VDA | Verband der Automobilindustrie

FAT | Forschungsvereinigung Automobiltechnik

FAT-Schriftenreihe 343

Objective assessment of database quality for use in the automotive research and development process

Objective assessment of database quality for use in the automotive research and development process

Johann Ziegler, Henrik Liers

Verkehrsunfallforschung an der TU Dresden GmbH

Albine Chanove, Maria Pohle

Fraunhofer-Institut für Verkehrs- und Infrastruktursysteme IVI

Das Forschungsprojekt wurde mit Mitteln der Forschungsvereinigung Automobiltechnik e.V. (FAT) gefördert.

Acknowledgments

The Research Association for Automotive Technology (FAT) has commissioned the Traffic Accident Research Institute at TU Dresden GmbH (VUFO) and the Fraunhofer Institute for Transportation and Infrastructure Systems (IVI) to work on the mentioned research project. Vehicle manufacturers and numerous suppliers have joined in the FAT to conduct pre-competitive and joint research under the framework of the German Association of the Automotive Industry (VDA).

The initiation of this research project was originated from the FAT AK3 working group by Dr.-Ing. Michael Düring (Volkswagen AG) and Michael Wagner (Continental AG). The two FAT project leaders were always valuable idea generators. Their constructive criticism and approaches contributed significantly to the success of this research project. Our first gratitude belongs to them.

Our second gratitude belongs to all members of FAT AK3 working group, who contributed to the development of the matching process with a scientific questionnaire and many approaches during the interim meetings.

Due to the support letter by Mrs. Prof. Dr.-Ing. Langowsky (FAT), many data providers could be won for the research project by providing meta data information on database. For this effort, we would like to thank Mrs. Prof. Dr.-Ing. Langowsky and all supporters.

- International Traffic Safety Data and Analysis Group (IRTAD)
- Initiative for the global harmonisation of accident data (IGLAD)
- Aalborg University - Traffic research group by H. Lahrmann
- Hellenic Institute of Transport (HIT) by D. Margaritis
- Insurance Institute for Highway Safety by B. Mueller for information on the VIPA project

I. Table of contents

II. Abstract

Numerous interdisciplinary aspects are considered in the research and development environment, particularly in the field of vehicle and traffic safety. The relevant issues of assisted, connected, and automated driving and the further development of passive and integral safety systems require reliable data sources. The sometimes very heterogeneous traffic and accident situation between countries and continents makes it necessary to take as many as possible data sources from several countries/regions into account. The aim of the present work is to develop a unified meta database that contains all necessary information for the research and development departments on a meta based level (no raw data) for several countries.

One of the main objectives is the research on international data sources in the field of traffic and vehicle safety. This includes national road accident statistics based on police accident data as well as highly detailed investigations in smaller regions (so-called In-depth data sources). At the beginning of this research project, the following countries are defined for further investigation: Germany, France, Greece, Czech Republic, Sweden, Denmark and USA. The research on data sources in the selected countries in the field of traffic and vehicle safety identified 32 data sources out of which 19 could be fully inventoried in the meta database. In total 64 data sources are identified worldwide.

The basis for the development of the meta database is the German In-Depth Accident Study (GIDAS). GIDAS is a cooperation project of the Federal Highway Research Institute (BAST) and the Research Association for Automotive Technology (FAT). For the development, fragments of the GIDAS database structure as well as knowledge and expertise in database development are used. The structure of the meta database is characterised by a general fact sheet information about the researched data source and its content on accident-relevant variables. In total, the meta database contains 15 data tables that are uniquely assigned via an identification number (primary key). Parallel to the meta database development, a codebook for the description of 237 variables contained in the meta database has been developed

In addition, a questionnaire is used to check the applicability of the developed meta database for specific questions from the German automotive industry. Therefore, the same set of variables from the meta database is used to code the questionnaire. For the objective assessment between the meta database and questionnaire a matching process is developed. The aim of the matching process is to indicate the covered percentage of necessary variables in the various data sources for each question. The perfect matches are represented by a 100 % coverage, suggesting that this data source may be one of the most appropriate ones for answering the question. The results of the matching process are collected in a result matrix

The meta database and the result matrix can be a useful tool to make the development process even more efficient by minimising the research time for suitable data sources to answer the relevant development questions in the field of the vehicle safety. Furthermore, the meta database can act as a platform to bring several data providers from different countries together and to encourage the global harmonisation of road accident data sources.

III. Task

The task of this research project is subdivided into several tasks, which are described more in detail in the following report:

- Identification of data sources in the environment of vehicle safety and road traffic accidents
- Contact and consultation of data providers with a detailed interview on the investigation methods, data content and quality management of the data source
- Development of a result matrix (meta database) incl. applicability assessment in a suitable form
- Checking the applicability of the data sources for certain research questions specified by the German automotive industry and suppliers
- Development of a matching process to compare the meta database and the questionnaire

IV. Abbreviations and indices

Symbol	Description
AADT	Average Annual Daily Traffic
ADAC	German Automobile Club (Allgemeiner Deutscher Automobil-Club e.V.)
AIS	Abbreviated Injury Scale
BAAC	Bulletin d'analyse d'accidents corporels (analysis report of road accidents involving physical injury)
BAST	Federal Highway Research Institute
CARE	Community Road Accident Database
CDS	Crashworthiness Data System
CDV	Transport Research Centre (Czech Republic)
CEESAR	Centre Eurpéen d'Etudes de Sécruité et d'Analyse des Risques
CERTH	Centre for Research and Technology Hellas
CISS	Crash Investigation Sampling System
CRSS	Crash Report Sampling System
CZ	Czech Republic
CZIDAS	Czech In-Depth Accident Study
DESTATIS	Federal Statistical Office
ELSAT	Hellenic statistical database
EU	European Union
EUSKA	Electronic accident type map (Elektronische Unfalltypensteckkarte)
FARS	Fatality Analysis Reporting System
FAT	Research Association for Automotive Technology (Germany)
GES	General Estimates System
GIDAS	German In-Depth Accident Study
HGV	Heavy goods vehicle
HIT	Hellenic Institute of Transport
ICAM	International Centre for Automotive Medicine (University of Michigan)
IDIADA	Institut d'Investigació Aplicada de l'Automòbil (Institute for Applied Automotive Research)
INTACT	Investigation Network and Traffic Accident Collection Techniques
IVI	Fraunhofer Institute for Transportation and Infrastructure Systems
LAB	Laboratory of Accident Analysis, Biomechanics and Human Behavior
MAIS	Maximum Abbreviated Injury Scale
MHH	Hannover Medical School
NASS	National Automotive Sampling System
NHTSA	National Highway Traffic Safety Administration
PDF	Portable Document Format
POLSAS	Police record system (Danish police case management system)
SCI	Special Crash Investigation
SQL	Structured Query Language
SUV	Sport utility vehicle

USA	United States of America
PTW	Powered two-wheelers
VDA	German Association of the Automotive Industry
VIN	Vehicle identification number
VIPA	Vulnerable Road User Injury Prevention Alliance
VRU	Vulnerable road user
VUFO	Traffic Accident Research Institute at TU Dresden GmbH
WHO	World Health Organization

1. **Introduction**

In the research and development environment of the automotive industry, numerous interdisciplinary aspects are considered, particularly in the field of vehicle and traffic safety. The relevant issues of assisted, connected, and automated driving and the further development of passive and integral safety systems require reliable data sources. The sometimes very heterogeneous traffic and accident situation between countries and continents makes it necessary to take as many as possible data sources from several countries/regions into account. However, it is not always obvious which data source is suitable for which kind of research question or development approach.

The aim of the present work is to develop a unified meta database that contains all necessary information for the research and development departments on a meta based level (no raw data) for several countries.

One of the main objectives is the research on international data sources in the field of traffic and vehicle safety. This includes national road accident statistics based on police accident data as well as highly detailed investigations in smaller regions (so-called In-depth data sources).

In addition to the development of a meta database, a questionnaire is used to check the applicability of the developed meta database for specific questions from the German automotive industry. For the objective assessment between the meta database and questionnaire a matching process is developed. The aim of the matching process is to indicate the covered percentage of necessary variables in the various data sources for each question. The results of the matching process are collected in a result matrix.

Based on the developed meta database and the matching process, the result matrix offers a possibility for an objective assessment of data sources and provides the opportunity for data providers to improve their data quantity and data quality. Furthermore, the meta database can act as a platform to bring several data providers from different countries together and to encourage the global harmonisation of traffic accident data sources.

2. Research on data sources

The research on data sources for road traffic accidents differs between national databases and in-depth databases. Regardless of the origin of the data (national or in-depth databases), the researched databases for this project are designated as data sources and the developed database as meta database (Figure 1).

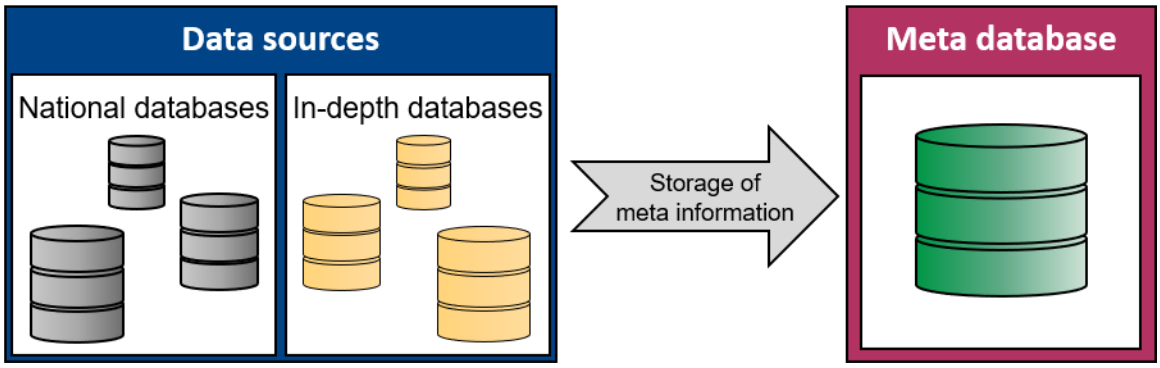


Figure 1: Scheme of the project with defined wording

The so-called national data sources are based on data collected by the police, which consist a large number of investigated accidents and give a macroscopic view of accident scenario. The use of a police reported data source on a national level allows a national coverage of the accident scenario and the acquisition of numerous cases.

The main focuses of the police investigation and the national data sources are the collection of general accident-related data to gather evidence for the later apportion of the blame for monitoring the general accident situation and assessing the infrastructure safety. In-depth data sources follow a different approach, where the data providers mainly want to investigate how the accident could happen without focusing without focusing on evidence for the later attribution of blame.

In comparison to the national data sources, in-depth data sources are mostly characterised by a smaller number of cases, but usually by a higher level of detail in the investigation of road traffic accidents. This allows a microscopic view of the accident scenario. In contrast to the police reported accidents, the accidents of the in-depth data sources are usually investigated by accident researcher and medical experts.

Depending on the sampling plan and the number of investigated cases, some in-depth data sources offer the opportunity to extrapolate their accident scenario on a national level to make representative statements.

At the beginning of the research project, it was agreed with the Research Association for Automotive Technology (FAT) that the detailed research on data sources should be focused on selected countries. The countries were selected as following: at least one representative European country per compass direction (north, east, south, west), plus Germany, and one non-European representative country. (Figure 2).

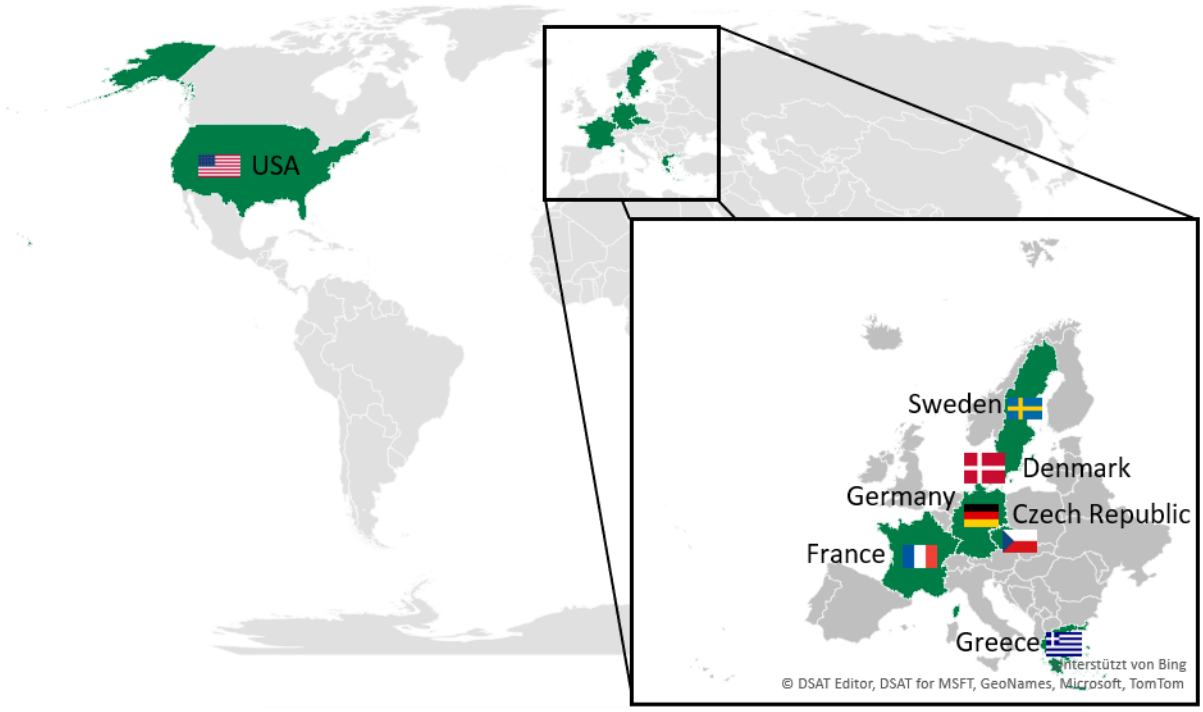


Figure 2: Selected countries for the detailed research on data sources

During the research process of data sources, the Swedish national authority explains that the access to the meta data of the police accident investigation is reserved only for Swedish research institutes. Subsequently, the research according to national data sources has been extended to Denmark to maintain the motivation of the country selection with one Nordic representative.

The complete selection contains Germany, France, Greece, Czech Republic, Sweden, Denmark and the United States of America (USA). In addition to the country selection, basic information about data sources in other countries should be included in the meta database.

2.1. National data sources

Based on the country selection, the following section investigates national data sources in particular. Basically, all data sources share a common parameter base, such as timeline, general accident properties, participant types and injury levels. However, various differences can be found on closer inspection[1].

The data sources are not only investigated at the level of the data scope but have to focus on the level of their meta data and data definitions as well. Therefore, the codebook, the general conditions of acquisitions, the context and the definitions have been reviewed. A first overview of the research data sources is given in Table 1.

Table 1 shows first differences of the national data sources e.g., the number of recorded variables or the definition of severity levels. These and other aspects will be discussed in the following sections.

Table 1: Investigated national accident data sources

Country	Germany (EUSKA)	Czech Republic (ASCZ)	France (BAAC)	Greece (ELSAT)	Denmark (Polsas)	EU (CARE)	USA (CRSS)	USA (FARS)
Geographical coverage	Federal states	Country				EU	Counties	Country
Accidents/year	300,000*	100,000	58,000	11,000	3,500**	1,300,000	50,000	33,600
Timeline	1990-today	1990-today	1975-today	1970-today	1990-today	1993-today	2016-today	1975-today
Number of variables	100	60	65	80	70	55	110	140
Conditions of survey	At least one victim, one vehicle, on the public road						At least one victim, one motor vehicle, on the public road	At least one fatality, one motor vehicle, on the public road
Severity definition	Time-based definition				Severity of the highest injury	Time-based definition and MAIS	KABCO scale	
Comment	Also cost estimations						Also cost estimations	Only fatal accidents. Also cost estimations

** As there is no global German police accident data source, the value "accidents/year" is from the DESTATIS table, only for traffic accidents with injuries. The other information are displayed based on the EUSKA format.*
*** At the time of the study, access to the police database was not granted: the value "accidents/year" is from the IRTAD database.*

[1] For the entire parameter set and the development of the data see the meta database (Section 3.1)

2.1.1. Geographical coverage

One of the first important aspects for national data sources is the geographical coverage. Most of them consist indeed of a coverage area that includes the whole country as well as the overseas territories. However, two countries (Germany and the USA) have a federal political organisation with different ways to set up a statistic strategy on road traffic accidents for the whole country.

In Germany, each federal state has to create and manage a police accident data source. However the German Accident Statistic Law (Straßenverkehrsunfallstatistikgesetz – StVUnfStatG) defines a minimum of to be collected accident information and their definition. This allows the German federal statistical office (DESTATIS) to unify and publish the road traffic accident data under an aggregated form.

Next to the minimum accident information, the data of each federal state provide further variables, especially the geographical coordinates, detailed information of accident causes or accident sketches. However, the additional information can differ between the states making an accurate and complete comparison difficult. For example, the accident data of Saxony reports the vehicle identification number (VIN) allowing to decode the vehicle manufacturer, model type, the model year or built-in assistance systems of motor vehicles. Further, the State of Baden-Württemberg records the usage of cycle-helmets or, if known, denotes the Average Annual Daily Traffic (AADT) at the accident site. Another example: Bavaria, Brandenburg, Saxony-Anhalt and Hessen have a more distinguished 3-digit-accident type, while Berlin uses raw motion lines for all participants to describe the accident constellation. In addition, Berlin provides 3d-scans of the accident site for accidents with severe and fatal injuries.

Another difference in the accident data can be the data format or the structure. Many databases define variables at an accident- or participant-level. In contrast, the State of Saarland records all information at the accident-level and thereby limits the number of accident participants to three. However, a format, which is used in the majority of the federal states is called EUSKA (Electronic accident type map) [1] and is inventoried in the meta database. EUSKA data includes the required information by the German Accident Statistics Law and allows for state-specific additional data.

Similar to the German federal states, each American federal state has to create and manage its own police accident data source. States are encouraged to follow the national agency recommendations of the National Highway Traffic Safety Administration (NHTSA) for the data collection, so that all states have the same parameters, but it is not mandatory [2]. In order to create a unified and single accident database, the NHTSA has two different databases, which are not compiled in the same way: one is an extrapolation of selected counties to obtain an estimation for all accidents (with/without personal injury), and the other one is an (almost) complete survey of fatalities.

The Crash Report Sampling System (CRSS database) receives its data from a nationally representative probability sample [3], selected from the different police-reported accidents that occur annually and result of property damage, injury, or death. These accident reports are chosen from 60 selected areas across the United States that reflect different properties: the geography, the

population density, the miles driven, and the number of accidents. CRSS analysts review the original accident reports, interpret and code their information in the common database. After the coding phase, quality checks are performed on the data, both automatically and manually to ensure validity and consistency. When these are completed, CRSS data files become available and exploitable, for road experts as for the public. Based on these representative samples, accident trends are estimated at a national level after an additional extrapolation.

The Fatality Analysis Reporting System (FARS database) only contains traffic accidents with fatalities from the 50 federal states, the District of Columbia and Puerto Rico.

Contrary to the CRSS database, it is not based on representative samples. Since the NHTSA managed to have a cooperative agreement with an agency in each state's government to provide information on all qualifying fatal accidents. In theory, all fatal accidents involving at least one motor vehicle are reported. FARS analysts present in each state are responsible for gathering, reviewing and coding the original data to the FARS format. The number of analysts varies by state, depending on the number of fatal accidents and the ease of obtaining data. Data is not only based on the accident database, but also on the vehicle registration and driver licensing files as well as medical data, including e.g., death certificates and hospital data.

Due to data protection regulations, and as they apply to German data sources, no personal identification data such as names, addresses or national insurance numbers are available in the US data sources either.

2.1.2. Data source size

A second important aspect for national data sources is the size of the data source, e.g., the investigated number of accidents per year. Most of them have more than 10,000 cases per year, which allow them to reach a high degree of representativity.

2.1.3. Database history

In Europe, the national accident data sources are usually more than 30 years old since the European Commission stated the creation, management and use of accident databases with its 1993 regulation [5]. Due to these regulations, the last European countries are obliged to set up an accident database.

In the United States, the first national accident data source GES (General Estimates System) appeared in 1988, and the FARS data appeared in 1975. However, following an overhaul of the variables and structure, the current version is from 2016. The previous national data source GES is now called CRSS and the previous in-depth data source NASS CDS (Crashworthiness Data System) is now called CISS (Crash Investigation Sampling System). Only the FARS database was not officially changed, although it was also modified.

2.1.4. General and specific content

In general, national data sources mainly provide information on data related to the place of the accident, the time when the accident occurred, at least the vehicle at fault (such as defined by the police officer) and at least the driver of this vehicle.

Thus, all sources contain data on time, weather, light, type of road, the area, crossing with regard to the place of the accident and its temporality. There are also data on the type of vehicle, the obstacle hit (if any), the type of accident, the age of the driver and the type of injury. However, there are differences in the level of detail. On this point, the number of variables in Table 1 gives an idea how detailed the data sources can be. For example, the Greek data source ELSAT (Hellenic statistical database) has several variables registering the road infrastructure with the presence of different road markings in the middle and on the side, with the mention of safety barriers and separation in the middle, and with the existence of special lanes, like a bicycle lane or a bus lane [6].

A specificity in the France data source BAAC is to record the driving manoeuvre of each participant instead of a pre-categorized accident constellation between the accident participants. Additionally, a special motorbike manoeuvre is distinguished: lane splitting (e.g., in traffic jams, in slowdown traffic on a two-lane road) [7].

Due to the high modal split of cyclists, the Danish data source POLSAS (police case management system) records specific information about the infrastructure related to cycle infrastructure [8]. Information can be provided on the existence of a cycle path at the accident scene, how it is designed and how it ends at the intersection. For example, at a T-intersection, the cycle path that was previously on the pavement, separated from the road, joins the road and disappears into the lane for cars.

One special feature of the Czech data source ASCZ (Accident Statistics Czech Republic) is that the car manufacturer of the vehicle responsible for the accident is registered. There are about 60 different manufacturers to choose from, depending on the type of vehicle (bus, car, motorbike, etc.) [9].

Another special feature of the Czech data source ASCZ as well as the German data sources is the cost estimation of the material damage in each accident. In both countries, the majority of cost estimations are made by the police officers and based on a visual analysis of the material damage.

The structure of the FARS and CRSS databases provides a chronology of harmful events, such as an impact with another vehicle, or an impact with a fixed object, or a fire in a vehicle (about 60 different events). It is possible for a given accident to record a list of so-called harmful events and order them from first to last in time. In this list, the most harmful event can also be coded. This is used to assess the most probable causes of the accident. Moreover, all collision events are recorded not only according to their temporality, but also according to the involvement of different actors. This is particularly useful in accidents with at least three vehicles. For example, a collision

event will be recorded first between vehicle A and B, and then between vehicle B and C, in the case of rear-end collisions.

2.1.5. Conditions of survey

For the European countries, the conditions of survey of an accident in the data sources are the same: the accident must involve at least one vehicle, and one victim on the public road.

In the US data sources, at least one motor-vehicle must be involved in an accident. In CRSS, there must be at least one motor-vehicle involved in an accident and at least one injured person. In FARS, there must be as well at least one motor-vehicle involved in an accident and at least one participant, which must have died within 30 days. Thereby, VRU accidents are only recorded if at least one participant has died according to the above conditions and at least one motor-vehicle was involved. Fatalities of single-VRU-accidents or accidents between VRU are not included [4].

2.1.6. Definition of the injury severity

Most of the police accident data sources in Europe use a time-based definition of the injury severity. Casualties are considered as fatalities if they die within 30 days after the accident. Participants are considered as seriously injured if they stay in a hospital more than 24 hours (stationary) after the accident, and slightly injured under 24 hours (ambulant) or if the casualties are treated by the rescue service directly (only) at the accident scene. However, this method is being questioned because some cases go unrecorded, for example single accidents of cyclists with treatments several days later [10]. Starting in 2013, various reports [11] [12] [13] have highlighted the weaknesses of a time-based definition and have advocated the use of a standard format for all, suggesting the trauma scale “Maximum Abbreviated Injury Scale” (MAIS). Although the Community Accident Report Database (CARE) stores the injury severity also under a time-based definition since its creation, the MAIS was newly implemented in 2015.One drawback of this new variable is that the completeness of its coding strongly depends on whether the member countries also report injuries under this scale.

The danish police database POLSAS proceeds a different way in the recording of the severity level. A "slight" degree is to be chosen if the present injury, or the condition of the person corresponds to a specific list. A "severe" level is to be chosen if the condition of the person corresponds to another list, e.g., if there is a visible fracture, dislocation or severe sprain. If a participant has more than one injury, the highest level is retained.

Both CRSS and FARS list their injuries according to the KABCO scale. This scale was created in 1966 and adopted by the states, although the definitions were left up to the state's discretion. While working on a common accidents criteria list, the scale was reviewed and the definitions settled in 2017 [13]. It is now composed of six different levels of injury severity, observed at the scene of the accident:

- K – Fatal injury;
- A – Suspected serious injury;
- B – Suspected minor injury;
- C – Possible injury;
- O – No apparent injury.

So-called serious injuries are defined as result in one or more of the following: severe laceration resulting in exposure of underlying tissues/muscle/organs or resulting in significant loss of blood, broken or distorted extremities, crush injuries, suspected skull, chest or abdominal injury, significant burns, unconsciousness or paralysis. The definition of level B is therefore based on whatever is lighter than level A, and level C is reserved for non-visible injuries, but still complaints of pain. Level O is for non-injured participants.

2.2. In-depth data sources

In-depth accident data sources are mainly investigated by interdisciplinary research units. They often consist of a technical team supplemented by a medical and/or psychological unit. The former investigates the technical aspects of the accident (e.g., infrastructure, vehicle) and the medical part investigates the injuries suffered by the persons in the accident and the physical/psychological backgrounds as well as the mental consequences caused by the accident.

Most data providers of in-depth data sources reconstruct their accident scenario in order to know the most probable accident sequences. Due to the increased effort in the data recording and reconstruction, the number of accidents is smaller than in national data sources.

In-depth investigations are mainly realised at selected locations where either more accidents occur or a selected type of road users are present or the boundary conditions like landscape, driver milage or age distribution are representative for a certain region or country. Based on a fixed sampling plan with a random accident selection and a defined weighting process, some data providers have the possibility to extrapolate their data for representative statements to the investigation area or even to certain regions.

According to the defined country selection, the following in-depth data sources have been researched in Europe (Figure 3).

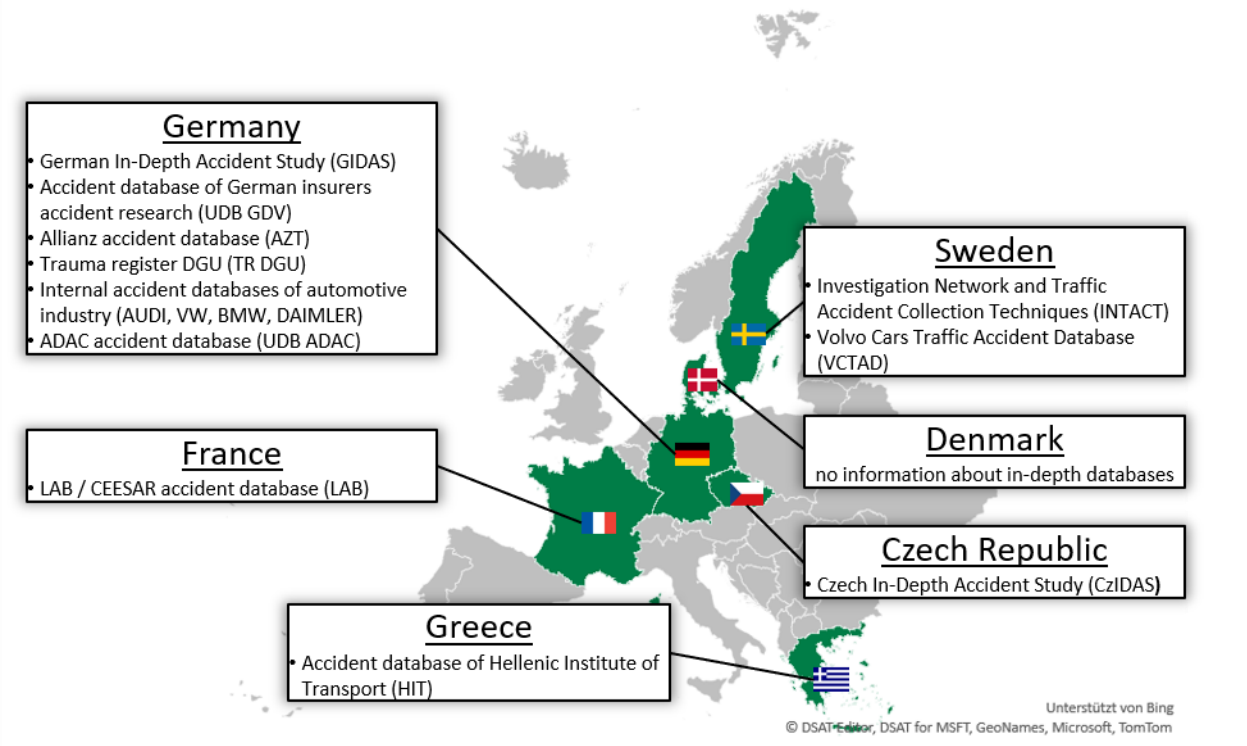


Figure 3: Research on in-depth data sources for the country selection in Europe

The accident data recording in Germany is operated on an in-depth level by various data providers. One of the most significant data sources for in-depth traffic accident research is the German In-Depth Accident Study (GIDAS).

GIDAS is a collaborative project of the Federal Highway Research Institute of Germany (Bundesanstalt für Straßenwesen / BASt) and FAT. Since mid-1999, the GIDAS project has collected on-scene accident cases in the areas of Hanover (by the MHH – Hannover Medical School) and Dresden (by the VUFO - Traffic Accident Research Institute at TU Dresden GmbH). The data collected in the GIDAS project is very extensive and serves as a knowledge base for various groups of interest [27]. The main feature of GIDAS is that the investigated data are representative for Germany due to the sampling plan for the accident selection and the weighting process [14].

Parallel to the access to the GIDAS data source, most of the automobile manufacturers in Germany also operate their own accident investigations. The focus of the internal accident research units is to investigate the accident scenario and accident behaviour of their latest vehicle models. Findings from this work are directly returned to the development departments for eventual adjustments to the current models and as baseline for new developments [15] [16] [17]. Further in-depth data sources are operated by insurance companies (e.g., German Insurers Accident Research UDV, Allianz) and automobile clubs (e.g., ADAC)

In 2007, a Swedish consortium of the government department for transport, the Chalmers University of Technology and the automotive industry set up the Investigation Network and Traffic Accident Collection Techniques (INTACT). The consortium was concerned with the establishment of a methodology for the in-depth investigation of road traffic accidents around the Gothenburg area, in which at least one of the following vehicle types must be involved

- Car (all categories)
- Light truck (≤3500 kg)
- Heavy goods vehicle (HGV) (>3500 kg)
- Bus (all categories)

Another selection procedure is when an ambulance is called to the accident scene due to assumed personal injury. The final report of this consortium was published in 2010 and involved in total 123 investigated accidents on the spot [18]. Further in-depth accident data sources are operated by Volvo Car Corporation [19] (Volvo Cars Traffic Accident Database - VCTAD) and Volvo Trucks [20] to the accident investigation of their latest vehicle models.

The research on the availability of in-depth data sources in Denmark did not provide any information.

Based on the knowledge and structure of GIDAS, the Transport Research Centre (CDV) from the Czech Republic, Skoda and IDIADA CZ set up the Czech In-Depth Accident Study so-called CZIDAS in 2011. CZIDAS collecting accidents on two investigation spots (Brno, Hradec Králové). Every year nearly 300 accidents are investigated. In addition to the investigation each case is reconstructed, and the participants are interviewed according to a psychological scheme [21].

The investigation of road traffic accidents at a microscopic level in Greece is operated by the Centre for Research and Technology Hellas (CERTH) or rather by the Hellenic Institute of Transport (HIT). The investigation is conducted around the site of the institute in Hellas [22].

In France, in-depth investigations for road traffic accidents are realised by the Laboratory of Accident Analysis, Biomechanics and Human Behavior (LAB) and CEESAR (Centre Eurpéen d'Etudes de Sécruité et d'Analyse des Risques) [23]. A request for meta information by codebooks or manuals has been denied. After a publication of the project results, a decision will be made whether meta information on the data source will be disclosed or not.

The research on in-depth data sources in the USA yielded in three data sources (Figure 4). The representative share of U.S. road accident data is government-owned and subject to the U.S. Department of Transportation (DOT). The NHTSA is the central authority for the investigation of road traffic accidents and data.

The National Centre for Statistics and Analysis (NCSA), a department within the NHTSA, is responsible for the preparation and analysis of statistical accident data. NCSA is the superordinate authority for various data collection programmes, whose research content varies between national and in-depth.

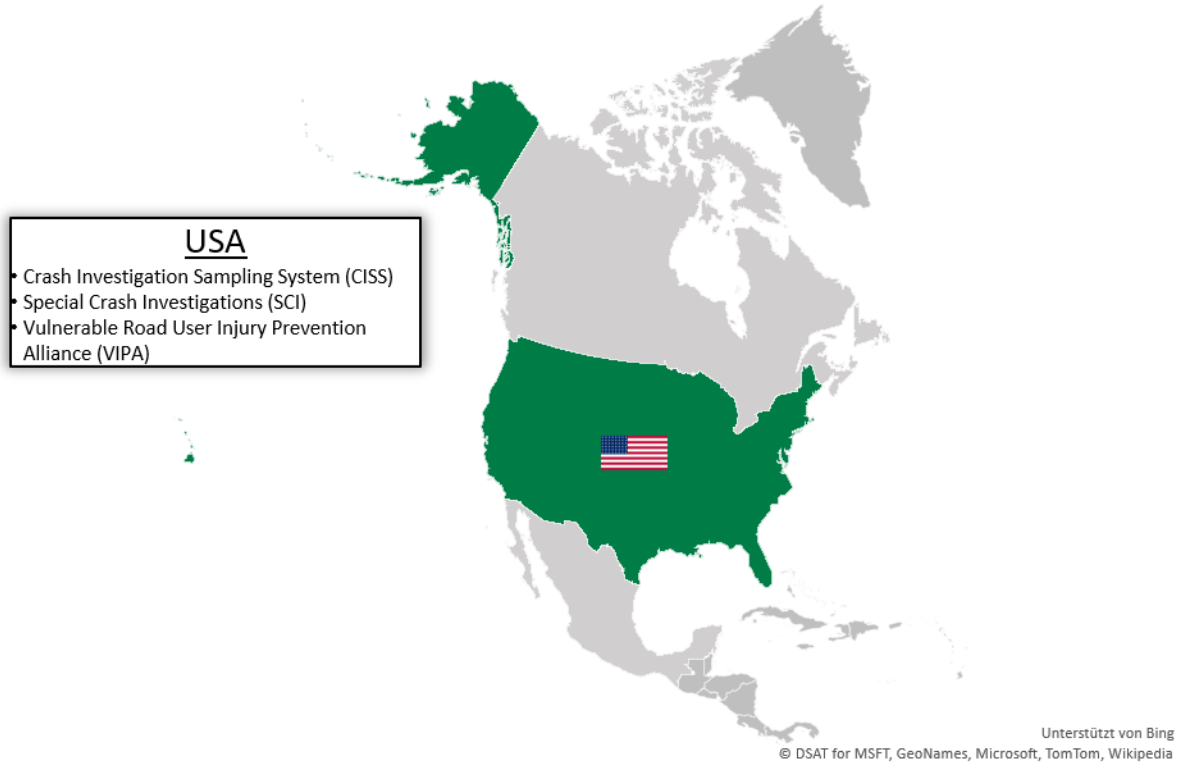


Figure 4: Research on in-depth data source in the United States of America

One of these data recordings with more in-depth content on road traffic accidents is the Crash Investigation Sampling System (CISS). CISS is based on the Crashworthiness Data System (CDS) of the National Automotive Sampling System (NASS), which was withdrawn in 2016 and replaced by CISS. The accident investigation is carried out by in-depth research units. The case selection is based on thousands of randomly selected traffic accidents from police accident reports, which are supposed to be representative for the USA in terms of geography, population, mileage and number of accidents. CISS only investigates accidents involving passenger cars, light trucks, sport utility vehicles (SUVs) and vans, in which at least one of the involved vehicles is towed and at least one of the involved persons is injured [24].

The Special Crash Investigations (SCI) is considered to investigate special accident scenarios. The cases for the SCI data source based on the data from the GES/CRSS, CDS/CISS and FARS. Depending on the research contract the cases are investigated for the SCI data source more in detail. The aim of these investigations is to publish a case number of at least 100 representative accidents per study. In the past, the SCI investigated accidents caused by carbon monoxide poisoning or cases in which babies or small children were left in the car and suffered health damage due to hyperthermia (overheating) or hypothermia. Other SCI studies focused on accidents involving vehicles where special airbags, child seats or alternative powertrains are used. Another SCI study investigated the accident scenario of school buses where at least one occupant was fatally injured [25].

The Vulnerable Road User Injury Prevention Alliance (VIPA) was founded by the International Centre for Automotive Medicine (ICAM) of the University of Michigan. The collection began in 2015, producing a detailed database of Michigan pedestrian and bicyclist accidents where police were called to the accident scene. The data of the VIPA database are not representative for the USA but should provide a rough impression of the accident scenario between VRUs and vehicles of U.S. vehicle fleet [26].

For the meta database the CISS database and SCI database are completely researched. The VIPA is registered based on scientific publications, because no response was received upon the request for access to the manual or codebook.

3. **Development of the meta database**

The basis of the meta database is the GIDAS structure, which is characterized by several levels, starting with the accident level, participant level, person level and the lowest level, the injury level. These four levels are linked by a CASE-ID, which is a consecutive number within a year and the respective investigation area [28].

For the development of the meta database, fragments of the GIDAS structure as well as knowledge and expertise in the database development are used [29]. Some of the GIDAS content, which is represented by an entire data table (e.g., tire data table, trailer data table) is limited to the essentials and is described for the meta database by one variable (e.g., tire data, trailer data).

3.1. Structure and implementation

In general, the meta database only contains the existence of certain information in the data sources through a binary coded variable. Raw data of the data sources are not included.

The main structure of the meta database is divided into two categories (Figure 5). The fact sheet information on data sources are stored in the part “GENERAL INFORMATION”. Information on the accident scenario, infrastructure data or vehicle information are inventoried in the “CONTENT” part of the meta database. In general, each area consists of several tables, which are linked to each other by a unique identification number (primary key) called “SOURCE_ID” and starting with the number 1. Each data source has its own “SOURCE_ID“ and is assigned continuously without restriction (i.e., by country/region or number of accidents per year).

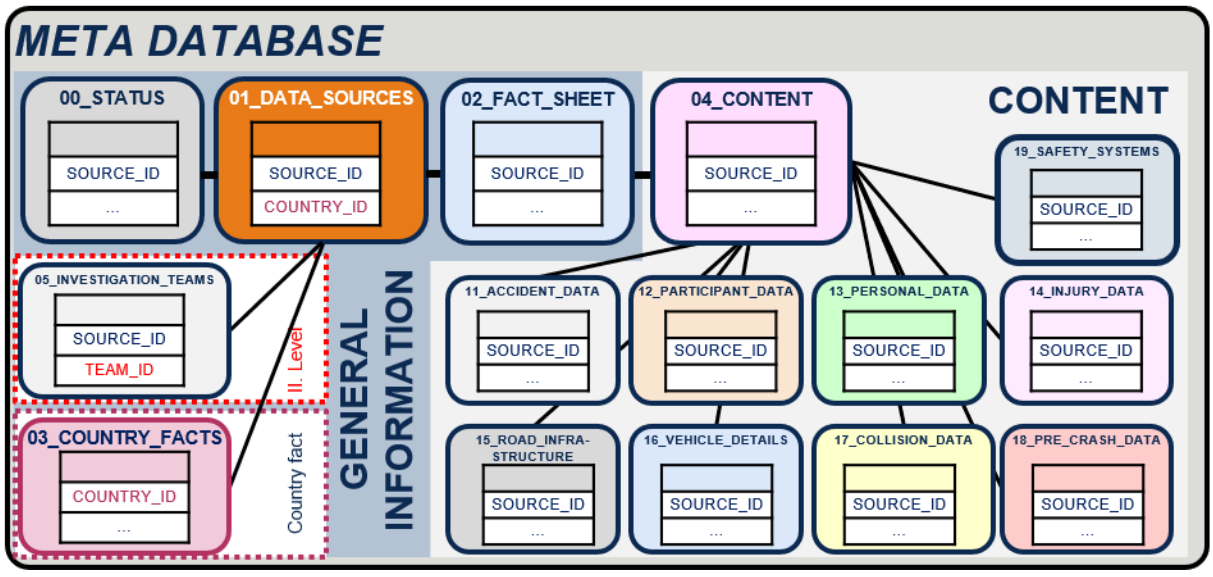


Figure 5: Schematic structure of the meta database

The category “GENERAL INFORMATION” is mainly characterized by the tables “01_DATA_SOURCES” and “02_FACT_SHEET”. The first mentioned table contain all important information on the data source, for example in which country the data source is located and its data provider. In addition to the type of data source (national or in-depth data source), this table contain information on the database format and its access possibilities.

In comparison to table “01_DATA_SOURCES”, the table “02_FACT_SHEET” contains mainly information on the data size (e.g., cases per year), the coverage (e.g., investigation area) and the beginning or ending of the accident investigation. It also provides an assessment of the representativity of the data source as well as contact details.

The assessment of the representativity is often based on mathematical methodologies that compare the investigated accident scenario with a higher-level or specific accident scenario (e.g., national data source). Based on the known sampling methodologies for the accident investigation of the individual data sources, the representativity of the data sources can be assessed. Three types of representativity are considered:

- data sources which achieve representativity by full sample survey (mainly police-recorded data),
- data sources with large sample sizes and parameter ranges which allow for extrapolation and thereby representative results and
- data sources with limited representatives.

Information on country-specific key figures is stored in table “03_COUNTRY_FACTS” and is linked to the table “01_DATA_SOURCES” by the “COUNTRY_ID”. Each country and each continent has its own “COUNTRY_ID“. In addition to the number of inhabitants, the table also include the number of all road traffic accidents. The researched accident numbers based on national statistics or the latest global status report on road safety from the World Health Organization (WHO) [32]. The published year of the statistics is also included.

Furthermore, the following (annual total) numbers are stored here:

- Accidents with personal injuries
- Accidents with fatalities
- Total number of fatalities in road traffic accidents

In order to ensure comparability between countries, the death rate per one million inhabitants and the year of data origin are given in the country fact table.

The methodology and which investigation tools are used for the accident collection are stored in the table “04_CONTENT”. The table “05_INVESTIGATION_TEAMS” provides the opportunity to add further information on the investigation units by the user. Furthermore, the table “05_INVESTIGATION_TEAMS” is the only table with a second primary key (second data level), because one data source can be investigated by several investigation units (e.g., GIDAS, CIDAS). Based on the table “02_FACT_SHEET”, the table “05_INVESTIGATION_TEAMS” collects general

information on the research units, such as the coverage, the accidents per year or investigation period (start or end of the investigation).

The tables with general information are directly connected to the “CONTENT” of the data source, which is based on a similar structure as GIDAS. The table “11_ACCIDENT_DATA” includes all characteristics of the accident which are necessary to answer general questions on the accident scenario. Further information on the accident participants and the kind of road users (e.g., vehicle or pedestrian) are stored in the table “12_PARTICIPANT_DATA”.

The table “13_PERSONAL_DATA” contains general information on the persons that are involved in the accident, independent of the kind of road users. In addition to the physical characteristics of persons, the table “13_PERSONAL_DATA” contains information on the usage of a seatbelt, a helmet usage or worn protective clothing.

Information on data of single injuries are stored in the table “14_INJURY_DATA”. Parallel to the information on the single injury types, the injury localization and the injury causes, the injury severity according to time-based definition and the abbreviated injury scale (AIS) are stored. The definition for the time-based injury severity is described in the section 2.1.6 - Definition of the injury severity”.

The AIS is an anatomical-based coding system and represent the threat to life associated with the injury. The entire AIS code is defined by the type of injury, the location, and the injury severity. The AIS injury severity code is determined by a scale of one to six, whereby AIS=1 represents a minor injury and AIS=6 represents an untreatable injury [30]. The coding of the AIS according to various update year (e.g., AIS 1998, AIS 2008, AIS 2015) is also included in the meta database.

If data sources contain information on infrastructures, road types or road conditions it is entered in table “15_ROAD_INFRASTRUCTURE”. Information on view obstacles, traffic signs and speed limit is also stored in the road and infrastructure table.

The table “16_VEHICLES_DETAILS” contains all general vehicle information. Specific information on the different vehicle types is stored as one variable per vehicle type (e.g., car details, truck details, PTW details). In addition to the general information of the vehicles, information on damage and deformation is also stored in the vehicle table. Damage to the vehicles is determined by visual inspection. In contrast, deformation is based on measured values that indicate the depth or bulge of the damage.

Parameters to evaluate the crash constellation are mainly based on reconstruction data of the accident scenario. Typically, for in-depth data sources, the most accident scenarios are rebuilt by computer-based software. Consequently, important information for the pre-crash and in-crash phase are generated. The information on the existence of such reconstructed data is stored in the table “17_COLLISION_DATA”. Describing variables for the crash constellation as well as information on the initial speeds or collision speeds are also part of this table.

For the evaluation of active safety systems, information on the pre-crash phase is necessary. After the reconstruction of the accident scenario, time-based parameters and data are generated via simulation tools to describe the pre-crash phase more in detail. The information on the

existence of pre-crash data is stored in the table “18_PRE_CRASH_DATA”. If a data source contains video sequences of the accident scenario, this information is also stored in the pre-crash table.

Information on installed active or passive safety systems in vehicles are important for the assessment and evaluation of such systems. Accordingly, the information on the existence of vehicle safety systems in a data source is stored in the table “19_SAFETY_SYTSEMS”. In addition to the information on existence of safety system, it is also important to know whether the safety systems were activated or not during the crash phase. Information on the activation status of safety systems is also stored in the table “19_SAFETY_SYSTEMS”.

Based on the above explained schematic structure, the meta database structure is transferred into a Microsoft® Access® database format (.mdb) (Figure 6). Microsoft® Access® is a database management system for medium-sized databases and can be accessed via queries based on the Structured Query Language (SQL) or other programming languages.

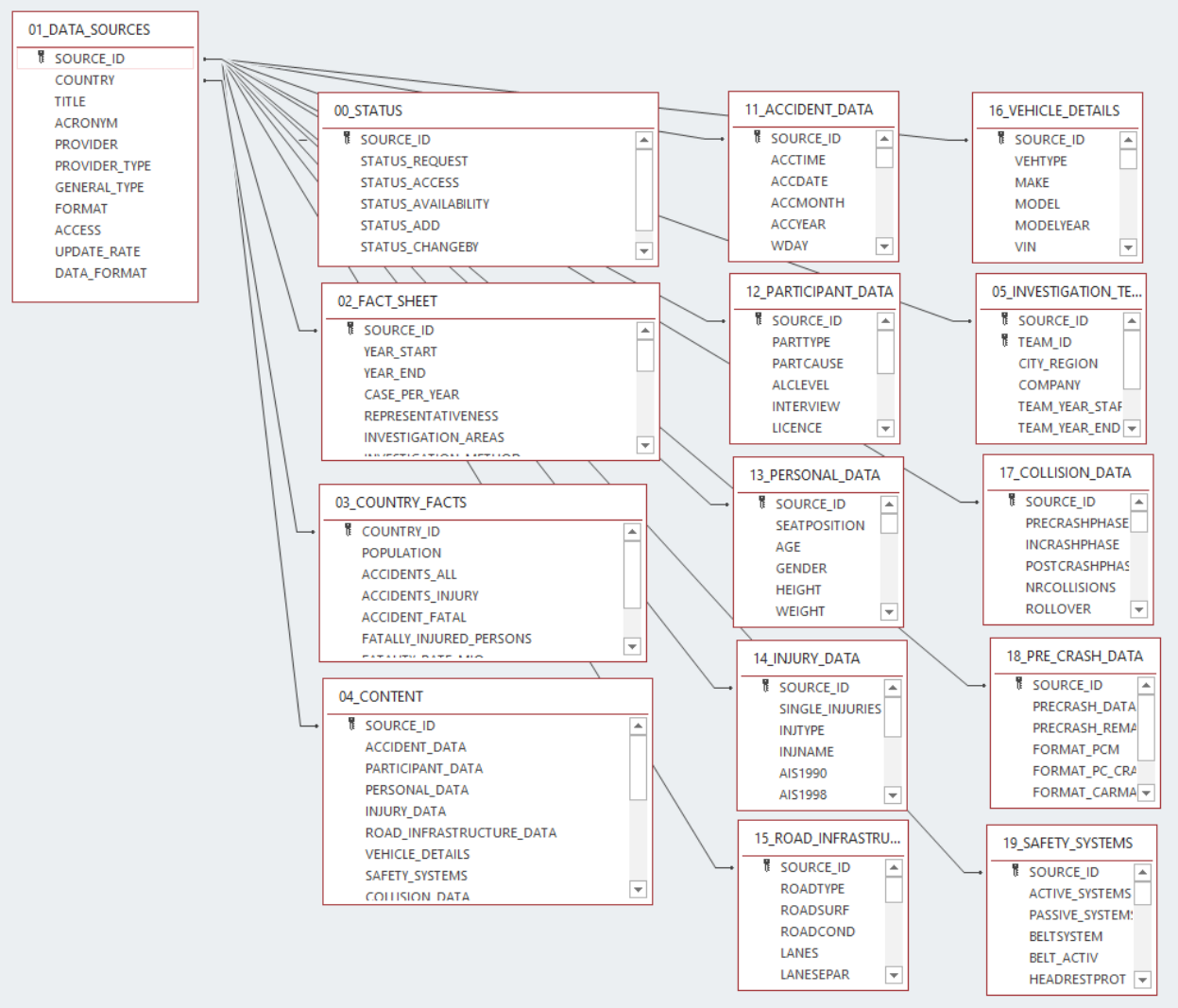


Figure 6: Structure of the meta database in Microsoft® Access®

Information about the researched data sources could be entered by using an input mask, which is directly linked to all data tables. Furthermore, this mask services as output formular (Figure 7).

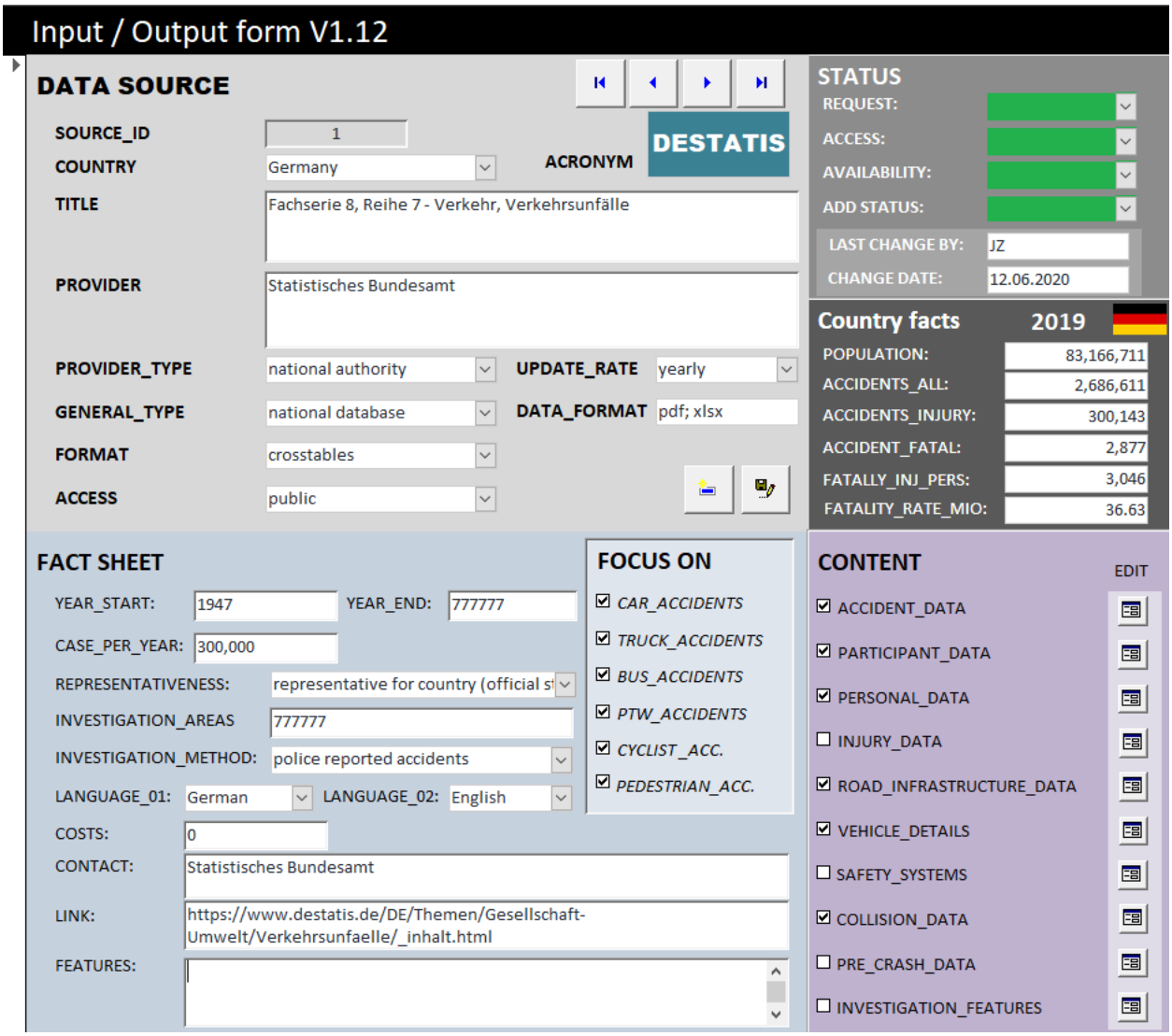


Figure 7: Input / Output mask for the researched data sources

The main page of the input / output mask shows the general information on the researched data sources (light grey box) and the associated fact sheet information (blue box). In addition to the status box (upper right box), the country facts below provide information on the accident scenario by each country according to latest researched statistic. The content box (violet box) gives an overview of the available content for each data source.

Several input masks on the lower right side of the content box offer the possibility to enter data in the specific content table (e.g., “10_ACCIDENT_DATA”, “11_PARTICIPANT_DATA”). Each button is directly linked to a separate input mask for the respective input table.

In total, the meta database contains 15 data tables and 10 in-/output masks. Parallel to the meta database development, a codebook for the description of 237 variables contained in the meta database has been developed.

3.2. Codebook

For the description of the tables with several variables and parameters a codebook is set up. The basis of the codebook is the GIDAS codebook [28] with the following three tables/categories

- RECORDS,
- VARIABLES and
- LABELS.

The table “RECORDS” contain the descriptions of all the tables in the meta database. The content of the tables is defined by variables and these are described in the “VARIABLES” table. The specifications of the variables are called labels, and these are explained in the “LABELS” table.

The connection between the tables is not directly linked by one common primary key (Figure 8). Each table contains its own primary key, which can also be found in the subordinate table.

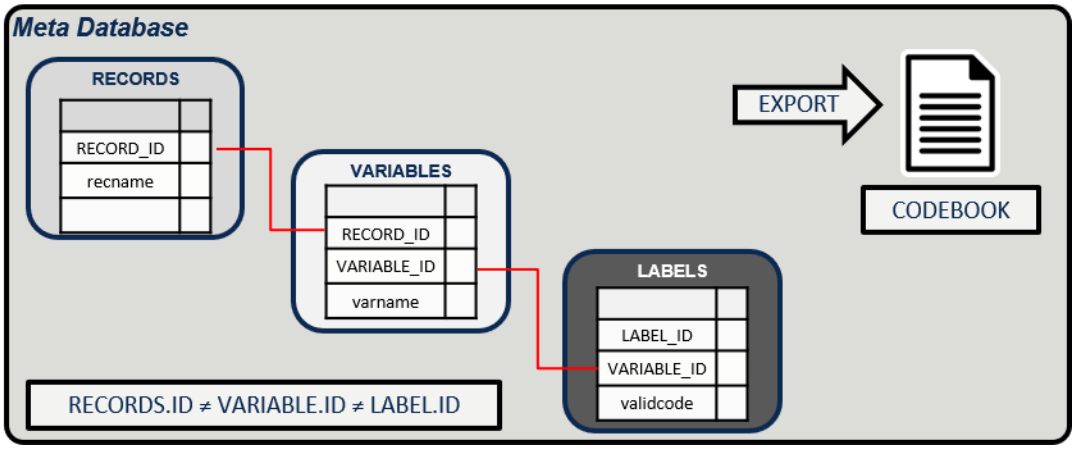


Figure 8: Structure of the codebook

Most variables in the meta database are defined by a binary code. The binary code 1 describes that an information is available in the data source. In contrast, the binary code 0 describes that an information is not available in a data source. Furthermore, some variables are more specified by further labels because the anticipated objective quality assessment for a data source based on the existence of an information alone is insufficient (e.g., representativity).

Variables, that must be extended either in the description or by further labels, are mainly identified in the answering of the questionnaire. For this purpose, the top 50 variables are described in more detail or are extended by additional labels in the codebook. The top 50 variables are determined by the number of times a variable is used to answer the 197 questions (Table 2). Superordinate variables (e.g., ACCIDENT_DATA, PARTICIPANT_DATA) are not included into the list.

Table 2: Top 50 variables from the answering of the questionnaire

Ranking	Variable	Short description	Number of binary code 1 (necessary for the answering)
1	INITIALSPEED	Initial speed	37
2	ACCTYPE	Type of accident	34
3	COLLSPEED	Collision speed	34
4	ACTIVE_SYSTEMS	Active safety systems available	25
5	ACCSEVERITY	Overall/Maximum accident severity	24
6	VEHTYPE	Vehicle type	23
7	PRECRASHPHASE	Information on the pre-crash phase	23
8	PARTTYPE	Type of participant	21
9	MANOEUVRE	Maneuver of the participant before the accident	21
10	CDC	Application of force during the crash (CDC coding)	21
11	CAUSATION	Information about (main) accident causes	19
12	COLLSIDE	Crash side of the vehicle or pedestrian	18
13	REPRESENTATIVENESS	Representativity	17
14	NRCOLLISIONS	Number of collisions	17
15	AEB_SYSTEM	Autonomous emergency braking	17
16	INJ_SEVERITY	Injury severity according to OEDC definition	15
17	COLLPOINT	Collision point (measured value)	15
18	CRUISE_CONTROL	Cruise control system	15
19	LDW	Lane departure warning	15
20	SINGLE_INJURIES	(General) availability of data for single injuries	14
21	INJCAUSE	Injury caused by	14
22	PASSIVE_SYSTEMS	Passive safety systems available	14
23	BSM	Blind spot monitor	14
24	ACCKIND	Kind of accident	13
25	DECELERATION	Deceleration or Acceleration	13
26	EES	Energy Equivalent Speed	13
27	DELTAV	Differential velocity	13
28	ACCTYPEA	Participant A according to accident type	12
29	ACCTYPEB	Participant B according to accident type	12
30	INJTYPE	Injury type	11
31	ESC	Electronic stability control	11
32	EDR_DATA	EDR information	10
33	CAUSER	Main causer of the accident	10
34	INJ_SEVERITY_MAIS	Injury severity according to MAIS	10
35	ABS	Anti-lock braking system	10
36	AEB_ACTIVE	Autonomous emergency braking active	10
37	PHOTOGRAPHIC_DOCU-MENTATION	Photographic documentation	9
38	LOCATION	Accident location	9
39	PARTCAUSE	Accident causations for each participant	9
40	INCRASHPHASE	Information of the in-crash phase in an accident	9
41	ATTENTION_ASSIST	Attention assistance	9
42	AUTOMATION_LEVEL	Automation level	9
43	CASE_PER_YEAR	Numbers of accidents per year	8
44	SEATPOSITION	Seat position / seat adjustments	8
45	VRUIMPACT	Impact of VRUs	8
46	INJCOLL	Injury caused collision	8
47	MAKE	Vehicle brand	8
48	MODEL	Vehicle model	8
49	BODYTYPE	Body type of the vehicle	8
50	CRASHWEIGHT	Crash weight of the vehicle	8

The list of the top 50 variables shows those variables that are currently requested more frequently in the questionnaire compiled by the German automotive industry. An extension of the questionnaire with further questions from other stakeholders could result in a different list of the top 50 variables.

In addition to the description of the variables, logical links of variables to other variables are also included. For example, a data source might not contain explicit information on the light conditions (day or night), but this information could be derived implicitly on the basis of accident time and date. Consequently, the variable light conditions" could still be selected as available" (binary code 1).

The codebook could be exported into a printed document via a programmed script to enable the user to work with both codebook and meta database at the same time. For this purpose, records and variables are declared as headings to facilitate the navigation. The chosen document format for published codebook version is the Portable Document Format (PDF).

4. **Objective assessment**

The development of the meta database provide a tool to work more effectively on existing research questions by suggesting suitable data sources to answer these questions. At the beginning of the meta database development a member-only, anonymized questionnaire with more than 170 questions from the German automotive industry has been set up. For the comparison of these questionnaire with the meta database, a methodology is developed. The main goal of the methodology is to generate the appropriate data sources for the respective questions. In addition, the questionnaire serves as support in identifying important variables and labels for answering the research questions.

Note: Answering the research questions in a scientific way is not part of this project and remain to the responsibility of the questioners. The comparison serves as guideline which data sources can be used to answer the respective question most effectively. The coding of the respective questions is based on the expertise of the processor.

4.1. Questionnaire

The first step, the questions of the questionnaire have to be inventoried. Therefore, a consecutive identification number called “QUESTION_ID” is implemented. Afterwards the questions are categorized as follows:

- Research question
- Fact sheet question (fact sheet information)
- Other / analysis question

Based on the existing research questions, superordinate questions (main research questions”) are formed in order to categorize these research questions with a similar subject (e.g., vehicle safety systems, autonomous driving, accident scenario of different kinds of road users). During the processing time the questionnaire is extended by the main research question to 193 questions. The distribution by the question categories is shown in Figure 9.

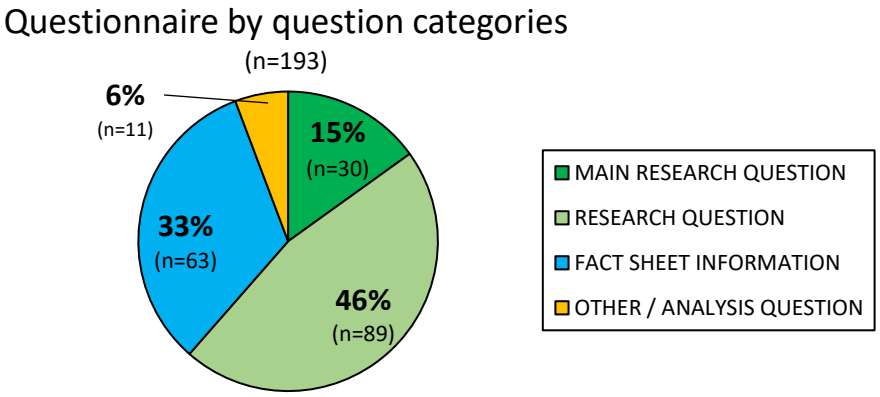


Figure 9: Distribution of the questions by categories

The combination of research questions and main research questions represent more than a half (>50%) of the questionnaire. One third of the questionnaire (33%) are questions to general information on data sources.

The category “other / analysis questions” contains questions that do not fit to the three other categories or that require further data sources with another background than road traffic accidents to answer these questions. This category accounts for 6% of the questionnaire.

In the following step, the questions are analysed. The syntactic analysis of each question analyses word by word and translate them into variables. Thereby, variables that are necessary to answer certain questions but are not contained in the meta database are added to the meta database. The semantic analysis of the questions identifies variables that are not directly mentioned in the question but are related to the question. Questions about driving manoeuvres before an accident automatically include information on the type of accidents and the participants according to the accident type, although they may not be directly mentioned in the question.

After categorizing the original questionnaire, few research questions had to be clarified by the FAT members. Therefore, the questionnaire is returned to the FAT members by the latest compliance guidelines. This iteration loop not only offers the opportunity to clarify some questions, but also to generate feedback and create a basis for discussion on further development approaches.

After the iteration loop a variable set with 237 variables is fixed. Based on these variables, each question of the questionnaire is linked to the respective variables. Therefore, a binary code (0/1) is used. The binary code 1 is coded, when the variable is necessary to answer the question. All other variables are coded by the binary code 0.

Based on the following main research question (QUESTION_ID = 207), the syntactic and semantic analysis is presented.

Which effect does the sitting position have on the injury severity?

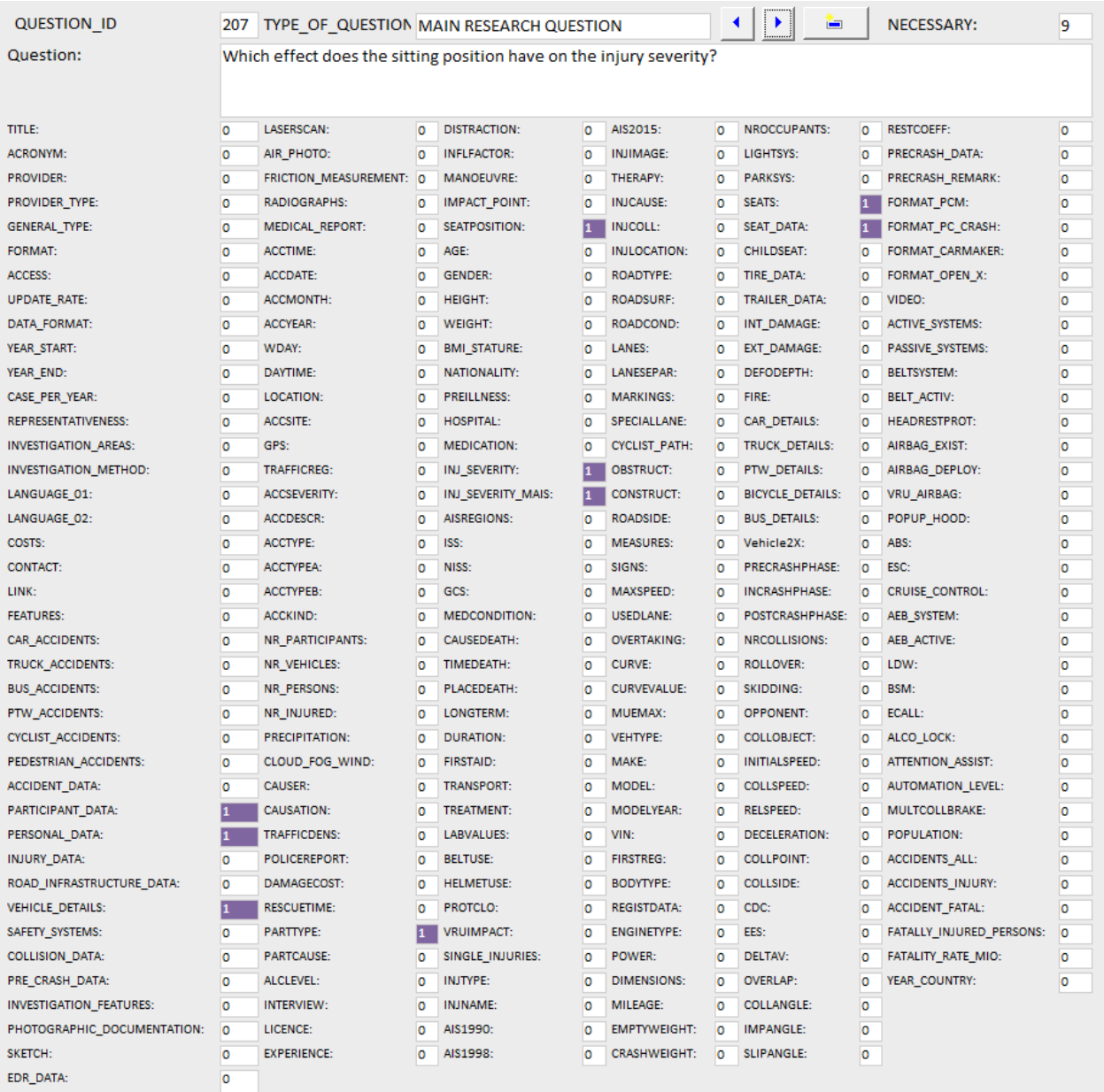
The syntactic analysis of the question indicates that the sitting position (SEATPOSITION) and information on the maximum injury severity of the occupants according to the time-based definition (INJ_SEVERITY) of the injury severity or to the AIS code (INJ_SEVERITY_MAIS) are necessary variables from a data source to answer the question.

The semantic analysis of the question indicates that for a robust assessment of the injury severity, it is also necessary to know, which occupant was sitting on which seat in the vehicle. Therefore, the information on the existence of the following variables is needed:

- seating arrangement in the vehicle (SEATS)
- information to the type of seats (SEAT_DATA)
- personal information (e.g., age, gender, weight) on the occupants (PERSONAL_DATA)
- participant information (PARTICIPANT_DATA)

In addition, the information on the existence of the general assignment of the participant to the respective vehicle (PARTICIPANT_DATA), the participant type (PARTTYPE) and general vehicle

information (VEHICLE_DETAILS) are also necessary information. Figure 10 summarises the necessary variables that have to be existence in a data source to answer the considered question.



QUESTION_ID	207	TYPE_OF_QUESTION	MAIN RESEARCH QUESTION	NECESSARY:	9
Question:	Which effect does the sitting position have on the injury severity?				

TITLE:	0	LASERSCAN:	0	DISTRACTION:	0	AIS2015:	0	NROCCUPANTS:	0	RESTCOEFF:	0
ACRONYM:	0	AIR_PHOTO:	0	INFLFACTOR:	0	INJIMAGE:	0	LIGHTSYS:	0	PRECRASH_DATA:	0
PROVIDER:	0	FRICTION_MEASUREMENT:	0	MANOEUVRE:	0	THERAPY:	0	PARKSYS:	0	PRECRASH_REMARK:	0
PROVIDER_TYPE:	0	RADIOGRAPHS:	0	IMPACT_POINT:	0	INJCAUSE:	0	SEATS:	1	FORMAT_PCM:	0
GENERAL_TYPE:	0	MEDICAL_REPORT:	0	SEATPOSITION:	1	INJCOLL:	0	SEAT_DATA:	1	FORMAT_PC_CRASH:	0
FORMAT:	0	ACCTIME:	0	AGE:	0	INJLOCATION:	0	CHILDSEAT:	0	FORMAT_CARMAKER:	0
ACCESS:	0	ACCDATE:	0	GENDER:	0	ROADTYPE:	0	TIRE_DATA:	0	FORMAT_OPEN_X:	0
UPDATE_RATE:	0	ACCMONTH:	0	HEIGHT:	0	ROADSURF:	0	TRAILER_DATA:	0	VIDEO:	0
DATA_FORMAT:	0	ACCYEAR:	0	WEIGHT:	0	ROADCOND:	0	INT_DAMAGE:	0	ACTIVE_SYSTEMS:	0
YEAR_START:	0	WDAY:	0	BMI_STATURE:	0	LANES:	0	EXT_DAMAGE:	0	PASSIVE_SYSTEMS:	0
YEAR_END:	0	DAYTIME:	0	NATIONALITY:	0	LANESEPAR:	0	DEFODEPTH:	0	BELTSYSTEM:	0
CASE_PER_YEAR:	0	LOCATION:	0	PREILLNESS:	0	MARKINGS:	0	FIRE:	0	BELT_ACTIV:	0
REPRESENTATIVENESS:	0	ACCSITE:	0	HOSPITAL:	0	SPECIALLANE:	0	CAR_DETAILS:	0	HEADRESTPROT:	0
INVESTIGATION_AREAS:	0	GPS:	0	MEDICATION:	0	CYCLIST_PATH:	0	TRUCK_DETAILS:	0	AIRBAG_EXIST:	0
INVESTIGATION_METHOD:	0	TRAFFICREG:	0	INJ_SEVERITY:	1	OBSTRUCT:	0	PTW_DETAILS:	0	AIRBAG_DEPLOY:	0
LANGUAGE_01:	0	ACCSEVERITY:	0	INJ_SEVERITY_MAIS:	1	CONSTRUCT:	0	BICYCLE_DETAILS:	0	VRU_AIRBAG:	0
LANGUAGE_02:	0	ACCDESCR:	0	AISREGIONS:	0	ROADSIDE:	0	BUS_DETAILS:	0	POPUP_HOOD:	0
COSTS:	0	ACCTYPE:	0	ISS:	0	MEASURES:	0	Vehicle2X:	0	ABS:	0
CONTACT:	0	ACCTYPEA:	0	NISS:	0	SIGNS:	0	PRECRASHPHASE:	0	ESC:	0
LINK:	0	ACCTYPEB:	0	GCS:	0	MAXSPEED:	0	INCRASHPHASE:	0	CRUISE_CONTROL:	0
FEATURES:	0	ACCKIND:	0	MEDCONDITION:	0	USEDLANE:	0	POSTCRASHPHASE:	0	AEB_SYSTEM:	0
CAR_ACCIDENTS:	0	NR_PARTICIPANTS:	0	CAUSEDEATH:	0	OVERTAKING:	0	NRCOLLISIONS:	0	AEB_ACTIVE:	0
TRUCK_ACCIDENTS:	0	NR_VEHICLES:	0	TIMEDEATH:	0	CURVE:	0	ROLLOVER:	0	LDW:	0
BUS_ACCIDENTS:	0	NR_PERSONS:	0	PLACEDEATH:	0	CURVEVALUE:	0	SKIDDING:	0	BSM:	0
PTW_ACCIDENTS:	0	NR_INJURED:	0	LONGTERM:	0	MUEMAX:	0	OPPONENT:	0	ECALL:	0
CYCLIST_ACCIDENTS:	0	PRECIPITATION:	0	DURATION:	0	VEHTYPE:	0	COLLOBJECT:	0	ALCO_LOCK:	0
PEDESTRIAN_ACCIDENTS:	0	CLOUD_FOG_WIND:	0	FIRSTAID:	0	MAKE:	0	INITIALSPEED:	0	ATTENTION_ASSIST:	0
ACCIDENT_DATA:	0	CAUSER:	0	TRANSPORT:	0	MODEL:	0	COLLSPEED:	0	AUTOMATION_LEVEL:	0
PARTICIPANT_DATA:	1	CAUSATION:	0	TREATMENT:	0	MODELYEAR:	0	RELSPEED:	0	MULTCOLLBRAKE:	0
PERSONAL_DATA:	1	TRAFFICDENS:	0	LABVALUES:	0	VIN:	0	DECELERATION:	0	POPULATION:	0
INJURY_DATA:	0	POLICEREPORT:	0	BELTUSE:	0	FIRSTREG:	0	COLLPOINT:	0	ACCIDENTS_ALL:	0
ROAD_INFRASTRUCTURE_DATA:	0	DAMAGECOST:	0	HELMETUSE:	0	BODYTYPE:	0	COLLSIDE:	0	ACCIDENTS_INJURY:	0
VEHICLE_DETAILS:	1	RESCUETIME:	0	PROTCLO:	0	REGISTDATA:	0	CDC:	0	ACCIDENT_FATAL:	0
SAFETY_SYSTEMS:	0	PARTTYPE:	0	VRUIMPACT:	1	ENGINETYPE:	0	EES:	0	FATALLY_INJURED_PERSONS:	0
COLLISION_DATA:	0	PARTCAUSE:	0	SINGLE_INJURIES:	0	POWER:	0	DELTAV:	0	FATALITY_RATE_MIO:	0
PRE_CRASH_DATA:	0	ALCLEVEL:	0	INJTYPE:	0	DIMENSIONS:	0	OVERLAP:	0	YEAR_COUNTRY:	0
INVESTIGATION_FEATURES:	0	INTERVIEW:	0	INJNAME:	0	MILEAGE:	0	COLLANGLE:	0		
PHOTOGRAPHIC_DOCUMENTATION:	0	LICENCE:	0	AIS1990:	0	EMPTYWEIGHT:	0	IMPANGLE:	0		
SKETCH:	0	EXPERIENCE:	0	AIS1998:	0	CRASHWEIGHT:	0	SLIPANGLE:	0		
EDR_DATA:	0										

Figure 10: Necessary variables in a data source to answer the question 207

In the end of the processing of the questionnaire by the syntactic and semantic analyses, a matrix of the questionnaire can be derived, which contains the 193 questions as rows and the 237 variables as columns (Table 3). Due to the fact that a few questions have been deleted and the category “main research questions” has been added from QUESTION_ID 200 onwards, the number of questions (n) differ from the “QUESTION_ID”.

Table 3: Result matrix of the questionnaire

n	QUESTION_ID	VARIABLE_1	VARIABLE_2	VARIABLE_3	VARIABLE_4	…	VARIABLE_237
1	1	1	1	0	1	…	0
2	2	1	0	0	1	…	1
3	3	1	1	1	1	…	1
...	…	…	…	…	…	…	…
193	229	0	0	1	0	…	0

n…number of questions

4.2. Matching process

The following part describes the theoretical background of the matching process. The basis of the matching process is the same variable set for the meta database and the questionnaire. Both tables differ from each other in the following points:

- Primary key: “SOURCE_ID“ vs. “QUESTION_ID”
- Content of the tables: various label dataset for the meta database vs. binary code for the questionnaire

In the development of the meta database, some variables are extended by several labels to specify them with information (e.g., velocity speed, representativity). In the first step of the matching process, the meta database is exported from Microsoft® Access® and is translated into a binary code to be able to match the content with the binary code of the questionnaire (Figure 11).

SOURCE_ID	VARIABLE_1	VARIABLE_2	VARIABLE_3	VARIABLE_4	…	VARIABLE_237
1	3	0	1	1	…	0
2	4	1	2	1	…	1
3	0	1	0	3	…	1
…	…	…	…	…	…	…

Transfer to binary codes

SOURCE_ID	VARIABLE_1	VARIABLE_2	VARIABLE_3	VARIABLE_4	…	VARIABLE_237
1	1	0	1	1	…	0
2	1	1	1	1	…	1
3	0	1	0	1	…	1
…	…	…	…	…	…	…

Figure 11: Transfer of the meta database into binary codes

After the translation of the meta database into binary codes, the questionnaire is added to the matching process and the content of both tables (meta database, questionnaire) are matched. Each “QUESTION_ID” with its coded variables is matched to each “SOURCE_ID“ and its variables (Figure 12).

QUESTION_ID	VARIABLE_1	VARIABLE_2	VARIABLE_3	VARIABLE_4	…	VARIABLE_237
1	1	1	0	1	…	0
2	1	0	0	1	…	1
3	1	1	1	1	…	1
…	…	…	…	…	…	…

SOURCE_ID	VARIABLE_1	VARIABLE_2	VARIABLE_3	VARIABLE_4	…	VARIABLE_237
1	1	0	1	1	…	0
2	1	1	1	1	…	1
3	0	1	0	1	…	1
MATCH	…	…	…	…	…	…

Figure 12: Scheme of the matching process

For the result of the matching process, a new table is created with the “QUESTION_ID” as row and the “SOURCE_ID“ as column. This table is called the result matrix (Figure 13).

QUESTION_ID	SOURCE_ID_1	SOURCE_ID_2	SOURCE_ID_3	…
1	2/3 = 66%	3/3 = 100%	2/3 = 66%	…
2	2/3 = 66%	3/3 = 100%	2/3 = 66%	…
3	3/5 = 60%	5/5 = 100%	3/5 = 60%	…
…	…	…	…	…

Figure 13: Result of the matching process

The content of the result matrix indicates the percentage of the “SOURCE_ID“ variables that match the necessary variable set to answer the respective question. For this purpose, the quotient of necessary variables from the question and available variables per data source is calculated. A complete match between one data source and one question results in a value of 100% (green box).

The matching process is implemented by a MATLAB® script and the result matrix is currently a static result, which is reimplanted as table in the meta database. Only a re-run of the MATLAB® script with the latest versions of the meta database and the questionnaire create a new result matrix.

In the last step, the generated result matrix is implemented into the meta data database via Microsoft® Access®.

4.3. Implementation of the result matrix

For an overview of the matching results the following forms have been created for each type of question in the meta database:

- 31_RESEARCH QUESTIONS - main research questions + research questions
- 32_FACT_SHEET_QUESTIONS - questions to fact sheet information
- 33_OTHER_QUESTIONS - other/analysis questions

For a better overview, the abbreviations of the respective data sources were used for the forms instead of the “SOURCE_ID“. In addition to the “QUESTION_ID” and question, each output mask contains a link to the coded questionnaire to review the coded variables. Result fields without a percentage value indicate that the data sources are not fully researched or entered at the time of the matching process.

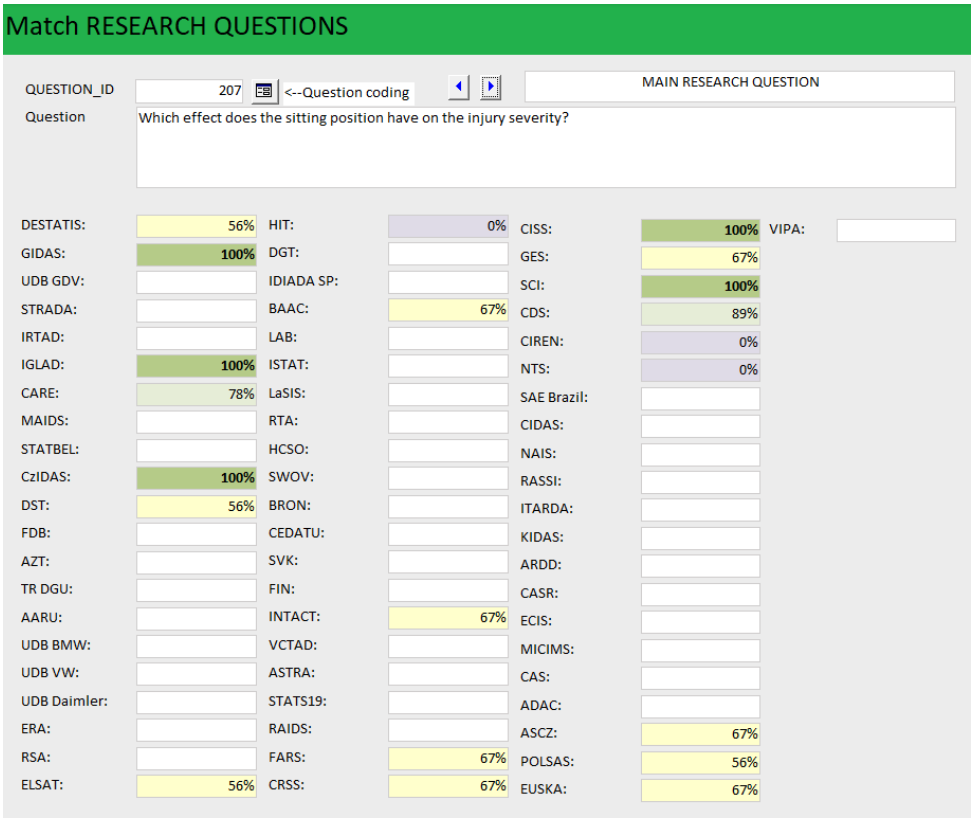


Figure 14: Output mask of the matching process in Microsoft® Access®

The questionnaire mainly contains research questions from the automotive industry and automotive supplier. If the questionnaire were extended by questions from governments, universities or other industries on road traffic accidents, the distribution of the result matrix could be different.

For a better understanding of the variables assigned to the respective questions, a mask for the questionnaire is also set up. The coded variables of each question of the questionnaire can either be reviewed via the output mask “30_QUESTIONNAIRE” or via the table “QUESTIONNAIRE”. Each implemented output mask of the result matrix is linked to the coded questionnaire via the button “Question coding” (Figure 15).

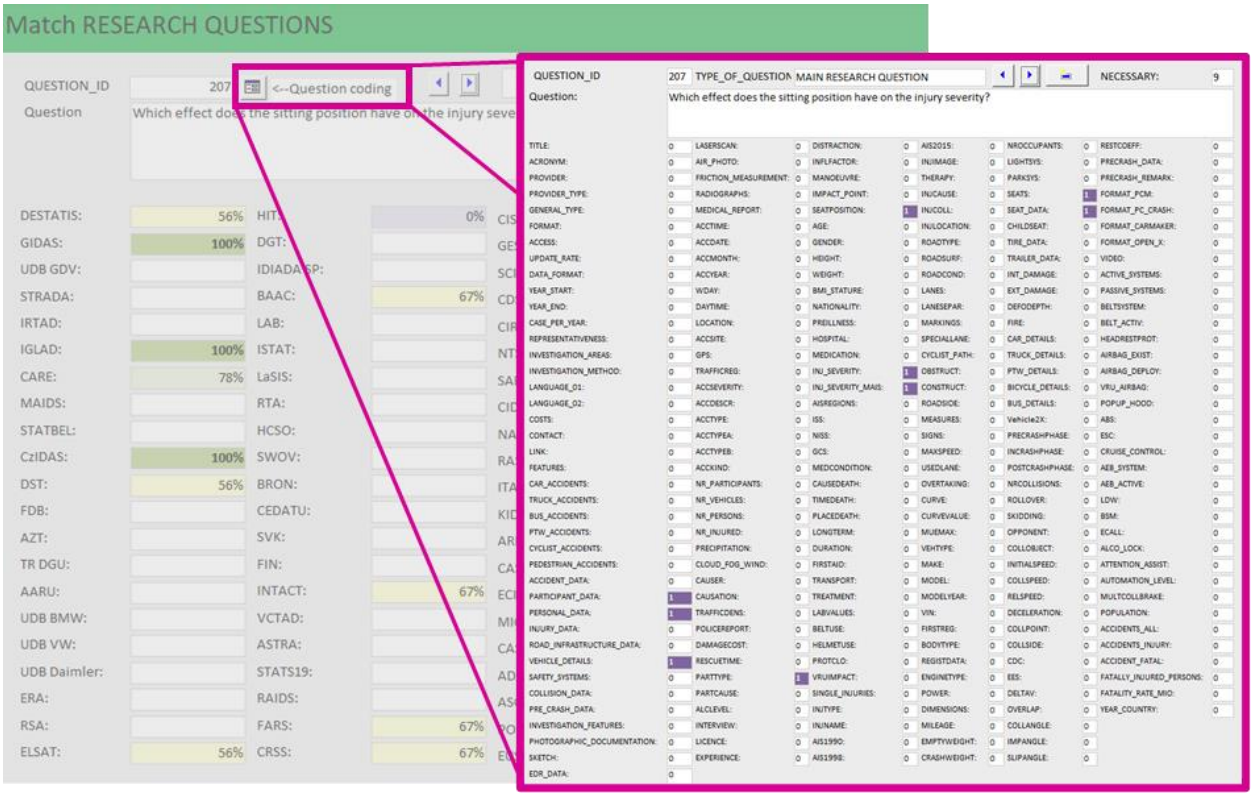


Figure 15: Output mask of the coded questionnaire (right)

From the output mask of the coded questionnaire, the user receives only the information on the existence of certain variables that are needed to answer the questions. The assessment of the data quality and the estimation of the percentage at which a data source is considered is the responsibility of the user. The extended label set of some variables should contribute to the user's decision

5. **Use of the meta database**

In addition to the presentation of the individual data sources by using the input/output mask, the implementation by Microsoft® Access® allows a simple variant to query the meta database for special requirements. Due to the fact that the meta consists of several tables each table can be queried separately by using the connection from the main table “01_DATA_SOURCES” via the “SOURCE_ID”.

For a better understanding of the usability and of the applicability, two example queries are created. The first example should give an overview of the countries, which are covered by the current meta database. The second example is more user-oriented and should identify the data sources depending on a variable/information. Regardless of the analytical approaches the main table "01_DATA_SOURCES" should be always the starting point.

In addition to the main table “01_DATA_SOURCES”, the codebook is helpful to give a better overview of the country coverage by the meta database. The codebook table enables the translation of the country codes into country names. Therefore, the variable “COUNTRY” of “01_DATA_SOURCES” is linked to the variable “validcode” of the table “Label”. The respective codebook entries can be assigned by uniquely variable identification. For this purpose, the “variablenid” in the table “Label” is set to 101 and is regarded as a condition for this query. Furthermore, the number of data sources per country is determined via the “SOURCE_ID” (Figure 16).

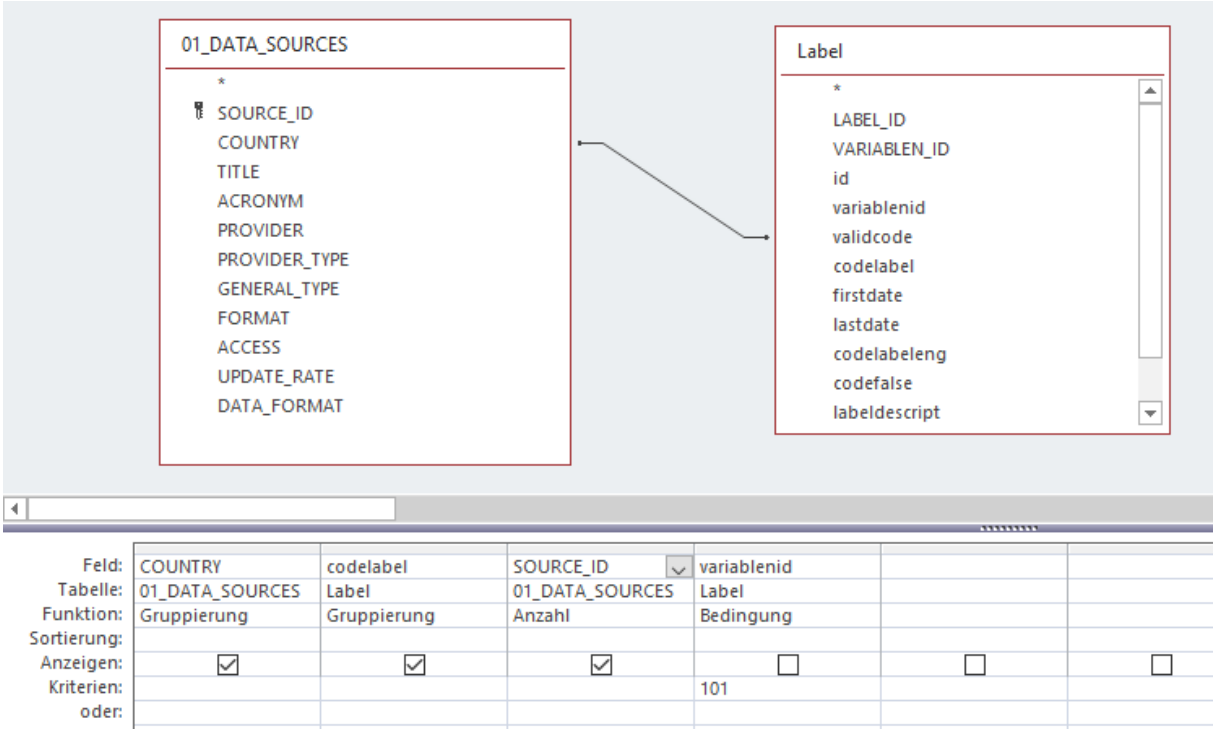


Figure 16: Microsoft® Access® query to the numbers of data sources per country

According to the query settings, the numbers of identified data sources are listed depending on the „COUNTRY_ID“ (Table 4).

Table 4: Country coverage of the meta database

„COUN-TRY_ID“	Country	Numbers of data sources	„COUN-TRY_ID“	Country	Numbers of data sources
1	World	2	87	India	1
4	Europe	2	91	Ireland	1
18	Australia	4	93	Italy	2
19	Austria	1	95	Japan	1
26	Belgium	1	103	Latvia	1
33	Brazil	1	132	Netherlands	2
46	China	2	133	New Zealand	1
54	Czech Republic	2	167	Slovakia	1
56	Denmark	2	172	South Korea	1
65	Estonia	1	174	Spain	2
69	Finland	1	178	Sweden	3
70	France	2	179	Switzerland	1
74	Germany	12	195	United Kingdom	2
76	Greece	2	196	USA	9
85	Hungary	1			

Afterwards the exported Table 4 can be used to generate a more visual overview of the country coverage by a choropleth map (Figure 17).

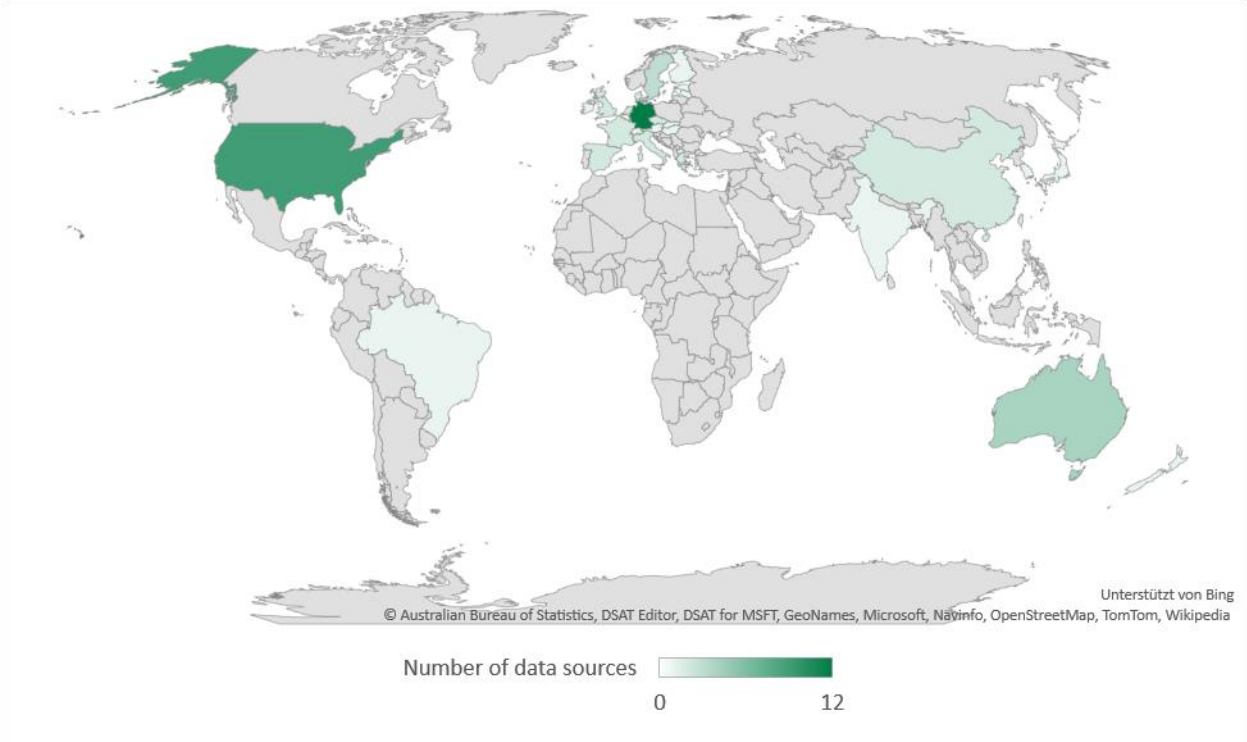


Figure 17: Choropleth map of the meta database via Microsoft®

The latest publication of the accident scenario in the European Union (EU) from 2018 reports that the share of fatally injured cyclists and pedestrians increase. [31] In the last decade from 2007 to 2016, the proportion of pedestrian fatalities increased by 1,4% and cyclist fatalities increased by 1,7% compared to the total accident scenario. Consequently, the shares of the other kind of road users indicate a decreasing trend (Figure 18).

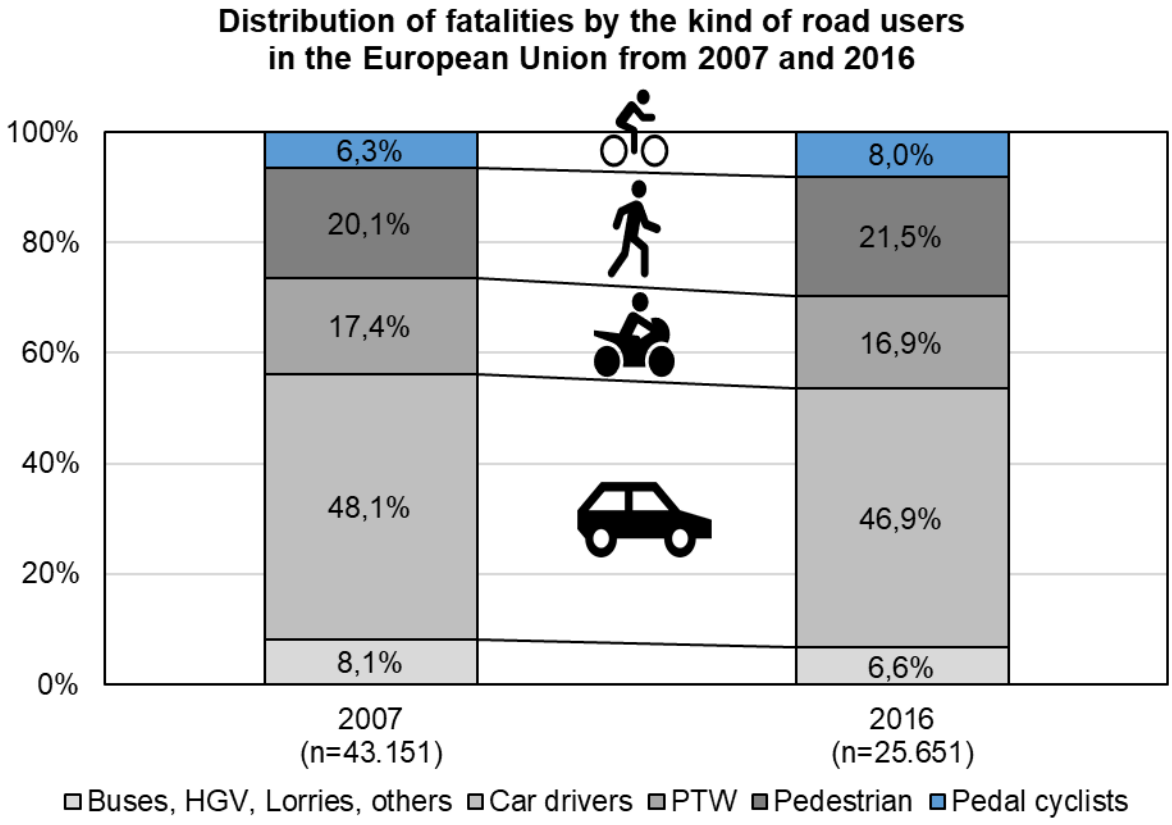


Figure 18: Distribution of fatalities in the European Union from 2007 and 2016 [31]

For the development of new safety systems and safety functions, the accident scenario with the participation of pedestrians or cyclists will become more important in the future. For this purpose, the meta database could serve as basis with information on the existence of pedestrian data or cyclist data in the EU. Cross-comparisons between different countries or regions of the EU can help to identify new safety aspects.

One of these fundamental aspects for the cyclist safety is the helmet usage and the injury severity in cyclist accidents with and without helmet. The second example shows which data sources of the current meta databank can provide information on the helmet usage and injury severity of cyclists.

In the first step, the structure of the first query example is extended. Therefore, the “SOURCE_ID” of table “01_DATA_SOURCES” is linked to the table “02_FACT_SHEET” and “13_PERSONAL_DATA”. For the selection of data sources with information on the cyclist helmet usage, the variable “CYCLIST_ACCIDENTS” in table “02_FACT_SHEET” and the variable

“HELMETUSE” in table “13_PERSONAL_DATA” have to be set to the binary code 1 or “true” (available). In the context, that the cyclist accident scenario should only be considered for European countries, the variable “COUNTRY_ID” have to be limited. For this purpose, the “COUNTRY_ID” for the world (COUNTRY = 1) and for the USA (COUNTRY = 196) are filtered out (Figure 19).

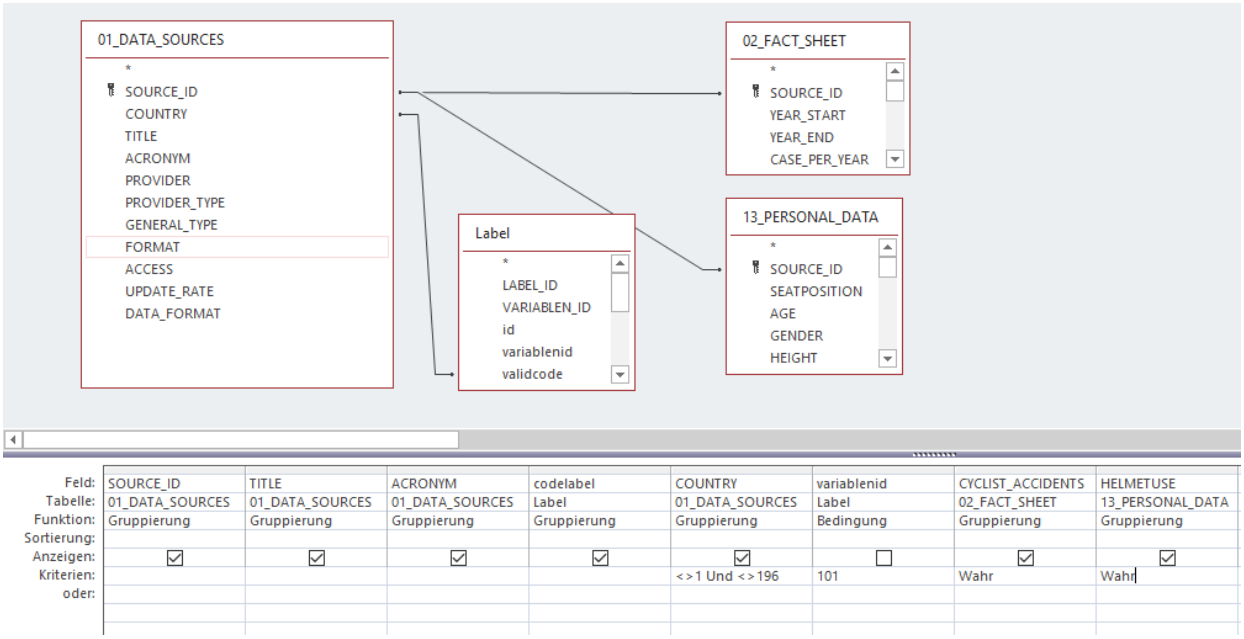


Figure 19: Example query for the cyclist helmet usage by European data sources

According to the current status of the meta database, six data sources could be used for various analyses, such as a comparison of the accident scenario between the data sources or countries. However, for such an analysis the user might need more than just the information on the existence of the helmet usage.

Table 5: Result for the example query for the cyclist helmet usage by European data sources

„SOURCE_ID“	TITLE	ACRONYM	COUNTRY (codelabel)	CYCLIST_ACCIDENTS	HELMETUSE
2	German In-Depth Accident Study	GIDAS	Germany	YES	YES
7	Community Road Accident Database	CARE	Europe	YES	YES
11	Czech In-Depth Accident Study	CZIDAS	Czech Republic	YES	YES
23	HIT accident database	HIT	Greece	YES	YES
26	Fichier BAAC	BAAC	France	YES	YES
80	Police case management system	POLSAS	Denmark	YES	YES

The next step could be to check if the access to the chosen data sources is already available or have the access to be provided by contacting the data provider. For this purpose, the meta database contains the information on the contact details (e.g., link to the data source, contact person).

The expansion of the meta database is a central component in the database development. For this purpose, the input/output mask (00_Input/Output mask) offers the possibility to implement new data source. By the activation of the button for a new data record, a new „SOURCE_ID“ and an empty data sheet are generated (Figure 20). To create a consecutive „SOURCE_ID“, the generated „SOURCE_ID“ is based on the last assigned ID. In the filling of the data set, the user needs to be careful that each variable field is filled, so that the matching process can be carried out. Empty fields for variables should be avoided.

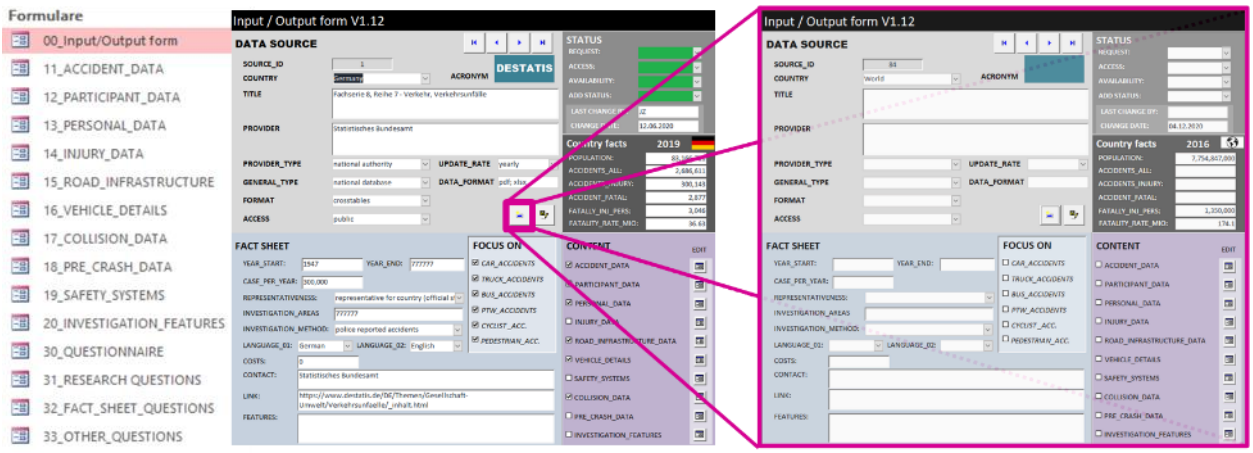


Figure 20: Implementation of new data source in the meta database

In addition to the implementation of new data sources, it is also conceivable that the questionnaire will be expanded for further authorities and their questions on accident databases in the future. Consequently, a function to the questionnaire is implemented, that allows to add new questions to the questionnaire (Figure 21).

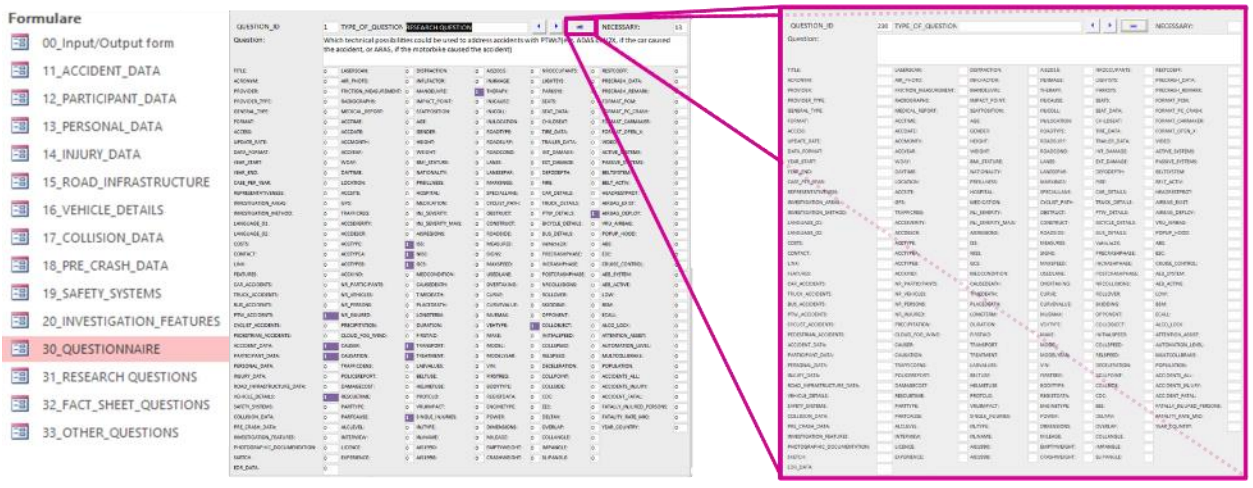


Figure 21: Implementation of new question in the questionnaire

Similar to the expansion function of the input/output mask (00_Input/Output mask), the questionnaire mask (30_QUESTIONNAIRE) contains a possibility to add a new question by coding the question according to the current variable set. The definition of the new “QUESTION_ID” based on the last assigned “QUESTION_ID”. The coding of the newly added questions is described in the section “4.1 - Questionnaire” and is based on the following binary codes:

1 = The variable is necessary to answering the question
0 = The variable is not necessary to answering the question

It is also important here that each variable field in “30_QUESTIONNAIRE” mask is filled, so that the matching process can be carried out. Empty fields should be avoided.

6. **Executive summary**

The goal of this research project was to objectively assess the quality of various accident data sources from different countries. For this purpose, a meta database was developed which based on the database structure of GIDAS. The meta database is mainly determined by the „SOURCE_ID“. Each researched data source was uniquely identified by this primary key.

In addition to the fact sheet information, the data sources were analysed to extract information on, for example, the accident scenario, the vehicle, the single injury level or the investigation features. In the development of the meta database, a distinction was made between national data sources, which are mainly based on police reported accidents, and in-depth data sources, which contain more details than the national data sources.

The research on data sources was based on a country selection that was set up in the beginning of the project. In addition to Germany, the country selection includes France, Greece, Czech Republic, Sweden, Denmark and the USA. Depending on the country, the respective data providers were contacted and asked for information. In addition to the detailed research, information on known data sources from other countries were also included in the meta database, but with a lower level of detail.

For the objective assessment of the data source quality, a questionnaire from the German automotive industry was set up with more than 170 research questions. In the first step, the questions were categorized. Afterwards a syntactic analysis of each question analysed word by word and translated them into the variables of the meta database. Thereby, variables that were necessary to answer certain questions but were not contained in the meta database are added. In addition to the syntactic analysis, the semantic analysis identified variables that were not directly mentioned in the question but were related to the question.

Parallel to the development of the meta database, a codebook for the description of the variables was set up. Most variables in the meta database and in the questionnaire were defined by a binary code, but some variables were more specified. Therefore, the top 50 variables were identified in the answering of the questionnaire and described in more detail or were extended by additional labels in the codebook.

The objective assessment of the data sources based on a matching process between meta database and the questionnaire. Therefore, a methodology was developed. The aim of the matching process was to indicate the percentage of necessary variables by each question which are cover by the various data sources. The perfect matches were represented by a 100% coverage, suggesting that this data source may be one of the most appropriate ones for answering the question. The results of the matching process were collected in a result matrix.

The result matrix provides the user an overview, which data sources are best matched to which research question. The scientific answers are up to the user. At this point, the user has to decide, which data sources are needed and which data providers have to be contacted to get access to the data sources.

7. **Outlook**

The development of the meta database for the objective assessment of various data sources provides a methodology to match the meta database content with research questions. Therefore, a compiled questionnaire from the German automotive industry has to be coded similar to the meta database. For the qualitative expansion of the current matching process, the objective estimation of some variables by non-binary labels should be considered.

In the analysis of the matching process, the comparison within the individual data sources is deliberately not considered. The subject-specific questionnaire is mainly based on research questions from development centres and research departments of the German automotive industry. Consequently, the result matrix tends more to in-depth data sources. In order to obtain a more comprehensive assessment and the comparison within the data sources, the questionnaire could be extended, for example by questions from governments, universities or by questions from the public interest on road traffic safety.

In addition to the qualitative improvements within the result matrix, the matching process has to be refined. The current matching process is realised by an external script that generated a static result matrix. Consequently, the matching process must be repeated if changes are made in the meta database or in the questionnaire. For this purpose, further development opportunities are seen in the implementation of the matching methodology into the meta database and the associated dynamic generation and implementation of the result matrix.

The current meta database, the questionnaire and the result matrix are based on an offline Microsoft® Access® database file, that needs to be updated at regular intervals. Changes made by the users in the meta database or in the questionnaire have to be collected and incorporated by a data host. In order to avoid the uncontrolled expansion of the meta database and a lot of effort in the coordination by the data host, the latest meta database structure could be implemented to a web-based platform. In the first step, the FAT has to determine a data host, which coordinate the development and the issuance of the authorised access for the web-based meta database. Furthermore, changes made by the users can be better monitored and every user has the latest update of meta database.

The continuous increase of data quantity and improvement of data quality is accompanied by a constant update of the already researched data sources as well as the expansion by further data sources. It is also conceivable to extent the meta database by exposure data for infrastructure, traffic flow, weather conditions, or driver behaviour.

The meta database and the result matrix can be a useful tool to make the development process even more efficient by minimising the research time after suitable data sources to answer the relevant development questions in the field of the vehicle safety. Furthermore, the meta database can provide as a platform to bring several data providers from different countries together and to encourage the global harmonisation of road accident data sources.

Appendix

List of literature

[1] Jörg Ortlepp, „EUska und ESN, Wirksame Instrumente für mehr Verkehrssicherheit“, Gesamtverband der Deutschen Versicherungswirtschaft e.V., 2017

[2] MMUCC Guideline, Model Minimum Uniform Crash Criteria, National Highway Traffic Safety Administration, Fifth Edition, 2017

[3] Crash Report Sampling System CRSS, Analytical User’s Manual, National Highway Traffic Safety Administration, 2018

[4] Fatality Analysis Reporting System FARS, Analytical User’s Manual, National Highway Traffic Safety Administration, 2019

[5] European Parlament, Einrichtung einer gemeinschaftlichen Datenbank, Nr. L 329/63, 1993

[6] Hellenic statistical authority 2010, Road accident data questionnaire, 2010

[7] Ministère de l’Intérieur, Guide de rédaction du Bulletin d'Analyse des Accidents Corporels de la circulation, ONISR, 2017

[8] Indberetning af faerdselsuheld – rapportnr 580, Vejdirektoratet, 2017

[9] Communication on the accident situation in Czech Republic, Czech Police, Prague Police direction, 2015

[10] "Towards a strategy on serious road traffic injuries", European Commission, 2013

[11] European Commission, “Serious Injuries”, European Commission, Directorate General for Transport, September 2015

[12] L.T. Aarts, J.J.F. Commandeur, R. Welsh et al., “Study on Serious Road Traffic injuries in the EU", October 2016

[13] National Highway Traffic Safety Administration (NHTSA). Model Minimum Uniform Crash Crite-ria (MMUCC), 5th Edition. Publication DOT HS 812 433. NHTSA, U.S. Department of Transportation, 2017

[14] H. Liers, “Traffic accident research in Germany and the German In-Depth Accident Study”, SI-AM conference, 2018

[15] Audi Accident Research Unit (AARU), “AARU Verkehrsunfallforschung am Universitätsklinikum Regensburg“, Informationsbroschüre, November 2017

[16] H. Bürkle, J. Dobberstein, Dr. A. Pneumaka, M. Muthanandam, “Accident Research - Data-driven Development and Technologoical Trends”, SIAM Conference, 2018

[17] Kolke, R. u. Rupp, A. (Hrsg.): Unfallforschung 2015, 1. ADAC Symposium für Unfallforschung und Sicherheit im Straßenverkehr. Hochschule für angewandte Wissenschaften Kempten, Schriftenreihe, Bd. 1. Göttingen: Cuvillier Verl. 2015

[18] E. Liljegren, H. Fagerlind, L. Hagström et al., “INTACT Final Report: Methodology Development for Advanced Accident Investigations - INTACT Version 1.0”, November 2010

[19] Volvo Cars, “Half a century in the service of safety: Volvo Cars’ Accident Research Team celebrates 50 years”, https://www.media.volvocars.com/global/en-gb/media/pressreleases/274029/half-a-century-in-the-service-of-safety-volvo-cars-accident-research-team-celebrates-50-years, 2020

[20] S. Kockum, R. Örtlund, A. Ekfijorden et al., “Volvo Trucks Safety Report 2017”, Volvo Trucks Accident Research Team, 2017

[21] M. Štěpán, ”CZIDAS, Czech In-depth Accident Study”, Applus+ IDIADA, September 2012

[22] Hellenic Institute of Transport (HIT), “Analysis of road accidents”, https://www.imet.gr/index.php/en/services-en-2/50-service-road-accidents-en

[23] Initiative for the global harmonisation of accident data (IGLAD), codebook, phase III, September 2019

[24] Radja, G. A., Noh, E.-Y., & Zhang, F. (2020, June). Crash Investigation Sampling System 2018 analytical user’s manual (Report No. DOT HS 812 958). National Highway Traffic Safety Administration.

[25] National Highway Traffic Safety Administration (NHTSA), Special Crash Investigations (SCI), https://www.nhtsa.gov/research-data/special-crash-investigations-sci

[26] S. Monfort, B. Mueller, “Update on pedestrian and bicyclist crashes in the US using the Vulnera-ble Road User Injury Prevention Alliance (VIPA) database”, 15th PraxisConference Pedestrian Protection, October 2020

[27] H. Liers, “Traffic accident research in Germany and the German In-Depth Accident Study”, SI-AM conference, 2018

[28] German In-Depth Accident Study (GIDAS), codebook, 06/2020

[29] German In-Depth Accident Study (GIDAS), database access by VUFO, version 06/2020

[30] Association for the Advancement of Automotive Medicine (AAAM), „The Abbreviated Injury Scale 2015“, 2018

[31] European Commission, “Annual Accident Report. European Commission”, Directorate General for Transport, June 2018

[32] Global status report on road safety 2018. Geneva: World Health Organization; 2018. Licence: CC BYNC-SA 3.0 IGO.

List of figures

List of tables

Bisher in der FAT-Schriftenreihe erschienen (ab 2015)

Nr.	Titel
270	Physiologische Effekte bei PWM-gesteuerter LED-Beleuchtung im Automobil, 2015
271	Auskunft über verfügbare Parkplätze in Städten, 2015
272	Zusammenhang zwischen lokalem und globalem Behaglichkeitsempfinden: Untersuchung des Kombinationseffektes von Sitzheizung und Strahlungswärmeübertragung zur energieeffizienten Fahrzeugklimatisierung, 2015
273	UmCra - Werkstoffmodelle und Kennwertermittlung für die industrielle Anwendung der Umform- und Crash-Simulation unter Berücksichtigung der mechanischen und thermischen Vorgeschichte bei hochfesten Stählen, 2015
274	Exemplary development & validation of a practical specification language for semantic interfaces of automotive software components, 2015
275	Hochrechnung von GIDAS auf das Unfallgeschehen in Deutschland, 2015
276	Literaturanalyse und Methodenauswahl zur Gestaltung von Systemen zum hochautomatisierten Fahren, 2015
277	Modellierung der Einflüsse von Porenmorphologie auf das Versagensverhalten von Al-Druckgussteilen mit stochastischem Aspekt für durchgängige Simulation von Gießen bis Crash, 2015
278	Wahrnehmung und Bewertung von Fahrzeugaußengeräuschen durch Fußgänger in verschiedenen Verkehrssituationen und unterschiedlichen Betriebszuständen, 2015
279	Sensitivitätsanalyse rollwiderstandsrelevanter Einflussgrößen bei Nutzfahrzeugen – Teil 3, 2015
280	PCM from iGLAD database, 2015
281	Schwere Nutzfahrzeugkonfigurationen unter Einfluss realitätsnaher Anströmbedingungen, 2015
282	Studie zur Wirkung niederfrequenter magnetischer Felder in der Umwelt auf medizinische Implantate, 2015
283	Verformungs- und Versagensverhalten von Stählen für den Automobilbau unter crashartiger mehrachsiger Belastung, 2016
284	Entwicklung einer Methode zur Crashsimulation von langfaserverstärkten Thermoplast (LFT) Bauteilen auf Basis der Faserorientierung aus der Formfüllsimulation, 2016
285	Untersuchung des Rollwiderstands von Nutzfahrzeugreifen auf realer Fahrbahn, 2016
286	χMCF - A Standard for Describing Connections and Joints in the Automotive Industry, 2016
287	Future Programming Paradigms in the Automotive Industry, 2016
288	Laserstrahlschweißen von anwendungsnahen Stahl-Aluminium-Mischverbindungen für den automobilen Leichtbau, 2016
289	Untersuchung der Bewältigungsleistung des Fahrers von kurzfristig auftretenden Wiederübernahmesituationen nach teilautomatischem, freihändigem Fahren, 2016
290	Auslegung von geklebten Stahlblechstrukturen im Automobilbau für schwingende Last bei wechselnden Temperaturen unter Berücksichtigung des Versagensverhaltens, 2016
291	Analyse, Messung und Optimierung des Ventilationswiderstands von Pkw-Rädern, 2016
292	Innenhochdruckumformen laserstrahlgelöteter Tailored Hybrid Tubes aus Stahl-Aluminium-Mischverbindun für den automobilen Leichtbau, 2017

293 Filterung an Stelle von Schirmung für Hochvolt-Komponenten in Elektrofahrzeugen, 2017

294 Schwingfestigkeitsbewertung von Nahtenden MSG-geschweißter Feinbleche aus Stahl unter kombinierter Beanspruchung, 2017

295 Wechselwirkungen zwischen zyklisch-mechanischen Beanspruchungen und Korrosion: Bewertung der Schädigungsäquivalenz von Kollektiv- und Signalformen unter mechanisch-korrosiven Beanspruchungs-bedingungen, 2017

296 Auswirkungen des teil- und hochautomatisierten Fahrens auf die Kapazität der Fernstraßeninfrastruktur, 2017

297 Analyse zum Stand und Aufzeigen von Handlungsfeldern beim vernetzten und automatisierten Fahren von Nutzfahrzeugen, 2017

298 Bestimmung des Luftwiderstandsbeiwertes von realen Nutzfahrzeugen im Fahrversuch und Vergleich verschiedener Verfahren zur numerischen Simulation, 2017

299 Unfallvermeidung durch Reibwertprognosen, 2017

300 Thermisches Rollwiderstandsmodell für Nutzfahrzeugreifen zur Prognose fahrprofilspezifischer Energieverbräuche, 2017

301 The Contribution of Brake Wear Emissions to Particulate Matter in Ambient Air, 2017

302 Design Paradigms for Multi-Layer Time Coherency in ADAS and Automated Driving (MULTIC), 2017

303 Experimentelle Untersuchung des Einflusses der Oberflächenbeschaffenheit von Scheiben auf die Kondensatbildung, 2017

304 Der Rollwiderstand von Nutzfahrzeugreifen unter realen Umgebungsbedingungen, 2018

305 Simulationsgestützte Methodik zum Entwurf intelligenter Energiesteuerung in zukünftigen Kfz-Bordnetzen, 2018

306 Einfluss der Kantenbearbeitung auf die Festigkeitseigenschaften von Stahl-Feinblechen unter quasistatisch und schwingender Beanspruchung, 2018

307 Fahrerspezifische Aspekte beim hochautomatisierten Fahren, 2018

308 Der Rollwiderstand von Nutzfahrzeugreifen unter zeitvarianten Betriebsbedingungen, 2018

309 Bewertung der Ermüdungsfestigkeit von Schraubverbindungen mit gefurchtem Gewinde, 2018

310 Konzept zur Auslegungsmethodik zur Verhinderung des selbsttätigen Losdrehens bei Bauteilsystemen im Leichtbau, 2018

311 Experimentelle und numerische Identifikation der Schraubenkopfverschiebung als Eingangsgröße für eine Bewertung des selbsttätigen Losdrehens von Schraubenverbindungen, 2018

312 Analyse der Randbedingungen und Voraussetzungen für einen automatisierten Betrieb von Nutzfahrzeuge im innerbetrieblichen Verkehr, 2018

313 Charakterisierung und Modellierung des anisotropen Versagensverhaltens von Aluminiumwerkstoffen für die Crashsimulation, 2018

314 Definition einer „Äquivalenten Kontakttemperatur“ als Bezugsgröße zur Bewertung der ergonomischen Qualität von kontaktbasierten Klimatisierungssystemen in Fahrzeugen, 2018

315 Anforderungen und Chancen für Wirtschaftsverkehre in der Stadt mit automatisiert fahrenden E-Fahrzeugen (Fokus Deutschland), 2018

316 MULTIC-Tooling, 2019

317 EPHoS: Evaluation of Programming - Models for Heterogeneous Systems, 2019

318 Air Quality Modelling on the Contribution of Brake Wear Emissions to Particulate Matter Concentrations Using a High-Resolution Brake Use Inventory, 2019

319 Dehnratenabhängiges Verformungs- und Versagensverhalten von dünnen Blechen unter Scherbelastung, 2019

320 Bionischer LAM-Stahlleichtbau für den Automobilbau – BioLAS, 2019

321 Wirkung von Systemen der aktiven, passiven und integralen Sicherheit bei Straßenverkehrsunfällen mit schweren Güterkraftfahrzeugen, 2019

322 Unfallvermeidung durch Reibwertprognosen - Umsetzung und Anwendung, 2019

323 Transitionen bei Level-3-Automation: Einfluss der Verkehrsumgebung auf die Bewältigungsleistung des Fahrers während Realfahrten, 2019

324 Methodische Aspekte und aktuelle inhaltliche Schwerpunkte bei der Konzeption experimenteller Studien zum hochautomatisierten Fahren, 2020

325 Der Einfluss von Wärmeverlusten auf den Rollwiderstand von Reifen, 2020

326 Lebensdauerberechnung hybrider Verbindungen, 2020

327 Entwicklung der Verletzungsschwere bei Verkehrsunfällen in Deutschland im Kontext verschiedener AIS-Revisionen, 2020

328 Entwicklung einer Methodik zur Korrektur von EES-Werten, 2020

329 Untersuchung zu den Einsatzmöglichkeiten der Graphen- und Heuristikbasierten Topologieoptimierung zur Entwicklung von 3D-Rahmenstrukturen in Crashlastfällen, 2020

330 Analyse der Einflussfaktoren auf die Abweichung zwischen CFD und Fahrversuch bei der Bestimmung des Luftwiderstands von Nutzfahrzeugen, 2020

331 Effiziente Charakterisierung und Modellierung des anisotropen Versagensverhaltens von LFT für Crashsimulation, 2020

332 Charakterisierung und Modellierung des Versagensverhaltens von Komponenten aus duktilem Gusseisen für die Crashsimulation, 2020

333 Charakterisierung und Meta-Modellierung von ungleichartigen Punktschweißverbindungen für die Crashsimulation, 2020

334 Simulationsgestützte Analyse und Bewertung der Fehlertoleranz von Kfz-Bordnetzen, 2020

335 Absicherung des autonomen Fahrens gegen EMV-bedingte Fehlfunktion, 2020

336 Auswirkung von instationären Anströmeffekten auf die Fahrzeugaerodynamik, 2020

337 Analyse von neuen Zell-Technologien und deren Auswirkungen auf das Gesamtsystem Batteriepack, 2020

338 Modellierung der Einflüsse von Mikrodefekten auf das Versagensverhalten von Al-Druckguss-komponenten mit stochastischem Aspekt für die Crashsimulation, 2020

339 Stochastisches Bruchverhalten von Glas, 2020

340 Schnelle, breitbandige Datenübertragung zwischen Truck und Trailer als Voraussetzung für das hochautomatisierte Fahren von Lastzügen, 2021

341 Wasserstoffkompatibilität von Aluminium-Legierungen für Brennstoffzellenfahrzeuge, 2021

342 Anforderungen an eine elektrische Lade- und Wasserstoffinfrastruktur für gewerbliche Nutzfahrzeuge mit dem Zeithorizont 2030, 2021

343 Objective assessment of database quality for use in the automotive research and development process, 2021

Impressum

Herausgeber	FAT Forschungsvereinigung Automobiltechnik e.V. Behrenstraße 35 10117 Berlin Telefon +49 30 897842-0 Fax +49 30 897842-600 www.vda-fat.de
ISSN	2192-7863

Verband der Automobilindustrie e.V. (VDA)
Behrenstraße 35, 10117 Berlin
www.vda.de
Twitter @VDA_online

Forschungsvereinigung Automobiltechnik e.V. (FAT)
Behrenstraße 35, 10117 Berlin
www.vda.de/fat

VDA | Verband der Automobilindustrie

FAT | Forschungsvereinigung Automobiltechnik