figure 9).

The **point of view** will always be from the "reconstructed point of view" when accident databases are used as data sources, which occurred 21 times. Processing sensor data

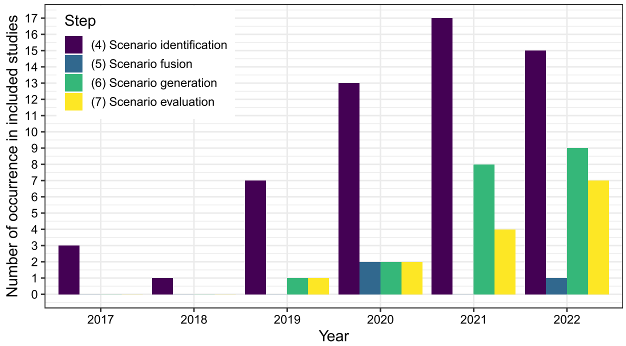


**FIGURE 7.** Distribution of the occurrence of the steps "(4) Scenario identification", "(5) Scenario fusion", "(6) Scenario generation", and "(7) Scenario evaluation" over the publication years of the 64 included studies. Steps 1-3 are not shown separately as they are usually dependent on steps 4-7.

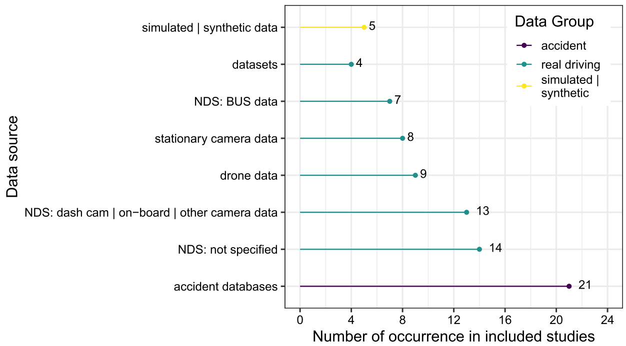


**FIGURE 8.** Distribution of different primary data sources used in the 64 included studies to generate scenarios. The data sources can be divided into three groups: Accident, real driving and simulated/synthetic data. In the real driving data, a further distinction is made between the different sensors used in the case of naturalistic driving studies (NDSs).

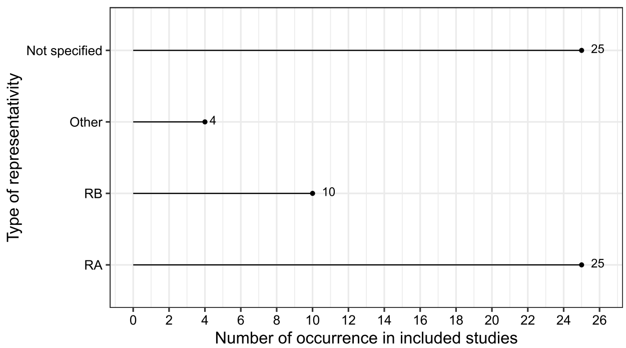


**FIGURE 9.** Distribution of the different types of representativity in the 64 included studies to generate scenarios. RA depicts representativity within each scenario set or space. RB describes the representativity in the natural population addressed by the scenario set/space.

with vehicle-specific dynamic parameters often leads to a road user based view, and camera footage mostly indicates a road user or a location based point of view. The global view is selected in cases that use global parameters such as traffic flow.

The results of the scenario generation process are presented in the **output** section. Eight times, the studies directly

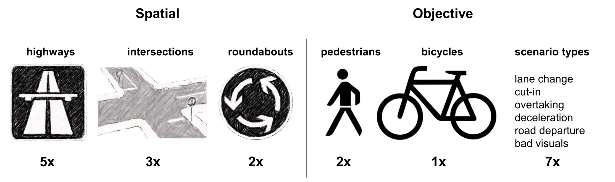


**FIGURE 10.** Distribution of spatial and objective related ODDs in the 64 included studies to generate scenarios (excerpt).

referenced the scenario output as a functional, logical, or concrete scenario structure [78]. Some approaches aim for functional scenarios, whereas others aim for completely parameterized concrete scenarios derived from functional scenarios. In addition, six studies described their output in a more detailed structure, such as OpenSCENARIO [52], [65], and pre-crash matrix [38] formats or matrices with scenario content [64]. A particular case was presented in [62], in which dynamic human poses were evaluated and categorized to generate pedestrian-crossing scenarios. The remaining approaches with direct scenario outputs have yet to state their output format; thus, they are only marked with a cross. Approximately 20% of the reviewed methods deliver trajectories or maneuvers. Approximately 10% of the methods output pre-crash conditions or crash characteristics but do not provide specific scenario content/formats.

The **ODD** was divided into spatially related (18x), objective-related (22x), and both (13x) or not specified (11x). The spatially related ODD is organized into highways, motorways, urban areas, intersections, and roundabouts. Regarding dynamic content, the objective-related ODD is organized into road users and maneuvers. If accident categories or variables, such as the 3-digit accident type, are used, as in [19], then the method can be categorized into both spatial and objective ODD. An overview of the ODDs is presented in Figure 10. Because only half of the reviewed approaches further address their generated scenarios to a specific ODD, the distributions would only be conditionally meaningful and are, therefore, not depicted in charts, but presented schematically.

Almost two-thirds of the scenario generation methods can be assigned to a pre-defined **scenario space** by searching for existing patterns using their algorithms. Repeated algorithms include Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and rule based approaches. The open scenario space appeared 16 times, indicating that the data were processed without any pre-defined categories. The scenario space of the remaining six studies was not addressed in detail [34], [36], [37], [45], [50], [58]. Frequently applied algorithms for scenario detection include CNNs [17], [18], [25], [56], and RNNs [51], [60], [72]. RNNs use the results of the current step as the input for the next step, and may not require labelled data for training. Features are learned from the memory of the previous input, which results in the problem of storing past information over an extended period [129]. Because RNNs predict the most likely result for the next step, they are often used with

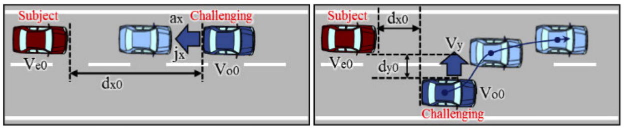


**FIGURE 11.** Deceleration scenario (left) and cut-in scenario (right) generated by [53] and [74]. Image by [74].

sequential data, such as text or video, for natural language processing, speech analysis, or entity extraction [128].

For scenario generation, CNNs are often used to extract the bounding boxes of different objects in video data. Another application of CNNs is the prediction of trajectories. RNNs are used to process the sequential motion information of vehicles and plan the next motion steps.

### C. EXEMPLARY RESULTS

For example, we introduce the connected and well-documented work of [53] and [74] which generated logical cut-in and deceleration scenarios on highways (see Figure 11) from real-world data using a rule-based scenario identification approach, called the "SAKURA approach" [74], [123].

This work is particularly interesting because it not only generates logical scenarios relevant for the validation of automated lane-keeping systems (ALKSs), but also because it uses publicly available data, such as the highD dataset [133], and compares the scenario generation approach for data from two countries (Germany and Japan).

Overall, [74] identified 8,822 deceleration scenarios in a Japanese NDS consisting of 1,047 h of recorded driving on highways (road user based point of view) and 26,846 deceleration scenarios in the German highD dataset (location based point of view). The deceleration scenarios were described by five main parameters (see Figure 11 (left)) aggregated using histograms: subject initial velocity $V_{e0}$, relative initial velocity $V_{e0} - V_{o0}$, initial longitudinal distance $d_{x0}$, maximum deceleration $a_x$ and mean jerk $j_x$ [74].

Moreover, [74] identified 1,561 cut-in scenarios in the Japanese NDS and two other Japanese datasets (one more NDS (350 h) and one location based traffic observation using fixed cameras) as well as 1,017 cut-in scenarios in the German highD dataset. The cut-in scenarios were also described using five main parameters (see Figure 11 (right)) aggregated using histograms, such as $V_{e0}$, $V_{e0} - V_{o0}$, $d_{x0}$, the initial lateral distance $d_{y0}$ and the lateral velocity $V_y$ [74].

To identify the scenarios, the SAKURA approach follows pre-defined rules. In the case of cut-in scenarios, for example, it was determined that the scenario duration should be between $2 - 16$s, the longitudinal velocity of the subject vehicle should be greater than the challenging vehicle's longitudinal velocity, and that $d_{x0}$ should be between $0 - 100$m [53]. Please see [53] for all rules of the applied SAKURA approach.

Overall, [74] showed that although there are differences in the traffic systems between Germany and Japan, there are also significant correlations between the majority of the corresponding scenario parameters. Thus, Japan and Germany can share a common set of test scenarios in this case to some extent [74].

## V. DISCUSSION

The following section presents a comparison with other reviews and discusses the RQs and the general results of the survey. Moreover, limitations and future work are outlined.

### A. DISCUSSION TO OTHER REVIEWS

Various surveys have addressed the topic of scenario generation for the evaluation of ADSs, whereas five surveys can be found in [107] and [121].

Zhang et al. [121] split the data sources for scenario generation into accident data and various types of data such as traffic and sensor data. They also referred to knowledge-based generation, and identified pre-crash situations to create logical and concrete scenarios.

In [130], dynamic scenario generation is divided into four methods: combinatorial testing, knowledge- and driving-behavior-based generation, and data-driven scenario generation. Roads were created independently from dynamic scenario content using field-collection data from OpenStreetMap files or remote sensing imagery.

Riedmaier et al. [131] provide a structured overview of knowledge-based and data-driven approaches. Furthermore, they compared the reviewed approaches systematically with ten evaluation criteria: Scenario Representativeness, Parameter Compatibility, Corner Case Identification, Scenario Space Coverage and Expansion, System Applicability, Computational Feasibility, Black-box Compatibility, Statement Reliability, and Assessment Transferability.

A scenario definition related to ASAM OpenX standards is given in [107]. The scenario data sources were divided into real, simulated, and empirical data. Moreover, the authors compared naturalistic driving datasets including scenario databases.

Cai et al. [120] proposed a framework for a data-driven scenario methodology, including data source collection, scenario identification and generation, and scenario evaluation. In this framework, the scenarios originate from natural driving, accidents, and virtual data. Different generation methods can be divided into diversity- and criticality-oriented generations. Diversity-oriented methods generate scenarios by random sampling or parameter variation to fill the entire ODD, with the drawback of generating "boring" scenarios. Criticality-oriented methods attempt to identify the most critical scenarios by using statistical methods or search algorithms. Additionally, various criticality metrics were created to evaluate the test scenarios.

All these surveys provide a taxonomy for relevant terms in scenario-based testing, for example, scenarios, scenes, or critical and challenging scenarios. They reviewed several approaches to scenario generation. However, we extended the data collection step in the previously proposed scenario generation processes by emphasizing the data selection (2) and collection steps (3) separately. Moreover, we added an optional fifth step of scenario fusion for the first time. Moreover, our survey differentiates the fact that only data-driven scenario generation methods are directly related to traffic and accident data, and provides two comprehensive lookup tables.

Generally, special attention must be paid to the terms "scenario database" and "scenario catalog", which are often used, but not yet specified. Although scenario databases mainly contain a collection of scenarios unrelated to a specific ADS, scenario catalogs are often created for a specific purpose, such as testing one type of ADS.

### B. RESEARCH QUESTIONS AND FUTURE WORK

**RQ1 asked** how the term representativity was used and defined in the context of scenario generation. To date, there is still no precise definition of "representativity" in relation to the coverage of real-world scenarios belonging to a natural population, although the resulting gap between the generated and real-world scenarios has often been discussed. Therefore, we propose the following definition of representativity in relation to a test scenario catalog for assessing specific ADS (SuTs') operating in clearly defined ODDs:

*A test scenario catalog represents reality at a given point in time and is thus representative of the ODD inherent to the SuT addressed by the catalog, if the distribution of the scenarios in the test scenario catalog and in real traffic matches at the given point in time. This is approximately achieved by correctly drawing and processing a sufficiently large random sample of scenarios from the ODD inherent to the SuT.*

This definition explicitly does not address the term "representative test scenario" because a single test scenario can never be representative of the real traffic situation. The term "representative test scenario" fits better to the representativity category RA introduced before. Moreover, owing to the complexity of the real traffic situation, a perfect representativity of a scenario catalog represents an ideal state that can only be approximated by limiting the considered section of the real traffic situation. Therefore, a test scenario catalog should always be designed for a specific SuT that addresses a clearly defined ODD. Furthermore, the data sources used to generate the test scenarios in the test scenario catalog must represent a sample of the population that can be derived from the SuT and its corresponding ODD. When drawing the sample, the number "*of characteristic combinations for each sample may not exceed the population itself, since otherwise the possibility of generalization is not given*" [83, p.368]. In addition, the scenario generation method must not distort the sample because of the arbitrary elimination of test scenarios. Finally, every scenario catalog must be continuously updated to remain representative.

In future research, we propose including the following aspects for every data collection:

- define the ODD and the targeted natural population before data collection starts;
- define standardized reports of sampling plans;
- consider potential requirements from the methods of scenario identification, fusion, and generation already during data collection, e.g., necessary metadata, as shown in [132];
- collect comprehensive metadata in general, for example, time, location, and road users involved, as in [116] and [132], to document the sample drawn.

**RQ2 asked** what types of data input sources were used for the data-driven scenario generation. The most commonly used data input sources are real driving and accident data, mostly from police accident reports. The highD dataset [133], which contains naturalistic vehicle trajectories driven on highways as recorded by drones, has been used in several studies [17], [18], [45], [48], [65], [74]. Moreover, the STATS19 UK accident database [134] was referenced three times [28], [29], [30]. Additionally, simulated data have been used in several approaches. This includes data created in virtual driving simulations with IPG Automotive CarMaker [135] in [54] and ground-truth data from the SVL Simulator [136] in [41].

In future research, we propose to:

- define publicly available, standardized research datasets, which help to compare the results of scenario generation;
- create datasets that share a common population, for example, to align real driving studies and accident data collection in time and space. Thus, the understanding of the scenarios observed increases;
- increase the number of available datasets by, for example, involving citizen volunteers in data collection. Therefore, citizens can help collect video-based traffic observation data ("recording from home") and thus cover as many different locations as possible in a cost-effective manner.

**RQ3 asked** which types of ODDs were currently addressed in the generated scenarios. The ODD in the generated scenarios was organized into spatial (highways, motorways, urban areas, intersections, and roundabouts), objective (pedestrians and bicycles, maneuver types), and both. No ODD has turned out to be mainly addressed. Furthermore, the primary application purpose, either development or validation and testing, is rarely discussed in the compared studies, which is important because the data source and generation approach depend on the chosen application.

In future research, we propose to:

- always define an ODD;
- use standardized ODD description formats, such as OpenODD [87];
- intensify the research focus on ODDs containing intersections and roundabouts, as well as vulnerable road users (bicycles and pedestrians). For example, this implies recording more publicly available datasets using traffic observations in urban areas. Care should be taken to ensure that vulnerable road users can also

be recognized by appropriate algorithms during post-processing, as they represent objects with a size of a few pixels, for example, when observing from a high altitude using drones [132]; a handbook for observing vulnerable road users in general is given by [137].

**RQ4 asked** what the most discussed scenario generation method was. First, no method reviewed addressed all seven steps of the scenario generation process introduced (see Figure 3). However, most of these methods use only one primary data source (58x), often in the form of real-driving data (55x), provided by existing datasets (35x), and apply rule-based approaches (28x) to identify and extract the scenarios. The scenario generation itself (step six), which aims to determine executable scenarios for testing, does not yet show a favorite method; six methods rely on estimating/sampling or combining parameters, and five methods use highly customized approaches relying on, e.g., driver models. Regarding the scenario evaluation (step seven), most methods assess the scenarios regarding their criticality (8x), followed by the coverage of the targeted scenario space (3x). However, slight modifications or data processing differences exist for each method, even among methods with the same data input sources. Evaluation and direct comparison are further aggravated by methods that have neither the same input nor the same output. To choose the ideal approach, one should consider the demands of the generation approach, which is facilitated by the overview and categorization presented in Tables 3 and 4, respectively. The corresponding methods [17], [21], [45], [51], [74] are worth considering for ADS designed for highways. If only drone data are accessible for scenario generation, the appropriate methods can also be filtered from Table 4.

In future research, we propose to:

- extend the research on the use of more than one primary data source ("scenario fusion") to maximize the information available per scenario (see Table 1);
- design standardized interfaces between the seven scenario generation steps to facilitate the exchange of and build on existing research results;
- recognize that combining different data-driven methods is required to best cover the scenario space. In addition, a combination of knowledge-based and data-driven approaches may be reasonable.

**Finally, RQ5 asked** whether data-driven generation of test scenarios from accident and traffic data could be standardized. Currently, standardization of the generation of test scenarios is not possible. The reasons for this are that most of the methods:

- …did not specify the type of ADS they could address.
- …did not specify the type of test environment (simulation, proving ground, etc.), they were applicable.
- …are based on existing datasets, which, in turn, are not yet collected in a standardized manner and thus provide, for example, different amounts of information.

Moreover, various problems may arise, resulting in an excessive number of undefined boundary conditions. One boundary condition is the input format of the data, which includes several parameters or file formats. In addition, the output structure is diverse among the many methods.

Some methods have already used the standardized scenario format OpenScenario, which is a promising way to structure the scenario output. Currently, every research facility or industry seems to use its own method, depending on the available data and their specific needs. This problem is intensified by the availability of different simulation methods and the lack of practical applications for these methods in the development or validation process.

Thus, we propose in future research to:

- compare existing scenario generation approaches based on the same input data ("round robin tests");
- develop methods using pre-defined case studies (given ADS, ODD, data set, validation method) to enable the comparison of methods;
- further standardize the terminology in the field of scenario-based testing;
- define not only standards for scenario generation but also for test execution and evaluation to help both the development of new ADSs and their safety validation;
- extend the research on methods that speed up the scenario generation process, such as dense reinforcement learning, to identify critical scenarios [138].

Last, this survey focused primarily on data-driven generation of test scenarios. Future surveys should take a closer look at the generation of test scenarios using (traffic) simulations.

### C. LIMITATIONS

Limitations arise due to the limited search period ending January 2023. In terms of answering the research questions, there may have been changes since 01/2023, especially with regard to RQ2–RQ4, which are based on quantitative analyses. Moreover, the inclusion and exclusion of studies are always subjective to a certain degree, even in the case of two experts reviewing independently. Furthermore, not all studies describe the methods or data sources used in detail, which makes basic categorization difficult.

Several studies have been conducted on scenario-based testing and scenario generation methods. However, the number of scenarios sufficient for safety validation still needs to be determined, whereby first approaches to solve this challenge exist [139]. Another limitation is the high number of generation methods with slight differences between each other: on the one hand, applying various data processing algorithms, and on the other hand, using different input and output variables.

## VI. CONCLUSION

Scenario-based testing will help validate the safety of ADSs. Based on the 64 studies reviewed, the scenario generation process was divided into seven steps: *(1) scope definition, (2) primary data source selection, (3) primary*

**TABLE 3.** Reviewed studies mapped to the scenario generation process following the seven identified steps.

| Ref | (1) Scope / SuT definition | (2) Primary data source selection | (3) Primary data collection | (4) Scenario identification | (5) Scenario fusion | (6) Scenario generation | (7) Scenario evaluation |
|---|---|---|---|---|---|---|---|
| [16] | - | 1 (real driving) | se | rule based | - | pe/ps/pc | criticality |
| [17] | - | 1 (real driving) | ed | supervised | - | - | - |
| [18] | - | 1 (real driving) | ed | unsupervised; supervised | - | - | - |
| [19] | x | 2 (accident; real driving) | sp | rule based | statistical matching | - | - |
| [20] | - | 1 (real driving) | se | rule based | - | - | - |
| [21] | - | 1 (real driving) | se | unsupervised; supervised | - | - | - |
| [22] | - | 1 (real driving) | ed | rule based | - | - | - |
| [23] | x | 1 (real driving) | ed | rule based; supervised | - | pe/ps/pc | criticality |
| [24] | - | 1 (accident; real driving) | se | rule based; unsupervised | - | - | - |
| [25] | - | 1 (real driving) | se | rule based; supervised | - | - | - |
| [26] | - | 1 (real driving) | se | rule based; supervised | - | - | - |
| [27] | - | 1 (real driving) | se | rule based | - | - | - |
| [28] | - | 1 (accident) | ed | supervised | - | co: others | - |
| [29] | - | 1 (accident) | ed | unsupervised | - | co: others | - |
| [30] | - | 1 (accident) | ed | unsupervised | - | co: others | - |
| [31], [32] | x | 1 (real driving) | ed | rule based | - | co: rl | criticality |
| [33] | - | 1 (accident) | ed | rule based; pre-defined | - | - | - |
| [34] | - | 1 (accident) | ed | rule based | - | pe/ps/pc | criticality; coverage |
| [35] | - | 1 (accident) | ed | rule based | - | - | - |
| [36] | - | 1 (real driving) | ed | rule based; supervised | - | - | diversity |
| [37] | - | 1 (accident) | ed | pre-defined | - | - | - |
| [38] | - | 1 (accident) | ed | unsupervised | - | - | - |
| [39] | x | 1 (real driving) | se | - | - | - | - |
| [40] | - | 1 (simulated \| synthetic) | se | supervised | - | - | - |
| [41] | - | 1 (simulated \| synthetic) | ed | rule based | - | co: sb | - |
| [42] | - | 2 (real driving; simulated \| synthetic) | se | rule based | - | - | - |
| [43] | x | 1 (accident; real driving) | ed | pre-defined | - | cm | criticality |
| [44] | - | 1 (real driving) | se | - | - | - | - |
| [45] | - | 1 (real driving) | ed | rule based | - | cm | cooperation |
| [46] | x | 2 (accident; real driving) | ed | rule based | - | pe/ps/pc | - |
| [47] | - | 1 (accident) | ed | rule based | - | pe/ps/pc | criticality |
| [48] | - | 1 (real driving) | ed | rule based; unsupervised | - | - | coverage |
| [49] | x | 1 (real driving) | se | rule based; unsupervised | - | - | - |
| [50] | - | 1 (real driving) | se | - | - | - | - |
| [51] | - | 1 (real driving) | se | rule based; supervised | - | - | - |
| [52] | - | 1 (real driving) | se | rule based | - | cm | - |
| [53] | - | 1 (real driving) | se | rule based | - | pe/ps/pc | criticality |
| [54] | x | 1 (simulated \| synthetic) | se | rule based | - | - | - |
| [55] | - | 1 (accident) | ed | unsupervised | - | - | - |
| [56] | - | 1 (real driving) | se | supervised | - | - | - |
| [57] | x | 1 (real driving) | se | rule based | - | - | - |
| [58] | - | 1 (real driving) | se | rule based | - | - | diversity |
| [59] | - | 1 (real driving) | se | rule based | - | - | - |

**TABLE 3.** *(Continued.)* **Reviewed studies mapped to the scenario generation process following the seven identified steps.**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| [60] | - | 1 (real driving) | ed | unsupervised; supervised | - | - | - |
| [61] | x | 1 (accident) | se | rule based | - | - | - |
| [62], [63] | - | 1 (real driving) | ed | unsupervised | - | cm | - |
| [64] | - | 1 (accident) | ed | pre-defined | - | - | - |
| [65] | x | 1 (real driving) | ed | rule based | - | - | - |
| [66] | - | 1 (real driving) | ed | - | - | co: sb | coverage; exposure |
| [67] | - | 2 (accident; real driving; simulated \| synthetic) | - | rule based | - | - | - |
| [68] | - | 1 (real driving) | se | rule based | - | cm | criticality |
| [69] | - | 1 (accident) | ed | unsupervised | - | - | - |
| [70] | - | 1 (accident) | sp; ed | pre-defined | - | - | - |
| [71] | - | 1 (real driving) | ed | pre-defined | - | - | - |
| [72] | - | 1 (real driving) | ed | rule based | - | co: rl | - |
| [73] | - | 1 (real driving) | ed | rule based | - | - | - |
| [74] | - | 1 (real driving) | ed | rule based | - | - | - |
| [75] | - | 2 (accident; real driving) | ed; se | rule based; pre-defined | record linkage | - | - |
| [76] | - | 2 (accident; real driving) | ed; se | rule based; pre-defined | statistical matching | - | - |

*data collection, (4) scenario identification, (5) scenario fusion, (6) scenario generation,* and *(7) scenario evaluation*. To date, none of the reviewed methods covers all seven steps. Moreover, most methods do not specify the ADS or ODD that is being addressed. The most frequently used primary data sources were real driving and accident data: 39 studies used real driving data, 21 used processed accident data, and six utilized virtually generated data. The ODD can be divided into spatial- and objective-related ODDs. Spatially related ODDs include five highways, three intersections, and two roundabout scenario generation methods. The objective-related ODDs were further divided into two methods for road user scenarios, one for pedestrian and one for bicycle scenario generation, and seven maneuver-related methods.

Future research should extend the research on the use of more than one primary data source to maximize the information available per scenario and to design standardized interfaces between the seven steps of scenario generation. In addition, new methods should be developed using pre-defined case studies to enable the comparison of methods. Furthermore, the terminology in the field of scenario-based testing must be further standardized. The representativity of the scenarios is another challenge that must be addressed using a precise definition for which the first proposal is given. With this, ADS users can trust the safety of their automated driving vehicles, which is a crucial step in transforming their mobility towards the needs of the future.

## APPENDIX

In the following, all important abbreviations for Tables 3 and 4 in the Appendix are introduced (again) for easier reading.

### A. TABLE 3: REVIEWED STUDIES MAPPED TO THE SCENARIO GENERATION PROCESS

- **(1) Scope / SuT definition**
  SuT: System under test
- **(2) Primary data source selection**
  minimum number of different data source types (data source type)
- **(3) Primary data collection**
  se: self-recorded on an experimental basis
  sp: self-recorded following a sampling plan
  ed: existing dataset
- **(6) Scenario generation**
  pe/ps/pc: parameter estimation/parameter sampling/ parameter combination
  co: criticality oriented
  rl: reinforcement learning
  sb: search-based
  cm: customized (e.g., driver models)

### B. TABLE 4: DETAILED CATEGORIZATION RESULTS OF THE REVIEWED STUDIES

1) **Data source**
   AVA: police-reported automated vehicle accidents
   IGLAD: Initiative for the global harmonization of accident data
   P: police accident data
   BUS: CAN-BUS data
   C: camera data
   D: drone data
   DC: dash cam data
   NDS: naturalistic driving study not specified
   OBC: on-board camera data

**TABLE 4.** Detailed categorization results of the reviewed studies.

| Ref | Data source (primary dynamics) | Representativity | View | Output | ODD | Scenario space |
|---|---|---|---|---|---|---|
| [16] | NDS | RA | road user based | scenario (c) | objective | pre-defined |
| [17] | D | - | location based | scenario (c) | spatial (highway; roundabout) | open; pre-defined (CNN) |
| [18] | D | - | global; reconstructed | scenario | - | pre-defined (CNN) |
| [19] | P; D; SC | RB | location based; global; reconstructed | trajectories/ maneuvers | both (3-digit accident type) | pre-defined |
| [20] | NDS; SC; D; TD | - | road user based; location based; global | trajectories/ maneuvers | - | pre-defined |
| [21] | NDS | - | road user based | trajectories/ maneuvers | spatial (highway); objective (OTM) | pre-defined |
| [22] | NDS | RB | road user based; global | trajectories/ maneuvers; values/ conditions | both | open |
| [23] | SPMD | RA | road user based; location based | scenario | objective | pre-defined |
| [24] | P; BUS; C | RA | location based; reconstructed | scenario (f; l; c) | both (information tables) | open |
| [25] | BUS | - | road user based | scenario | - | pre-defined (CNN) |
| [26] | OBC | - | road user based | trajectories/ maneuvers | objective (CIM; LCM) | pre-defined |
| [27] | BUS | - | road user based | scenario | - | pre-defined |
| [28] | P | RB | global; reconstructed | values/ conditions | both (19 traffic variables) | pre-defined |
| [29] | P | RA | reconstructed | scenario | both (categorical variables) | pre-defined |
| [30] | P | RA | reconstructed | scenario | both (6 clusters with variables) | pre-defined |
| [31], [32] | NDS | RA | road user based | scenario | both (varying parameters) | open |
| [33] | P | - | location based | scenario; trajectories/ maneuvers | spatial | - |
| [34] | P(AVA) | other | reconstructed | scenario (f; l; c) | spatial (5 locations) | pre-defined |
| [35] | P | RB | reconstructed | scenario | spatial (intersection accidents) | pre-defined |
| [36] | DC | - | road user based | scenario | objective | - |
| [37] | AVA | other | global; reconstructed | scenario | - | - |
| [38] | IGLAD | RB | reconstructed | scenario (PCM) | spatial (7 cluster fields) | pre-defined |
| [39] | NDS | other | road user based; reconstructed | scenario | - | open |
| [40] | TS | RA | road user based; location based | scenario | both (variation of traffic and lanes) | pre-defined |
| [41] | simulated | RA | road user based; global | scenario | objective | pre-defined |
| [42] | SC; simulated | RA | location based | trajectories/ maneuvers | spatial (roundabouts); objective (bicycles) | pre-defined |
| [43] | P; DC (Youtube) | RA | reconstructed | scenario (l; c) | both (12 categories) | pre-defined |
| [44] | NDS | RA | road user based; global | trajectories/ maneuvers | spatial | open |
| [45] | D | RA | location based | scenario | spatial (motorway) | - |
| [46] | P; DC | RB | road user based; reconstructed | scenario (f) | both (8 accident categories) | pre-defined |
| [47] | P | RA | reconstructed | scenario (f; l; c) | - | open |

**TABLE 4.** ***(Continued.)*** **Detailed categorization results of the reviewed studies.**

| | | | | | | |
|---|---|---|---|---|---|---|
| [48] | D | RA | location based | scenario (l) | objective | pre-defined |
| [49] | NDS | - | road user based; global | scenario (l) | spatial | open |
| [50] | D | - | location based | trajectories/ maneuvers | - | - |
| [51] | BUS; DC | - | road user based | detected scenarios in dataset | spatial (highway) | pre-defined (RNN) |
| [52] | BUS | - | road user based | scenario (OSC); trajectories/ maneuvers | spatial | open |
| [53] | SC; NDS | RA | road user based; location based | scenario (l) | objective (CIM; DCM) | pre-defined |
| [54] | simulated | RB | road user based; global | scenario | objective | open |
| [55] | P | RA | reconstructed | values/ conditions (CC) | spatial (intersection accidents) | pre-defined |
| [56] | C | - | road user based | trajectories/ maneuvers | objective | pre-defined (FastRCNN) |
| [57] | C; BUS | - | road user based; global | scenario | objective | open |
| [58] | NDS | - | global | scenario | - | - |
| [59] | C; NDS | - | road user based | scenario | objective | pre-defined |
| [60] | C | RA | road user based | scenario (grid-based) | spatial | pre-defined (RNN) |
| [61] | P | - | reconstructed | values/ conditions | objective (road departure crashes) | pre-defined |
| [62], [63] | DC; SC | RB | road user based; location based | scenario (human poses) | objective (pedestrian behavior) | pre-defined |
| [64] | AVA | other | reconstructed | scenario (road and actor matrices) | - | open |
| [65] | D | RA | location based | scenario (OSC) | both | pre-defined |
| [66] | NDS | RA | road user based; global | scenario | objective | pre-defined |
| [67] | P; BUS; simulated | RA | road user based; reconstructed | scenario | - | open |
| [68] | real driving datasets | RA | road user based | trajectories/ maneuvers | both (trajectories) | pre-defined |
| [69] | P | RA | reconstructed | values/ conditions | objective | open; pre-defined |
| [70] | P | - | reconstructed | values/ conditions (CC) | objective (bad visual accidents) | open |
| [71] | SC | - | location based | scenario | spatial | open |
| [72] | NDS | RA | reconstructed | scenario | objective | pre-defined (RNN) |
| [73] | SPMD | - | road user based; location based | scenario | 6 scenarios | pre-defined |
| [74] | D | RA | location based | scenario (c) | spatial (highway); objective (CIM; DCM) | pre-defined |
| [75] | P; DC | - | road user based; reconstructed | values/ conditions | spatial (urban accidents) | pre-defined |
| [76] | P; SC | RB | location based; reconstructed | values/ conditions | spatial (intersection accidents) | pre-defined |

SC: stationary camera data
SPMD: safety pilot model deployment dataset
TD: trajectory dataset
Euro NCAP: European New Car Assessment Program
TS: traffic simulator

2) **Representativity**

RA: representativity in the scenario space
RB: representativity in the natural population

3) **Output**

c: concrete scenarios
f: functional scenarios
l: logical scenarios

PCM: pre-crash matrix
OSC: OpenSCENARIO files
CC: crash characteristics

4) **ODD**
both: objective and spatial ODD specified
CIM: cut-in maneuver
DCM: deceleration maneuver
LCM: lane change maneuver
OTM: overtaking maneuver

5) **Scenario space**
CNN: convolutional neural network
RCNN: region-based convolutional neural network
RNN: recurrent neural network

## ACKNOWLEDGMENT

The authors would like to thank all the reviewers who have continuously improved the paper with their helpful feedback.

## REFERENCES

[1] R. Okuda, Y. Kajiwara, and K. Terashima, "A survey of technical trend of ADAS and autonomous driving," in *Proc. Tech. Papers Int. Symp. VLSI Design, Autom. Test*, Hsinchu, Taiwan, Apr. 2014, pp. 1–4, doi: 10.1109/VLSI-DAT.2014.6834940.

[2] Mercedes-Benz AG, Germany. *Mercedes-Benz Drive Pilot*. Accessed: Mar. 12, 2023. [Online]. Available: https://www.mercedes-benz.de/passengercars/technology/drive-pilot.html

[3] T. Otto and R. Auerswald, "Toolbox for test planning and test realization of scenario-based field tests for automated and connected driving," in *Automatisiertes Fahren*. Cham, Switzerland: Springer, 2020, pp. 165–180.

[4] C. Amersbach and H. Winner, "Defining required and feasible test coverage for scenario-based validation of highly automated vehicles," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 425–430.

[5] H. Elrofai, J.-P. Paardekooper, E. D. G. S. H. Kalisvaart, and O. O. D. Camp. (2018). *StreetWise: Scenario-based Safety Validation of Connected Automated Driving*. [Online]. Available: https://publications.tno.nl/publication/34626550/AyT8Zc/TNO-2018-streetwise.pdf

[6] C. Neurohr, L. Westhofen, T. Henning, T. de Graaff, E. Möhlmann, and E. Böde, "Fundamental considerations around scenario-based testing for automated driving," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Las Vegas, NV, USA, Oct. 2020, pp. 121–127.

[7] P. Junietz, W. Wachenfeld, K. Klonecki, and H. Winner, "Evaluation of different approaches to address safety validation of automated driving," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, HI, 2018, USA, Nov. 2018, pp. 491–496.

[8] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *Int. J. Surgery*, vol. 88, Aug. 2021, Art. no. 105906.

[9] G. Bagschik, T. Menzel, A. Reschka, M. Maurer, "Szenarien für Entwicklung, Absicherung und Test von automatisierten Fahrfunktionen," in *Proc. Workshop Fahrerassistenz Und Automatisiertes Fahren FAS*, 2017, pp. 1–13.

[10] M. Brini, E. Arnoux, B. Foyer, G. Bresson, L. Durville, F. Gaillard, and M. Pajon, "Moove & mosar projects: A scenario library for designing & validating ads," in *Proc. Driving Simul. Virtual Reality Conf. Exhibition*, 2020, pp. 13–17.

[11] J. Hiller, S. Kalyagina, and L. Eckstein, "Creating a framework for the representativeness assessment of automated driving safety assurance," in *Proc. ITS World Congr.*, 2022, pp. 1–14.

[12] Z. Tan, Y. Che, L. Xiao, W. Hu, P. Li, and J. Xu, "Research of fatal car-to-pedestrian precrash scenarios for the testing of the active safety system in China," *Accident Anal. Prevention*, vol. 150, Feb. 2021, Art. no. 105857.

[13] X. Wang, X. Li, and F. Cui, "Construction method of AEB pedestrian scenario based on accident data," in *Proc. CICTP*, Dec. 2020, pp. 13–24.

[14] X. Wang, Y. Peng, T. Xu, Q. Xu, X. Wu, G. Xiang, S. Yi, and H. Wang, "Autonomous driving testing scenario generation based on in-depth vehicle-to-powered two-wheeler crash data in China," *Accident Anal. Prevention*, vol. 176, Oct. 2022, Art. no. 106812.

[15] Z. Zhang, J. Shi, K. Guo, and J. Wang, "Research on construction method and application of autonomous driving test scenario database," in *CICTP 2020*. Reston, VA, USA: Amer. Soc. Civil Eng., 2020, pp. 311–323.

[16] Y. Akagi, R. Kato, S. Kitajima, J. Antona-Makoshi, and N. Uchida, "A risk-index based sampling method to generate scenarios for the evaluation of automated driving vehicle safety," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Auckland, New Zealand, Oct. 2019, pp. 667–672.

[17] L. Balasubramanian, F. Kruber, M. Botsch, and K. Deng, "Open-set recognition based on the combination of deep learning and ensemble method for detecting unknown traffic scenarios," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jul. 2021, pp. 674–681.

[18] L. Balasubramanian, J. Wurst, M. Botsch, and K. Deng, "Traffic scenario clustering by iterative optimisation of self-supervised networks using a random forest activation pattern similarity," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Nagoya, Japan, Jul. 2021, pp. 682–689.

[19] M. Bäumler and G. Prokop, "Validating automated driving systems by using scenario-based testing: The Fuse4Rep process model for scenario generation as part of the Dresden method," *Zeitschrift Für Verkehrssicherheit*, vol. 68, no. 3, pp. 226–230, 2022. [Online]. Available: https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-900106

[20] T. Braun, L. Ries, F. Körtke, L. Turner, S. Otten, and E. Sax, "Collection of requirements and model-based approach for scenario description," in *Proc. 7th Int. Conf. Vehicle Technol. Intell. Transp. Syst.*, 2021, pp. 634–645.

[21] H. de Carvalho Pinheiro, D. Cruz Stanke, A. Ferraris, M. Carello, G. Gabiati, I. Camuffo, and M. Grillo, "Autonomous driving scenario generation in overtake manoeuvres through data fusion," in *Mechanisms and Machine Science*. SpringerCham, Switzerland: Springer 2021, pp. 786–794.

[22] C. Chang, D. Cao, L. Chen, K. Su, K. Su, Y. Su, F.-Y. Wang, J. Wang, P. Wang, J. Wei, G. Wu, X. Wu, H. Xu, N. Zheng, and L. Li, "MetaScenario: A framework for driving scenario data description, storage and indexing," *IEEE Trans. Intell. Vehicles*, vol. 8, no. 2, pp. 1156–1175, Feb. 2023.

[23] N. E. Chelbi, D. Gingras, and C. Sauvageau, "Worst-case scenarios identification approach for the evaluation of advanced driver assistance systems in intelligent/autonomous vehicles under multiple conditions," *J. Intell. Transp. Syst.*, vol. 26, no. 3, pp. 284–310, May 2022.

[24] C. Cheng, "Design and implementation of test scenario library data system for autonomous vehicles," in *Proc. SAE Tech. Paper*, 2020, p. 5219.

[25] P. Elspas, Y. Klose, S. Isele, J. Bach, and E. Sax, "Time series segmentation for driving scenario detection with fully convolutional networks," in *Proc. 7th Int. Conf. Vehicle Technol. Intell. Transp. Syst.*, 2021, pp. 56–64.

[26] P. Elspas, J. Langner, M. Aydinbas, J. Bach, and E. Sax, "Leveraging regular expressions for flexible scenario detection in recorded driving data," in *Proc. IEEE Int. Symp. Syst. Eng. (ISSE)*, Vienna, Austria, Oct. 2020, pp. 1–8, doi: 10.1109/ISSE49799.2020.9272025.

[27] P. Elspas, J. Lindner, M. Brosowsky, J. Bach, and E. Sax, "Towards a scenario database from recorded driving data with regular expressions for scenario detection," in *Proc. 8th Int. Conf. Vehicle Technol. Intell. Transp. Syst.*, 2022, pp. 400–409.

[28] E. Esenturk, S. Khastgir, A. Wallace, and P. Jennings, "Analyzing real-world accidents for test scenario generation for automated vehicles," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jul. 2021, pp. 288–295.

[29] E. Esenturk, D. Turley, A. Wallace, S. Khastgir, and P. Jennings, "A data mining approach for traffic accidents, pattern extraction and test scenario generation for autonomous vehicles," *Int. J. Transp. Sci. Technol.*, vol. 12, no. 4, pp. 955–972, Dec. 2023, doi: 10.1016/j.ijtst.2022.10.002.

[30] E. Esenturk, A. G. Wallace, S. Khastgir, and P. Jennings, "Identification of traffic accident patterns via cluster analysis and test scenario development for autonomous vehicles," *IEEE Access*, vol. 10, pp. 6660–6675, 2022.

[31] S. Feng, Y. Feng, H. Sun, S. Bao, Y. Zhang, and H. X. Liu, "Testing scenario library generation for connected and automated vehicles, Part II: Case studies," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 5635–5647, Sep. 2021.

[32] S. Feng, Y. Feng, C. Yu, Y. Zhang, and H. X. Liu, "Testing scenario library generation for connected and automated vehicles, Part I: Methodology," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1573–1582, Mar. 2021.

[33] A. Gambi, V. Nguyen, J. Ahmed, and G. Fraser, "Generating critical driving scenarios from accident sketches," in *Proc. IEEE Int. Conf. Artif. Intell. Test. (AITest)*, Newark, CA, USA, Aug. 2022, pp. 95–102.

[34] Q. Goss, Y. AlRashidi, and M. I. Akbas, "Generation of modular and measurable validation scenarios for autonomous vehicles using accident data," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jul. 2021, pp. 251–257.

[35] I. Han, "Scenario establishment and characteristic analysis of intersection collision accidents for advanced driver assistance systems," *Traffic Injury Prevention*, vol. 21, no. 6, pp. 354–358, Aug. 2020.

[36] A. Harder, J. Ranjit, and M. Behl, "Scenario2 vector," in *Proc. ACM/IEEE 12th Int. Conf. Cyber-Phys. Syst.*, May 2021, pp. 167–176.

[37] J. C. Holland and A. Sargolzaei, "Verification of autonomous vehicles: Scenario generation based on real world accidents," in *Proc. SoutheastConf.*, Mar. 2020, pp. 1–17, doi: 10.1109/southeast-con44009.2020.9368284.

[38] W. Hu, X. Xu, Z. Zhou, Y. Liu, Y. Wang, L. Xiao, and X. Qian, "Mining and comparative analysis of typical pre-crash scenarios from IGLAD," *Accident Anal. Prevention*, vol. 145, Sep. 2020, Art. no. 105699.

[39] S. B. Jayapal Gowdu, P. Aust, A. Schwind, F. Hau, and M. A. Hein, "Evaluation of scenario-based automotive radar testing in virtual environment using real driving data," in *Proc. IEEE 25th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2022, pp. 2379–2384.

[40] I. R. Jenkins, L. O. Gee, A. Knauss, H. Yin, and J. Schroeder, "Accident scenario generation with recurrent neural networks," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, HI, 2018, USA, Nov. 2018, pp. 3340–3345.

[41] D. Kaufmann, L. Klampfl, F. Klück, M. Zimmermann, and J. Tao, "Critical and challenging scenario generation based on automatic action behavior sequence optimization: 2021 IEEE autonomous driving AI test challenge group 108," in *Proc. IEEE Int. Conf. Artif. Intell. Test.*, Aug. 2021, pp. 118–127.

[42] A. Keler, P. Malcolm, G. Grigoropoulos, S. A. Hosseini, H. Kaths, F. Busch, and K. Bogenberger, "Data-driven scenario specification for AV–VRU interactions at urban roundabouts," *Sustainability*, vol. 13, no. 15, p. 8281, Jul. 2021.

[43] D. Kibalama, P. Tulpule, and B.-S. Chen, "AV/ADAS safety-critical testing scenario generation from vehicle crash data," in *Proc. SAE Technical Paper*, 2022, p. 104.

[44] L. Klitzke, J. Meyer, T. Leune, C. Koch, and F. Köster, "DAGMaR: A DAG-based robust road membership estimation framework for scenario mining," in *Proc. 6th Int. Conf. Internet Things, Syst. Manage. Secur. (IOTSMS)*, Oct. 2019, pp. 358–365.

[45] C. Knies and F. Diermeyer, "Data-driven test scenario generation for cooperative maneuver planning on highways," *Appl. Sci.*, vol. 10, no. 22, p. 8154, Nov. 2020.

[46] W. Ko, S. Park, S. Park, H. Jeong, and I. Yun, "Development of freeway-based test scenarios for applying new car assessment program to automated vehicles," *PLoS ONE*, vol. 17, no. 7, Jul. 2022, Art. no. e0271532.

[47] W. Ko, S. Park, J. Yun, S. Park, and I. Yun, "Development of a framework for generating driving safety assessment scenarios for automated vehicles," *Sensors*, vol. 22, no. 16, p. 6031, Aug. 2022.

[48] N. Kolb, F. Hauer, M. Golagha, and A. Pretschner, "Data-driven assessment of parameterized scenarios for autonomous vehicles," in *Computer Safety, Reliability, and Security*. SpringerCham, Switzerland: Springer 2022, pp. 350–364.

[49] J. Langner, H. Grolig, S. Otten, M. Holzäpfel, and E. Sax, "Logical scenario derivation by clustering dynamic-length-segments extracted from real-world-driving-data," in *Proc. 5th Int. Conf. Vehicle Technol. Intell. Transp. Syst.*, 2019, pp. 458–467.

[50] T. Moers, C. Klas, A. Zlocki, and L. Vater, "Real world scenarios for the safety validation and development of highly automated driving functions," in *Fahrerassistenzsysteme Und Automatisiertes Fahren*. Düsseldorf, Germany: VDI Verlag, 2022, pp. 121–130.

[51] F. Montanari, H. Ren, and A. Djanatliev, "Scenario detection in unlabeled real driving data with a rule-based state machine supported by a recurrent neural network," in *Proc. IEEE 93rd Veh. Technol. Conf. (VTC-Spring)*, Apr. 2021, pp. 1–5, doi: 10.1109/VTC2021-Spring51267.2021.9449032.

[52] F. Montanari, C. Stadler, J. Sichermann, R. German, and A. Djanatliev, "Maneuver-based resimulation of driving scenarios based on real driving data," in *Proc. IEEE Intell. Vehicles Symp.*, Nagoya, Japan, Jul. 2021, pp. 1124–1131.

[53] H. Nakamura, H. Muslim, R. Kato, S. Préfontaine-Watanabe, H. Nakamura, H. Kaneko, H. Imanaga, J. Antona-Makoshi, S. Kitajima, N. Uchida, E. Kitahara, K. Ozawa, and S. Taniguchi, "Defining reasonably foreseeable parameter ranges using real-world traffic data for scenario-based safety assessment of automated vehicles," *IEEE Access*, vol. 10, pp. 37743–37760, 2022.

[54] D. Nalic, A. Eichberger, G. Hanzl, M. Fellendorf, and B. Rogic, "Development of a co-simulation framework for systematic generation of scenarios for testing and validation of automated driving systems," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 1895–1901.

[55] P. Nitsche, P. Thomas, R. Stuetz, and R. Welsh, "Pre-crash scenarios at road junctions: A clustering method for car crash data," *Accident Anal. Prevention*, vol. 107, pp. 137–151, Oct. 2017.

[56] Park, Wen, Sung, and Cho, "Multiple event-based simulation scenario generation approach for autonomous vehicle smart sensors and devices," *Sensors*, vol. 19, no. 20, p. 4456, Oct. 2019.

[57] S.-W. Park, K. Patil, W. Wilson, M. Corless, G. Choi, and P. Adam, "Creating driving scenarios from recorded vehicle data for validating lane centering system in highway traffic," in *Proc. SAE Tech. Paper*, 2020, p. 718.

[58] R. Pfeffer, J. He, and S. Eric, "Development and implementation of a concept for the meta description of highway driving scenarios with focus on interactions of road users," in *Proc. 6th Int. Conf. Vehicle Technol. Intell. Transp. Syst.*, 2020, pp. 440–447.

[59] C. Reichenbächer, M. Rasch, Z. Kayatas, F. Wirthmüller, J. Hipp, T. Dang, and O. Bringmann, "Identifying scenarios in field data to enable validation of highly automated driving systems," in *Proc. 8th Int. Conf. Vehicle Technol. Intell. Transp. Syst.*, 2022, pp. 134–142.

[60] L. Ries, J. Langner, S. Otten, J. Bach, and E. Sax, "A driving scenario representation for scalable real-data analytics with neural networks," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 2215–2222.

[61] D. Shen, Q. Yi, L. Li, R. Tian, S. Chien, Y. Chen, and R. Sherony, "Test scenarios development and data collection methods for the evaluation of vehicle road departure prevention systems," *IEEE Trans. Intell. Vehicles*, vol. 4, no. 3, pp. 337–352, Sep. 2019.

[62] J. Spooner, V. Palade, M. Cheah, S. Kanarachos, and A. Daneshkhah, "Generation of pedestrian crossing scenarios using ped-cross generative adversarial network," *Appl. Sci.*, vol. 11, no. 2, p. 471, Jan. 2021.

[63] J. Spooner, V. Palade, A. Daneshkhah, and S. Kanarachos, "Using generative adversarial networks and non-roadside video data to generate pedestrian crossing scenarios," in *Proc. 20th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2021, pp. 1413–1420.

[64] C. Stark, C. Medrano-Berumen, and M. I. Akbas, "Generation of autonomous vehicle validation scenarios using crash data," in *Proc. SoutheastConf.*, Mar. 2020, pp. 1–6, doi: 10.1109/Southeast-Con44009.2020.9249662.

[65] A. Tenbrock, A. König, T. Keutgens, and H. Weber, "The ConScenD dataset: Concrete scenarios from the highD dataset according to ALKS regulation UNECE R157 in OpenX," in *Proc. IEEE Intell. Vehicles Symp. Workshops*, Nagoya, Japan. London, U.K.: Concrete, Jul. 2021, pp. 174–181.

[66] S. Thal, R. Henze, R. Hasegawa, H. Nakamura, H. Imanaga, J. Antona-Makoshi, and N. Uchida, "Generic detection and search-based test case generation of urban scenarios based on real driving data," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2022, pp. 694–701.

[67] P. Themann, A. Pütz, A. Zlocki, and L. Eckstein, "Database of relevant traffic scenarios as a tool for the development and validation of automated driving," in *Proc. 13th Int. Symp. Adv. Vehicle Control*, 2016, pp. 489–496.

[68] L. Töttel, M. Zipfl, D. Bogdoll, M. R. Zofka, and J. M. Zöllner, "Reliving the dataset: Combining the visualization of road users' interactions with scenario reconstruction in virtual reality," in *Proc. 6th Int. Conf. Intell. Transp. Eng.*, 2022, pp. 436–454.

[69] H. Watanabe, L. Tobisch, J. Rost, J. Wallner, and G. Prokop, "Scenario mining for development of predictive safety functions," in *Proc. IEEE Int. Conf. Veh. Electron. Saf. (ICVES)*, Sep. 2019, pp. 1–7, doi: 10.1109/ICVES.2019.8906293.

[70] T. Winkle, C. Erbsmehl, and K. Bengler, "Area-wide real-world test scenarios of poor visibility for safe development of automated vehicles," *Eur. Transp. Res. Rev.*, vol. 10, no. 2, Jun. 2018, Art. no. 32.

[71] Z. Xinxin, L. Fei, and W. Xiangbin, "CSG: Critical scenario generation from real traffic accidents," in *Proc. IEEE Intell. Vehicles Symp.*, Oct. 2020, pp. 1330–1336.

[72] L. Xu, C. Zhang, Y. Liu, L. Wang, and L. Li, "Worst perception scenario search for autonomous driving," in *Proc. IEEE Intell. Vehicles Symp.*, Oct. 2020, pp. 1702–1707.
[73] D. Zhao, Y. Guo, and Y. J. Jia, "TrafficNet: An open naturalistic driving scenario library," in *Proc. IEEE 20th Int. Conf. Intell. Transp. Syst.*, Oct. 2017, pp. 1–8, doi: 10.1109/ITSC.2017.8317860.
[74] A. Zlocki, A. König, J. Bock, H. Weber, H. Muslim, H. Nakamura, S. Watanabe, J. Antona-Makoshi, and S. Taniguchi, "Logical scenarios parameterization for automated vehicle safety assessment: Comparison of deceleration and cut-in scenarios from Japanese and German highways," *IEEE Access*, vol. 10, pp. 26817–26829, 2022.
[75] S. Krause and F. Busch, "Identification of hazards in urban driving situations using accident categorizations," in *Proc. IEEE Intell. Vehicles Symp.*, Oct. 2020, pp. 156–161.
[76] M. Bäumler, L. Dziuba-Kaiser, Z. Yin, M. Lehmann, and G. Prokop, "Use information you have never observed together: Data fusion as a major step towards realistic test scenarios," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2020, pp. 1–8, doi: 10.1109/ITSC45102.2020.9294649.
[77] *Road Vehicles—Safety of the Intended Functionality*, Standard ISO 21448:2022, 2022.
[78] T. Menzel, G. Bagschik, and M. Maurer, "Scenarios for development, test and validation of automated vehicles," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2018, pp. 1821–1827.
[79] C. Siebke, M. Bäumler, K. Blenz, M. Lehmann, M. Ringhand, M. Mai, and G. Prokop, "Predicting the impact on road safety of an intersection AEB at urban intersections. Using a novel virtual test field for the assessment of conflict prevention between cyclists/pedelecs and cars," *Transp. Res. Interdiscipl. Perspect.*, vol. 17, Jan. 2023, Art. no. 100728.
[80] M. Scholtes, L. Westhofen, L. R. Turner, K. Lotto, M. Schuldes, H. Weber, N. Wagener, C. Neurohr, M. H. Bollmann, F. Körtke, J. Hiller, M. Hoss, J. Bock, and L. Eckstein, "6-layer model for a structured description and categorization of urban traffic and environment," *IEEE Access*, vol. 9, pp. 59131–59147, 2021.
[81] G. Bagschik, T. Menzel, and M. Maurer, "Ontology based scene creation for the development of automated vehicles," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1813–1820.
[82] D. Nalic, T. Mihalj, A. Eichberger, M. Bäumler, and M. Lehmann, "Scenario based testing of automated driving systems: A literature survey," in *Proc. FISITA World Congr. Tech. Programme*, Sep. 2021, p. 96.
[83] M. Lehmann, M. Bäumler, G. Prokop, and D. Hamelow, "Use of a criticality metric for assessment of critical traffic situations as part of SePIA," in *Proc. Int. Stuttgarter Symp.*, 2019, pp. 1154–1167.
[84] H. Watanabe, T. Maly, J. Wallner, T. Dirndorfer, M. Mai, and G. Prokop, "Methodology of scenario clustering for predictive safety functions," in *Tagung Automatisiertes Fahren*. Munich, Germany: TU München Lehrstuhl für Fahrzeugtechnik mit TÜV SÜD Akademie, 2019.
[85] ASAM, Germany. *ASAM OpenSCENARIO*. Accessed: Jan. 26, 2023. [Online]. Available: https: //www.asam.net/standards/detail/openscenario/
[86] ASAM, Germany. *ASAM OpenDRIVE*. Accessed: Jan. 26, 2023. [Online]. Available: https: //www.asam.net/standards/detail/opendrive/
[87] ASAM, Germany. *ASAM Standards*. Accessed: Jan. 26, 2023. [Online]. Available: https: //www.asam.net/standards/
[88] A. Pütz, A. Zlocki, J. Bock, and L. Eckstein, "System validation of highly automated vehicles with a database of relevant traffic scenarios," in *Proc. 12th ITS European Congr.*, 2017, pp. 19–22.
[89] H. Winner, K. Lemmer, T. Form, and J. Mazzega, "PEGASUS-first steps for the safe introduction of automated driving," in *Road Vehicle Automation*. Cham, Switzerland: Springer 2019, pp. 185–195.
[90] M. Köbe, "Eine Bewertungsmethodik zur Inspektion automatisierter Fahrfunktionen. Szenarien-basierte Plattform zur Inspektion automatisierter Fahrfunktionen. Das Projekt SePIA," in *Hochschule Fuer Angewandte Wissenschaften Kempten*. Dresden, Germany: Schriftenreihe, 2017.
[91] A. Leitner, "ENABLE-S3: Project introduction," in *Validation and Verification of Automated Systems*. Cham, Switzerland: Springer 2020, pp. 13–23.
[92] Sakura. *Sakura Research Project*. Accessed: Jan. 26, 2023. [Online]. Available: https://www.sakura-prj.go.jp/
[93] *Headstart*. Accessed: Mar. 10, 2023. [Online]. Available: https://www.headstart-project.eu/
[94] *Deutsches Zentrum Fur Luft-und Raumfahrt E.V. SET Level: The Safety of Automated Driving*. Accessed: Mar. 10, 2023. [Online]. Available: https://setlevel.de/en
[95] H. Elrofai, J.-P. Paardekooper, E. de Gelder, S. Kalisvaart, and O. O. den Camp, *Scenario-based Safety Validation of Connected and Automated Driving*. The Hague, Netherlands: TNO, 2018.
[96] Vedecom Institut. (2024). *MOOVE—Monitoring Outillé Pourle Vehicule Dans Son Environnement*. Accessed: Jan. 23, 2024. [Online]. Available: https://www.vedecom.fr/wp-content/uploads/FICHE-MOOVE-VF-2.pdf
[97] *L3Pilot Driving Automation*. Accessed: Mar. 10, 2023. [Online]. Available: https://l3pilot.eu/
[98] *VV Methoden*. Accessed: Mar. 10, 2023. [Online]. Available: https://www. vvm-projekt.de/
[99] *HI-DRIVE*. Accessed: Mar. 21, 2023. [Online]. Available: https://www.hi-drive.eu/
[100] *V4Safety Project*. Accessed: Mar. 10, 2023. [Online]. Available: https://v4safetyproject.eu/
[101] Bundesministerium für Digitales und Verkehr. (2023). *Sicherheit Des Vernetzte Und Automatisierten Straenverkehrs—Siva*. Accessed: Jul. 17, 2023. [Online]. Available: https://bmdv.bund.de/SharedDocs/DE/Artikel/DG/mfund-projekte/sivas.html
[102] Y. Page, F. Fahrenkrog, A. Fiorentino, J. Gwehenberger, T. Helmer, M. Lindman, O. O. D. Camp, L. V. Rooij, S. Puch, M. Fränzle, U. Sander, and P. Wimmer, "A comprehensive and harmonized method for assessing the effectiveness of advanced driver assistance systems by virtual simulation: The PEARS initiative," in *Proc. 24th Int. Tech. Conf. Enhanced Saf. Vehicles (ESV)*, 2015, pp. 1–12.
[103] IAMTS. (2024). *International Alliance for Mobility Testing and Standardization*. Accessed: Jan. 23, 2024. [Online]. Available: https://www.iamts.org/en
[104] (2024). *Safety PoolTM—The Global Initiative For Certifiable AV Safety*. Accessed: Feb. 28, 2024. [Online]. Available: https://www.safetypool.ai
[105] T. Mihalj, D. Nalic, S. Arefnezhad, and A. Eichberger, "Hazards identification using scenario-based testing with respect to ISO PAS 21448 and ISO 26262," in *Proc. IEEE 26th International Conference Intelligent Transportation Systems (ITSC)*, 2023, pp. 5764–5770.
[106] W. Ding, C. Xu, M. Arief, H. Lin, B. Li, and D. Zhao, "A survey on safety-critical driving scenario generation—A methodological perspective," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 7, pp. 1–18, May 2023.
[107] H. Ren, H. Gao, H. Chen, and G. Liu, "A survey of autonomous driving scenarios and scenario databases," in *Proc. 9th Int. Conf. Dependable Syst. Their Appl. (DSA)*, Wulumuqi, China, Aug. 2022, pp. 754–762.
[108] Verkehrsunfallforschung an der TU Dresden GmbH. *GIDAS—German In-Depth Accident Study*. Accessed: Mar. 10, 2023. [Online]. Available: https://www.gidas.org/start.html
[109] Q. Chen and B. Dai, "Characteristics and casualty analysis of two-wheeler accidents in China, data source: The China in-depth accident study (CIDAS)," in *Proc. SAE Tech. Paper Ser.*, Apr. 2018, pp. 1–14.
[110] B. Deng, H. Wang, J. Chen, X. Wang, and X. Chen, "Traffic accidents in Shanghai-general statistics and in-depth analysis," in *Proc. 23rd Int. Tech. Conf. Enhanced Saf. Vehicles (ESV)*, 2013, p. 182.
[111] S. Kim, J. Lee, and Y. Youn, "A study on the construction of the database structure for the Korea in-depth accident study," *Trans. Korean Soc. Automot. Engineers*, vol. 22, no. 2, pp. 29–36, Mar. 2014.
[112] NHTSA. (2023). *National Automotive Sampling System: General Estimates System (GES)*. Accessed: Jul. 17, 2023. [Online]. Available: https://www.nhtsa.gov/crash-data-systems/national-automotive-sampling-system
[113] J. Bakker, H. Jeppsson, L. Hannawald, F. Spitzhüttl, A. Longton, and E. Tomasch, "IGLAD-International harmonized in-depth accident data," in *Proc. 25th Int. Conference Enhanced Saf. Vehicles (ESV)*, 2017, p. 248.
[114] IGLAD Initative. (2023). *IGLAD—Initiative for the Global Harmonisation of Accident Data: Data Description*. Accessed: Jul. 17, 2023. [Online]. Available: http://www.iglad.net/web/page.aspx?refid=2
[115] J. M. Hankey, M. A. Perez, and J. A. McClafferty. (2016). *Description of the SHRP 2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets*. [Online]. Available: http://hdl.handle.net/10919/70850
[116] M. Bäumler and M. Lehmann. (2023). *ListDB: Representative Datasets of Video-based Traffic Observations*. Accessed: Jul. 17, 2023. [Online]. Available: https://opara.zih and https://opara.zih.tu-dresden.de/xmlui/handle/123456789/5882

[117] (2017). Statistisches Bundesamt Deutschland. *Qualitütsbericht: Statistik der Verkehrsunfülle*. [Online]. Available: https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Verkehrsunfaelle/strassenverkehrsunfaelle.pdf?_blob=publicationFile
[118] M. Pfeiffer and J. Schmidt, "Statistical and methodological foundations of the GIDAS accident survey system," in *Proc. 2nd Int. Conf. ESAR Expert Symp. Accident Res.*, 2007, pp. 81–87.
[119] K. Czarnecki, "Operational design domain for automated driving systems," in *Waterloo Intelligent Systems Engineering (WISE) Lab*. Waterloo, Canada: University of Waterloo, 2018.
[120] J. Cai, W. Deng, H. Guang, Y. Wang, J. Li, and J. Ding, "A survey on data-driven scenario generation for automated vehicle testing," *Machines*, vol. 10, no. 11, p. 1101, Nov. 2022.
[121] X. Zhang, J. Tao, K. Tan, M. Törngren, J. M. G. Sánchez, M. R. Ramli, X. Tao, M. Gyllenhammar, F. Wotawa, N. Mohan, M. Nica, and H. Felbinger, "Finding critical scenarios for automated driving systems: A systematic mapping study," *IEEE Trans. Softw. Eng.*, vol. 49, no. 3, pp. 991–1026, Mar. 2023.
[122] M. Bäumler and G. Prokop, "Test scenario fusion: How to fuse scenarios from accident and traffic observation data," *IEEE Access*, vol. 12, pp. 16354–16374, 2024, doi: 10.1109/ACCESS.2023.3340442.
[123] A.-M. Jacobo, U. Nobuyuki, Y. Kunio, O. Koichiro, K. Eiichi, and T. Satoshi, "Development of a safety assurance process for autonomous vehicles in Japan," in *Proc. ESV Conf.*, 2019, pp. 1–14.
[124] SAE Int. (2021). *J3016 202104: Taxonomy and Definitions for Terms Related To Driving Automation Systems for On-Road Motor Vehicles*. Accessed: Jul. 18, 2023. [Online]. Available: https://www.sae.org/standards/content/j3016underscore202104/
[125] *Road Vehicles-Functional Safety*, Standard ISO 26262, 2018.
[126] A. J. Braverman, "Data fusion," in *Encyclopedia of Quantitative Risk Analysis and Assessment*. Hoboken, NJ, USA: Wiley, 2008.
[127] N. K. Chauhan and K. Singh, "A review on conventional machine learning vs deep learning," in *Proc. Int. Conf. Comput., Power Commun. Technol. (GUCON)*, Greater Noida, India, Sep. 2018, pp. 347–352.
[128] J. Yu, A. de Antonio, and E. Villalba-Mora, "Deep learning (CNN, RNN) applications for smart homes: A systematic review," *Computers*, vol. 11, no. 2, p. 26, Feb. 2022.
[129] P. Dhruv and S. Naskar, "Image classification using convolutional neural network (CNN) and recurrent neural network (RNN): A review," in *Machine Learning and Information Processing*. SpringerCham, Switzerland: Springer 2020, pp. 367–381.
[130] Y. Zhu, J. Wang, F. Meng, and T. Liu, "Review on functional testing scenario library generation for connected and automated vehicles," *Sensors*, vol. 22, no. 20, p. 7735, Oct. 2022.
[131] S. Riedmaier, T. Ponn, D. Ludwig, B. Schick, and F. Diermeyer, "Survey on scenario-based safety assessment of automated vehicles," *IEEE Access*, vol. 8, pp. 87456–87477, 2020.
[132] M. Bäumler, M. Lehmann, and G. Prokop, "Generating representative test scenarios: The fuse for representativity (FUSE4Rep) process model for collecting and analysing traffic observation data," in *Proc. 27th ESV Conf.*, 2023, pp. 1–25, doi: 10.26128/2024.
[133] RWTH Aachen Univ. *The Highway Drone Dataset*. Accessed: Mar. 14, 2023. [Online]. Available: https://www.highd-dataset.com/
[134] U.K. Dept. for Transp. (2023). *Road Safety Data*. Accessed: Mar. 14, 2023. [Online]. Available: https://www.data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data
[135] IPG Automot. *CarMaker*. Accessed: Mar. 14, 2023. [Online]. Available: https://ipg-automotive.com/de/produkte-loesungen/software/carmaker/
[136] LG Electron. *SVL Simulator*. Accessed: Mar. 14, 2023. [Online]. Available: https://www.svlsimulator.com/
[137] E. Polders and T. Brijs. (2018). *Horizon 2020 EC Project, InDeV*. [Online]. Available: https://documentserver.uhasselt.be/bitstream/1942/27208/2/Handbook-InDeV.pdf
[138] S. Feng, H. Sun, X. Yan, H. Zhu, Z. Zou, S. Shen, and H. X. Liu, "Dense reinforcement learning for safety validation of autonomous vehicles," *Nature*, vol. 615, no. 7953, pp. 620–627, Mar. 2023.
[139] F. Hauer, T. Schmidt, B. Holzmüller, and A. Pretschner, "Did we test all scenarios for automated and autonomous driving systems?" in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 2950–2955.

**MAXIMILIAN BÄUMLER** received the Diploma degree from TU Dresden, Germany, in 2017, and the M.B.A. degree from Collège des Ingénieurs, in 2018. Since 2018, he has been a Researcher with the Chair of Automobile Engineering, TU Dresden, in the field of scenario generation for assessing automated driving systems. His current research interests include the fusion of police accidents and video-based traffic observation data to generate representative test scenarios. For more information, see www.mbaeumler.de.

**FELIX LINKE** received the Diploma degree from TU Dresden, in 2024. In his specialization in automotive engineering, he focused on scenario-based testing of automated driving systems. As a Student Trainee, he was with Fraunhofer IVI in the research group on vehicle and road safety and reconstructing accidents from police data.

**GÜNTHER PROKOP** received the Dr.-Ing. degree from the Technical University of Munich, Germany, in 1998. After more than ten years of work in German Automotive Industry, he became a Professor with the Chair of Automobile Engineering, TU Dresden, in 2010. From 2018 to 2022, he was the Dean of the "Friedrich List" Faculty of Transport and Traffic Sciences, TU Dresden.