Funded under European Union's Horizon 2020 research and innovation programme - Grant agreement No 724029

RECOMMENDATIONS FOR A COMMON DATA COLLECTION SYSTEM AND DEFINITIONS

Project	SAFER AFRICA
Work Package:	WP4 - Pan-African Road safety knowledge and data centre
Deliverable:	D4.2 - Recommendations for a common data collection system and definitions
Version	V2
Date	January 2018
Report authors:	P.Thomas, R.Welsh, K.Folla, A.Laiou, S.Mavromatis, G.Yannis, D.Usami, E.Meta

DOCUMENT CONTROL SHEET

Document title	Recommendations for a common data collection system and definitions
Work package:	Wp4 - Pan-African Road safety knowledge and data centre
Deliverable	D4.2
Version	V2
Last version date	
Status	Under review
File Name	SA-NTUA-WP4_D4.2
Number of pages	
Dissemination level	Public
Responsible author	P.Thomas, R.Welsh
Editors	K.Folla, A.Laiou, S.Mavromatis, G.Yannis, D.Usami, E.Meta

VERSIONING AND CONTRIBUTION HISTORY

Version	Date	Author/Editor	Contributions	Description / Comments
1st	July 2017			Under review
2nd	January 2018			Under review

Table of Contents

Executive Summary

A key assignment within the SaferAfrica project is to thoroughly assess the needs of stakeholders involved in road safety in terms of knowledge and information tools and convey a clear view of current road safety practices followed in Africa.

As a starting point, two-fold surveys as well as existing road safety analysis documents were exploited. The **findings** from these sources on road safety data, data collection systems and definitions provide a thorough **overview** of the **current situation in African countries**, where it is clear that there are significant **deficiencies** and **shortcomings**. In addition, available data are not always comparable due to the different basic road safety definitions as well as the different collection and process methods.

Therefore, it is **necessary** to **define** a **common set of data** that are necessary to understand and **assess road safety** and a **common methodology to collect** them. These tools will help acquire both accurate and comparable road safety data that can be used for evidence-based decision making.

The objective of the present deliverable is to provide recommendations and guidelines for a minimum set of harmonised data collection procedures and standard definitions that could be applied in the short- to medium term to improve African data collection systems. On that purpose, relative manuals from European and international projects were exploited by giving emphasis on the collection systems and definitions of three types of data: accident data, exposure data and road safety performance indicators. The recommendations for all types of data consist of a **minimum set of data elements and a common collection system**. However, due to limited experience, unavailability and lack of standardization in the collection process of such data for most African countries, a 2-fold priorities scenario is proposed on each data type, based on a combination of usefulness and ease to collect.

As far as **road accident data** are concerned, the police plays the major role in the data collection process, since they are the first who record the needed data, finalise them after the period of 30 days and forward them to the responsible national authority. The data collection form is recommended to be revised frequently, include detailed information on the vehicles and road users involved in the accident, as well as adopt all existing standardized international definitions of variables and values.

Concerning **road fatalities**, the international 30-days definition is recommended to be adopted by the African countries. On that purpose, the countries that are not currently utilizing such a definition should modify the data collection process and develop appropriate conversion factors. Underreporting is also an issue that should be tackled, so that the databases are further improved and comparability of the data among the countries is reached. It is recommended that road accident data are adjusted by means of linking Police data with hospital data.

Regarding the exposure and performance indicators, the respective variables and values are recommended to be defined in such a way that they will be compatible to the accident data. The **exposure measures** concern two groups of data, the road traffic estimates and the road user at risk estimates. The recommendations of the present report include a list of primary data that should be collected in order to calculate the exposure indicators, as well as additional information that could be collected at a next stage. The collection processes examined concern travel surveys and traffic count systems, while national registers may also provide with useful and commonly used exposure data, such as population, drivers' population, vehicle fleet etc.

Two mainly data collection methods exist for estimating the **road safety performance indicators**: the first one concerns observational techniques, while the second needs national statistics and data collected by national registers. Specific recommendations are given for each of the examined core areas; namely drink-driving, speed, use of protective systems and vehicles safety. In general, these recommendations concern the survey requirements (design requirements, measurement requirements, period of surveys etc.), data analysis and documentation and reporting of the final results.

Aiming to **examine the implementation process** of the recommendations for a common data collection system and definitions, certain **direct** as well as **general requirements** need to be met.

Within the SaferAfrica project, the recommendations need to be rapidly conversed to the involved local authorities of each African country. Therefore, a **network of national experts** should be defined and spread out geographically to cover Africa.

On the other hand, the **general implementation roadmap** consists of certain prerequisites, which besides capacity consolidation of the authorities and dedicated budget, involves summary sampling and costing as well as the formation of a **Pan-African coordinative organization.**

1 Introduction

1.1 Overview of survey on road safety data, data collection systems and definitions

A key assignment within the SaferAfrica project is to thoroughly assess the needs of stakeholders involved in road safety in terms of knowledge and information tools and convey a clear view of current road safety practices followed in Africa.

For this purpose, two-fold surveys as well as existing road safety analysis documents were exploited.

The surveys consisted of a brief questionnaire in order to point out the current status in each country in terms of basic road safety aspects and definitions, followed by an extensive one where, besides other concerns, detailed demands and views from road safety experts not necessarily directly involved in decision-making in each examined African country were recorded. These surveys were filled in by representatives from 20 and 21 countries respectively. As far as the extensive survey is concerned, questions on road safety management, data collection practices, road safety resources and basic road safety data were raised, developed appropriately to reflect the current conditions in Africa. The majority of the replies, 29 experts from 21 countries, were received from governmental representatives.

Furthermore, existing road safety analysis documents were exploited; namely the Global Status Report on Road Safety (WHO, 2015) and the IRF World Road Statistics 2016 (IRF, 2016) reports.

Based on the experts' responses it was found that there is a significant demand for data and knowledge in order to be used for road safety-related decision making. Currently, such information is poorly available in African countries. This fact makes the work of road safety professionals difficult, therefore, their discontent was expressed. In several cases, it was found that road safety professionals are not even aware on the availability status of items that they consider to be relevant to their work. In general, stakeholders seem to be poorly informed about the availability of road safety data and tools.

Among the most important findings among the respondents, is the fact that sustainable systems to collect and manage data on road accidents, fatalities and injuries are in place for many but not all the examined countries.

Exposure indicators were found in the examined countries' national observatories, where 5 countries out of 10 seem to include exposure data in their national road safety observatories.

Approximately 50% of the examined countries have in place a sustainable system for the collection and management of data on behavioural indicators emphasizing on speeding and alcohol impaired driving. Availability of safety belt wearing rates were found to be somehow lower.

The assessment of the existing road safety data collection systems in African countries revealed similarities but mostly differences since besides the existence of formal systems for recording road accidents for almost all countries, the data collection practices from the road safety monitoring and evaluation points of view are addressed in various ways.

Regarding the critical aspect of a common definition for road accident fatalities, serious injuries and work related accidents, it was found that although the existence of a common fatality definition (mainly) was highly prioritized such a classification is not available in all the examined countries. Another highlighted issue of general concern is the underreporting of road accidents for which the accessibility to relevant data, though regarded as a priority of major importance for the majority of the stakeholders, is only partially available. Road accident databases that link Police and hospital data may serve as a potential solution to the underreporting issue. Such a perspective for joined databases, although once again highly acknowledged by the respondents, at present, seems not available to the majority of stakeholders. Identifying high-risk sites are considered more important compared to performing in-depth accident analysis, where regarding the latter, the existence of a common methodology seems rather limited.

Exposure data although appreciated by more than 50% of the stakeholders are fully available to approximately 20% of them.

Information on road users' behavioural aspects and attitudes were found to be highly prioritized by more than 70% of road safety stakeholders in all countries. However, availability of such information is rather limited to almost 30% of stakeholders. The same percentages more or less in terms of priority and availability ratings respectively were found regarding information on road accident causation factors. From the road infrastructure point of view, data on road safety audits – inspections were greatly appreciated by the stakeholders, although such information is currently available to less than 10% of the respondents.

Concerning road safety data availability in Africa, it is widely known that there is a serious lack of road safety data on African countries, and even when data are available, for example through international databases, little is known about data collection systems, data definitions, etc. Only few countries dispose suitable time series of road fatality data and especially for the latest available decade 2005-2014, only 21 African countries have available data for more than 5 years. The greatest lack in data concerns risk exposure and safety performance indicators, for which few countries have collected such data.

Moreover, a second issue concerns the comparability of the data and the potential of using different databases in a complementary way. Concerning the fatality data, the review revealed that different definitions are used among the countries. The WHO provides the primary data as received by the national sources in the country profiles of the reports, which adjust them to the 30-days definition and publish them in the statistical tables. However, these data are not directly comparable because of differences in the quality of data collection process among the countries. In order to take into account under-reporting issues and achieve comparability, the WHO has developed statistical models to estimate the number of fatalities. In addition, the comparison of the two databases showed that while the IRF uses the 30-days definition for the killed persons in road accidents, the data that publishes are those reported by the national sources, which use different definitions. Thus, the data cannot be comparable among the countries, without being processed before, while attention is needed when combining the two databases.

Concerning the data on exposure and road safety performance, the comparability of the countries with available data is not totally reliable, since the data refer to different years, with a difference of more than 10 years in some cases (e.g. road network density). Moreover, there is not much information on the collection methods that ensures an appropriate comparison.

However, the available data are presented in tables and related figures drafted in order to obtain an approximate picture of the road safety situation in African countries. There are clear differences on road safety performance of the countries in terms of road safety outcomes, which are also obvious when examining the motorization level or the characteristics of the road infrastructure of the countries. However, the lack of data on road safety performance and traffic exposure do not permit to come to some first conclusions so far.

The examination of the existing situation regarding road safety data, data collection systems and definitions in African countries based on the survey results, provides some important insight on deficiencies of current practices which might partially explain poor road safety performance in these countries. Furthermore, in combination with the special characteristics of these countries, common deeper problems in structures and policies may be identified.

A number of the questioned issues for many African countries are collected for the first time and can be very useful to road safety decision-makers to take into consideration for future actions. In addition, identification of the specific problems may enhance participation of the African countries in road safety initiatives and undertaking a more active role which will promote their efforts towards the improvement of road safety in the area.

The above findings of the survey on road safety data, data collection systems and definitions provide a thorough overview of the current situation in African countries. All in all, it is clear that there are significant deficiencies and shortcomings in several critical issues related to road safety data collection system and definitions. In addition, available data are not always comparable due to the different basic road safety definitions as well as the different collection and process methods. Therefore, it is necessary to define a common set of data that are necessary to understand and assess road safety and a common methodology to collect them. These tools will help acquire both accurate and comparable road safety data that can be used for evidence-based decision making.

1.2 Current international developments

The United Nations has established a set of Sustainable Development Goals to be implemented by 2030. Road safety is included in Goal 3 for Health with a target aiming, by 2020, at halving the number of global traffic fatalities and injuries.

Although no country is untouched by the problem of road traffic deaths and injuries, low-income countries have fatality rates more than double those in high-income countries, and account for a disproportionate number of deaths relative to their level of motorization. The African region has the highest rate fatalities per 100,000 population and Europe the lowest.

For this reason, reliable road safety data are essential to understand, assess and monitor the nature and magnitude of the road safety problem and the related solutions, to set ambitious and achievable safety target, to design and implement effective safety policies and measure their effectiveness. Improvement made to the quality of road safety data can improve the quality of data driven policy decisions.

The International Traffic Safety Data and Analysis Group (IRTAD) of the International Transport Forum (ITF) has identified the collection and analysis of road safety data as a critical tool to design effective road safety policies. A minimum set of road safety data is required to analyse road safety. It is recommended that road safety data is collected at three levels:

- Final outcome data, including the number of persons killed and injured by type of road users, location and time
- Data on road Safety Performance Indicators (SPIs), focusing on the safety performance of vehicles, road infrastructure and post-crash care and road user behaviours.

Regarding the latter, the following are a minimum set of SPIs:

- Speed
- seatbelt wearing and use of child restraint systems
- helmet wearing by users of powered two-wheelers
- drinking and driving.

The application of a uniform methodology for producing national and harmonised SPIs has been also promoted by the European Commission for Member States countries.
Seven problem areas in road safety were selected for the development of SPIs in EU countries: alcohol and drug-use; speeds; protective systems; daytime running lights; vehicles (passive safety); roads (infrastructure) and the trauma management system; then, for each safety area, quantitative SPIs are defined.
Each EU country is encouraged to build the necessary systems of data collection for producing national SPIs, in each one of the predefined safety fields, and to make them comparable on a European level.

1.3 Objective

The objective of this deliverable is to provide recommendations and guidelines for a minimum set of harmonised data collection procedures and standard definitions that could be applied in the short- to medium term to improve African data collection systems. For this purpose, key relevant international sources such as the Manual of the WHO on Data Systems (2011) and the EU-funded research project SafetyNet (2008) are exploited. The recommendations are also based on the findings of the survey on road safety data, data collection systems and definitions in African countries previously presented. Emphasis is given on the collection systems and definitions of three types of data: accident data, exposure data and road safety performance indicators.

1.4 Methodology

Concerning the accident data, the report "Data systems - A road safety manual for decision makers and practitioners" by WHO was exploited, as well as the Deliverable of the EU SafetyNet project "CADaS - The Common Accident Data Set". As a first step, certain recommendations are given on the collection process of the accident data, focusing on the role of Police and on the storage of the data by the responsible national authority. Moreover, the principal selection criteria for the definition of the variables are given alongside with a proposed structure of the data. Finally, a minimum set of variables and values with their definitions are proposed based on the CADaS structure and definitions.

As far as exposure data are concerned, some general recommendations on the collection procedures are formed, while specific definitions and data collection recommendations are given for the following indicators: population, driver population, road length, vehicle fleet, vehicle kilometres and person kilometres. The aforementioned suggestions are based on the EU SafetyNet project deliverable "Risk Exposure Data – Recommendations for collection and exploitation".

Finally, the deliverable of the EU SafetyNet project "Road Safety Performance Indicators Manual" was exploited for the recommendations on the data collection procedures and definitions of road safety performance indicators. Initially, recommendations are provided with regard to the basic principles of conducting a survey targeted to collect such data, supported by definitions and survey recommendations for road safety performance indicators related to drink-driving, speeding, use of protection systems and vehicles' safety.

2 Accident Data

2.1 Introduction

Reliable and consistent road accident data are a valuable and necessary prerequisite for the support of decision making aimed at the improvement of road safety. Based on the WHO report on Data Systems (2011), some steps are given in order to strengthen an existing road accident system or design and implement a new one. The basic targets are considered similar when designing a common data collection system based on the national existing ones. These steps are the following:

- Establishing a working group, which will review and discuss the road safety goals set already by the national lead agency in terms of data requirements for monitoring and achieving each one.
- Choosing a course of action, which is a range of strategies aiming to strengthen road safety systems depending on the different needs and characteristics of each region or country. The main strategies concern:
  - the improvement of data quality and system performance of road accident systems coming from police data
  - the improvement of health facility-based data on road injuries.
  - the improvement of the vital registration system and particularly the death registration system
  - the combination of existing data sources in order to obtain more accurate estimates on the magnitude and effects of road injuries
- Defining the recommended minimum data elements and definitions, based on specific selection criteria.

The recommendation for a common accident data collection system consists of a minimum set of standardised data elements, which will allow for comparable road accident data to be available in Africa. Moreover, such an African common data set besides serving the national needs of each country' s organizations and authorities, should also be comparable with international data and thus, provide with reliable data the international data systems. On that purpose, knowledge and best practices from developed countries could be transferred, taking into account the particular local needs and conditions.

For the development of a common data collection system, a two-step approach is most commonly recommended:
a) improvement and harmonisation of existing data and methods
b) collection of new harmonised data

The common dataset composed of minimum data elements (variables) will be a key tool for ensuring the appropriate data are captured to enable analysis, and for maximizing consistency and compatibility of data collected across different jurisdictions/ countries. Uniformity of accident data is especially important when combining sub-national datasets and for international comparisons.

2.2 Data definitions and standards

One of the greatest limitations when examining international comparisons of road accident figures is the incompatibility of data, which is due to either different collection procedures or different definitions of the variables and values used.

Concerning **road fatalities**, the uniform international definition of persons killed in road accidents is defined as "*the persons who died within 30 days from the day of the accident".* At present this definition is used by a number of African countries and is suggested to be adopted by the remaining ones. On that purpose, some countries have to modify the data collection process and develop appropriate conversion factors, for the conversion of the number of road accident fatalities prior to the adoption of the common definition.

On the other hand, definitions of **injury severity** may present important differences among countries. Furthermore, the minimum injury for which an accident is recorded by the Police is different in each country. Especially, the distinction between seriously and slightly injured persons presents important differences among countries.

One of the main problems of each national road accident data file is that not all injury accidents are recorded. **Underreporting** is an issue of general concern in Africa and affects the degree to which the statistical output of a country's data system reveals the actual situation of road safety. Thus underreporting delivers a biased database in terms of fatalities and serious injuries. Road accident databases that link Police and hospital data may serve as a potential solution to the underreporting issue.

However, additional inaccuracies in reporting the various variables and values contained in the national road accident data collection form may exist. Such vagueness, which are inherent to the nature of these variables and values, result from the conditions under which the primary information is collected by the police officer as well as the way this information is filled-in later on. Such inaccuracies may also raise due to inadequate training of the Police force collecting the information.

Moreover, two main sources of data incompatibility can be identified and should be handled:
- incompatibilities due to missing or incomplete national definitions (e.g. for weather conditions)
- incompatibilities due to different definitions in different countries (e.g. for road types).

The establishment of international rules for road accident data variables, values, structure and definitions has been recommended by several international research projects and some efforts for harmonising accident data at international level have already taken place (e.g. CARE system). The data structure, definitions and formats for the most common variables in road safety analyses can be also used as a basis for the development of an African common data set. This structure with the respective variables and values is presented in the following sections.

However, it should be noted that when planning the introduction of new variables or modifying the existing ones, changes to the definitions and values of existing data elements should be minimized, as these can create problems with the consistency and comparability of data over time. On the other hand, if definition or data element changes are made, then the date of change should be clearly noted in official records, allowing for some misclassification during the transition period.

2.2.1 Accident data elements

The accident data elements describe the overall characteristics of the accident.

A1. Accident ID
Definition: The accident identification number is a number which will allow the accident record to be cross-referenced to road, traffic unit and person records. It consists of three distinct fields, the country code, the year and the accident number.
Obligation: Mandatory
Data type: Numeric or character string
Comments: This value is usually assigned by the police as they are responsible at the accident scene. Other systems may reference the incident using this number.

A2. Accident date
Definition: The date (day, month and year), on which the accident occurred.
Obligation: Mandatory
Data type: Numeric (DDMMYYYY)
Comments: If a part of the accident date is unknown, the respective places are filled in with 99 (for day and month). Absence of year should result in an edit check. Important for seasonal comparisons, time series analyses, management/administration, evaluation and linkage.

A3. Accident time
Definition: The time at which the accident occurred, using the 24 hour-clock format (00.00-23:59).
Obligation: Mandatory
Data type: Numeric (HH:MM)
Comments: Midnight is defined as 00:00 and represents the beginning of a new day. Variable allows for analyses of different time periods.

A4. Accident municipality and region
Definition: The municipality and county or equivalent entity in which the accident occurred.
Obligation: Mandatory
Data type: Character string
Comments: Important for analyses of local and regional programmes and critical for linkage of the accident file to other local/regional data files (hospital, roadway, etc.). Also important for inter-regional comparisons.

A5. Accident location
Definition: The exact location where the accident occurred. Optimum definition is route name and GPS/GIS coordinates if there is a linear referencing system (LRS), or other mechanism that can relate geographic coordinates to specific locations in road inventory and other files. The minimum requirement for documentation of accident location is the street name, the reference point, the distance from the reference point and direction from the reference point.
Obligation: Mandatory

Data type: Character string, to support latitude/longitude coordinates, linear referencing method, or link node system.
Comments: Critical for problem identification, prevention programmes, engineering evaluations, mapping and linkage purposes.

A6. Accident type
Definition: The accident type is characterized by the first injury or damage-producing event of the accident.
Obligation: Mandatory
Data type: Numeric
Data values:

1 Accident with a pedestrian: Accident between a vehicle and at least one pedestrian.
2 Accident with a parked vehicle: Accident between a moving vehicle and a parked vehicle. A vehicle with a driver that is just stopped is not considered as parked.
3 Accident with a fixed obstacle: Accident with a stationary object (i.e. tree, post, barrier, fence, etc).
4 Non-fixed obstacle: Accident with a non-fixed object or lost load.
5 Animal: Accident between a moving vehicle and an animal.
6 Single vehicle accident /non-collision: Accident in which only one vehicle is involved and no object was hit. Includes vehicle leaving the road, vehicle rollover, cyclists falling etc.
7 Accident with two or more vehicles: Accident Accidents where two or more moving vehicles are involved.
8 Other accident: Other accident types not described above.

Comments: If the road accident includes more than one event, the first should be recorded, through this variable. If more than one value is applicable, only the one that corresponds best to the first event should be selected. Important for understanding accident causation, identifying accident avoidance countermeasures.

A7. Impact type
Definition: Indicates the manner in which the road motor vehicles involved initially collided with each other. The variable refers to the first impact of the accident, if that impact was between two road motor vehicles.
Obligation: Mandatory
Data type: Numeric
Data values:

1 No impact between motor vehicles: There was no impact between road motor vehicles. Refers to single vehicle accident, collisions with pedestrians, animals or objects.
2 Rear end impact: The front side of the first vehicle collided with the rear side of the second vehicle.
3 Head on impact: The front sides of both vehicles collided with each other.
4 Angle impact – same direction: Angle impact where the front of the first vehicle collides with the side of the second vehicle.
5 Angle impact – opposite direction: Angle impact where the front of the first vehicle collides with the side of the second vehicle.
6 Angle impact – right angle: Angle impact where the front of the first vehicle collides with the side of the second vehicle.

7 Angle impact – direction not specified: Angle impact where the front of the first vehicle collides with the side of the second vehicle.
8 Side by side impact – same direction: The vehicles collided side by side while travelling in the same direction.
9 Side by side impact – opposite direction: The vehicles collided side by side while travelling in opposite directions.
10 Rear to side impact: The rear end of the first vehicle collided with the side of the second vehicle.
11 Rear to rear impact: The rear ends of both vehicles collided with each other.

Comments: Useful for identifying structural defects in vehicles.

A8. Weather conditions

Definition: Prevailing atmospheric conditions at the accident location, at the time of the accident.
Obligation: Mandatory
Data type: Numeric
Data values:

1 Clear (No hindrance from weather, neither condensation nor intense movement of air. Clear and cloudy sky included)
2 Rain (heavy or light)
3 Fog, mist or smoke
4 Sleet, hail
5 Severe winds (Presence of winds deemed to have an adverse effect on driving conditions)
6 Other weather condition
7 Unknown weather condition

Comments: Allows for the identification of the impact of weather conditions on road safety. Important for engineering evaluations and prevention programmes.

A9. Light conditions

Definition: The level of natural and artificial light at the accident location, at the time of the accident.
Obligation: Mandatory
Data type: Numeric
Data values:

1 Daylight: Natural lighting during daytime.
2 Twilight: Natural lighting during dusk or dawn. Residual category covering cases where daylight conditions were very poor.
3 Darkness: No natural lighting, no artificial lighting
4 Dark with street lights unlit: Street lights exist at the accident location but are unlit.
5 Dark with street lights lit: Street lights exist at the accident location and are lit.
9 Unknown: Light conditions at time of accident are unknown.

Comments: Information about the presence of lighting is an important element in analysis of spot location or in network analysis. Additionally, important for determining the effects of road illumination on night-time accident accidents to guide relevant future measures.

2.2.2 Accident data elements derived from collected data

AD1. Accident severity

Definition: Describes the severity of the road accident, based on the most severe injury of any person involved.

Obligation: Mandatory

Data type: Numeric

Data values:

1. Fatal: At least one person was killed immediately or died within 30 days as a result of the road accident.
2. Serious/severe injury: At least one person was hospitalized for at least 24 hours because of injuries sustained in the accident, while no one was killed.
3. Slight/minor injury: At least one of the participants of the accident was hospitalized less than 24 hours or not hospitalized, while no participant was seriously injured or killed.

Comments: Provides a quick reference to the accident severity, summarizing the data given by the individual personal injury records of the accident. Facilitates analysis by accident severity level.

2.2.3 Road data elements

The road related data elements describe the characteristics of the road and associated infrastructure at the place and time of the accident.

R1. Type of roadway

Definition: Describes the type of road, whether the road has two directions of travel, and whether the carriageway is physically divided. For accident occurring at junctions, where the accident cannot be clearly allocated in one road, the road where the vehicle with priority was moving is indicated.

Obligation: Mandatory

Data type: Numeric

Data values:

1. Motorway/freeway: Road with separate carriageways for traffic in two directions, physically separated by a dividing strip not intended for traffic. Road has no crossings at the same level with any other road, railway or tramway track, or footpath. Specially sign-posted as a motorway and reserved for specified categories of motor vehicles.
2. Express road: Road with traffic in two directions, carriageways not normally separated. Accessible only from interchanges or controlled junctions. Specially sign-posted as an express road and reserved for specified categories of motor vehicles. Stopping and parking on the running carriageway are prohibited.
3. Urban road, two-way: Road within the boundaries of a built-up area (an area with sign-posted entries and exits). Single, undivided street with traffic in two directions, relatively lower speeds (often up to 50 km/h), unrestricted traffic, with one or more lanes which may or may not be marked.
4. Urban road, one-way: Road within the boundaries of a built-up area, with entries and exits sign-posted as such. A single, undivided street with traffic in one direction, relatively lower speeds (often up to 50 km/h).
5. Road outside a built-up area: Road outside the boundaries of a built-up area (an area with sign-posted entries and exits).
6. Restricted road: A roadway with restricted access to public traffic. Includes cul-de-sacs, driveways, lanes, private roads.

8 Other: Roadway of a type other than those listed above.
9 Unknown: Not known where the incident occurred.

Comments: Important for comparing accident rates of roads with similar design characteristics, and for conducting comparative analyses between motorway and non-motorway roads.

R2. Road functional class

Definition: Describes the character of service or function of the road where the first harmful event took place. For accident occurring at junctions, where the accident cannot be clearly allocated in one road, the road where the vehicle with priority was moving is indicated.

Obligation: Mandatory

Data type: Numeric

Data values:

1 Principal arterial: Roads serving long distance and mainly interurban movements. Includes motorways (urban or rural) and express roads. Principal arterials may cross through urban areas, serving suburban movements. The traffic is characterized by high speeds and full or partial access control (interchanges or junctions controlled by traffic lights). Other roads leading to a principal arterial are connected to it through side collector roads.
2 Secondary arterial: Arterial roads connected to principal arterials through interchanges or traffic light controlled junctions supporting and completing the urban arterial network. Serving middle distance movements but not crossing through neighborhoods. Full or partial access control is not mandatory.
3 Collector: Unlike arterials, collectors cross urban areas (neighbourhoods) and collect or distribute the traffic to/from local roads. Collectors also distribute traffic leading to secondary or principal arterials.
4 Local: Roads used for direct access to the various land uses (private property, commercial areas etc). Low service speeds not designed to serve interstate or suburban movements.

R3. Speed limit

Definition: The legal speed limit at the location of the accident.

Obligation: Mandatory

Data type: Numeric

Data values:

nnn: The legal speed limit as provided by road signs or by the country's traffic laws for each road category, in kilometres per hour (km/h).

999 unknown: The speed limit at the accident location is unknown.

Comments: For accident occurring at junctions, where the accident cannot be clearly allocated in one road, the speed limit for the road where the vehicle with priority was moving is indicated.

R4. Road obstacles

Definition: The presence of any person or object which obstructed the movement of the vehicles on the road. Includes any animal standing or moving (either hit or not), and any object not meant to be on the road. Does not include vehicles (parked or moving vehicles, pedestrians) or obstacles on the side of the carriageway (e.g. poles, trees).

Obligation: Mandatory

Data type: Numeric

Data values:

1 Yes: Road obstacle(s) present at the accident site.
2 No: No road obstacle(s) present at the accident site.
9 Unknown: Unknown presence of any road obstacle(s) at the accident site.

R5. Road surface conditions
Definition: The condition of the road surface at the time and place of the accident.
Obligation: Mandatory
Data type: Numeric
Data values:

1 Dry: Dry and clean road surface.
2 Slippery: Slippery road surface due to existence of sand, gravel, mud, leaves, oil on the road. Does not include snow, frost, ice or wet road surface.
3 Wet, damp: Wet road surface. Does not include flooding.
4 Flood: Still or moving water on the road.
5 Other: Other road surface conditions not mentioned above.
6 Unknown: The road surface conditions were unknown.

Comments: Important for identification of high wet-surface accident locations, for engineering evaluation and prevention measures.

R6. Junction
Definition: Indicates whether the accident occurred at a junction (two or more roads intersecting) and defines the type of the junction. In at-grade junctions all roads intersect at the same level. In not-at-grade junctions roads do not intersect at the same level.
Obligation: Mandatory
Data type: Numeric
Data values:

1 At-grade, crossroad: Road intersection with four arms.
2 At-grade, roundabout: Circular road.
3 At-grade, T or staggered junction: Road intersection with three arms. Includes T intersections and intersections with an acute angle.
4 At-grade, multiple junction: A junction with more than four arms (excluding roundabouts).
5 At-grade, other: Other at-grade junction type not described above.
6 Not at grade: The junction includes roads that do not intersect at the same level.
7 Not at junction: The accident has occurred at a distance greater than 20 metres from a junction.
9 Unknown: The accident location relative to a junction is unknown.

Comments: Accident occurring within 20 metres of a junction are considered as accident accidents at a junction. Important for site-specific studies and identification of appropriate engineering countermeasures.

R7. Traffic control at junction
Definition: Type of traffic control at the junction where accident occurred. Applies only to accident accidents that occur at a junction.
Obligation: Mandatory if accident occurred at a junction (R6)
Data type: Numeric
Data values:

1 Authorized person: Police officer or traffic warden at intersection controls the traffic. Applicable even if traffic signals or other junction control systems are present.
2 Stop sign: Priority is determined by stop sign(s).
3 Give-way sign or markings: Priority is determined by give-way sign(s) or markings.
4 Other traffic signs: Priority is determined by traffic sign(s) other than 'stop', 'give way' or markings.
5 Automatic traffic signal (working): Priority is determined by a traffic signal that was working at the time of the accident.
6 Automatic traffic signal (out of order): A traffic signal is present but out of order at time of accident.
7 Uncontrolled: The junction is not controlled by an authorized person, traffic signs, markings, automatic traffic signals or other means.
8 Other: The junction is controlled by means other than an authorized person, signs, markings or automatic traffic signals.

Comments: If more than one value is applicable (e.g. traffic signs and automatic traffic signals) record all that apply.

R8. Road curve

Definition: Indicates whether the accident occurred inside a curve, and what type of curve.
Obligation: Mandatory
Data type: Numeric
Data values:
1 Tight curve: The accident occurred inside a road curve that was tight (based on the judgment of the police officer).
2 Open curve: The accident occurred inside a road curve that was open (based on the judgment of the police officer).
3 No curve: The accident did not occur inside a road curve.
9 Unknown: It is not defined whether the accident occurred inside a road curve.

Comments: Useful for identification and diagnosis of high-accident locations, and for guiding changes to road design, speed limits, etc.

R9. Road segment grade

Definition: Indicates whether the accident occurred on a road segment with a steep gradient.
Obligation: Mandatory
Data type: Numeric
Data values:
1 Yes: The accident occurred at a road segment with a high grade.
2 No: The accident did not occur at a road segment with a high grade.
9 Unknown: It is not defined whether the accident occurred at a road segment with a high grade.

Comments: Useful for identification and diagnosis of high-accident locations, and for guiding changes to road design, speed limits, etc.

2.2.4 Vehicle data elements

The vehicle data elements describe the characteristics and events of the vehicle(s) involved in the accident.

V1. Vehicle number
Definition: Unique vehicle number assigned to identify each vehicle involved in the accident.
Obligation: Mandatory
Data type: Numeric, sequential two-digit number
Comments: Allows the vehicle record to be cross-referenced to the accident record and person records.

V2. Vehicle type
Definition: The type of vehicle involved in the accident.
Obligation: Mandatory
Data type: Numeric
Data values:
1 Bicycle: Road vehicle with two or more wheels, generally propelled solely by the energy of the person on the vehicle, in particular by means of a pedal system, lever or handle.
2 Other non-motor vehicle: Other vehicle without engine not included in the list above.
3 Two/three wheel motor vehicle: Two or three-wheeled road motor vehicle (includes mopeds, motorcycles, tricycles and all-terrain vehicles).
4 Passenger car: Road motor vehicle other than a two or three-wheeled vehicle, intended for the carriage of passengers and designed to seat no more than nine (driver included).
5 Bus/coach/trolley: Passenger-carrying vehicle, most commonly used for public transport, inter-urban movements and tourist trips, seating more than nine persons. Includes vehicles connected to electric conductors and which are not rail-borne.
6 Light goods vehicle (<3.5 t): Smaller (by weight) motor vehicle designed exclusively or primarily for the transport of goods.
7 Heavy goods vehicle (≥3.5 t): Larger (by weight) motor vehicle designed exclusively or primarily for the transport of goods.
8 Other motor vehicle: Other vehicle not powered by an engine and not included in the two previous lists of values.
9 Unknown: The type of the vehicle is unknown or it was not stated.

Comments: Allows for analysis of accident risk by vehicle type and road user type (in combination with Type of road user, P20). Important for evaluation of countermeasures designed for specific vehicles or to protect specific road users.

V3. Vehicle make
Definition: Indicate the make (distinctive name) assigned by motor vehicle manufacturer.
Obligation: Mandatory if the vehicle is a motorized vehicle. Not applicable to bicycles, tricycles, rickshaws and animal-powered vehicles.
Data type: Character string. Alternatively, a list of motor vehicle makes can be composed, with a code corresponding to each. Such a list allows for more consistent and reliable recording, as well as for easier interpretation of the data.
Comments: Allows for accident analyses related to the various motor vehicle makes.

V4. Vehicle model
Definition: The code assigned by the manufacturer to denote a family of motor vehicles (within a make) that have a degree of similarity in construction.

Obligation: Mandatory if the vehicle is a motorized vehicle. Not applicable to bicycles, tricycles, rickshaws and animal-powered vehicles

Data type: Character string. Alternatively, a list of motor vehicle models can be composed, with a code corresponding to each. Such a list allows for more consistent and reliable recording, as well as for easier interpretation of the data.

Comments: Record the name of the model as referred to in the country in which the accident occurred. Allows for accident analyses related to the various motor vehicle models.

V5. Vehicle model year

Definition: The year assigned to a motor vehicle by the manufacturer.

Obligation: Mandatory if the vehicle is a motorized vehicle. Not applicable to bicycles, tricycles, rickshaws and animal-powered vehicles

Data type: Numeric (YYYY)

Comments: Can be obtained from vehicle registration. Important for use in identifying motor vehicle model year for evaluation, research, and accident comparison purposes.

V6. Engine size

Definition: The size of the vehicle's engine is recorded in cubic centimeters (cc).

Obligation: Mandatory, if vehicle is motorized. Not applicable to bicycles, tricycles, rickshaws and animal-powered vehicles.

Data type: Numeric

Data values:

nnnn: Size of engine

9999: Unknown engine size

Comments: Important for identifying the impact of motor vehicle power on accident risk.

V7. Vehicle special function

Definition: The type of special function being served by this vehicle regardless of whether the function is marked on the vehicle.

Obligation: Mandatory, if vehicle is motorized. Not applicable to bicycles, tricycles, rickshaws and animal-powered vehicles.

Data type: Numeric

Data values:

1 No special function: No special function of the vehicle.

2 Taxi: Licensed passenger car for hire with driver, without predetermined routes.

3 Vehicle used as bus: Passenger road motor vehicle used for the transport of people.

4 Police / military: Motor vehicle used for police / military purposes.

5 Emergency vehicle: Motor vehicle used for emergency purposes (includes ambulances, fire service vehicles etc).

8 Other: Other special functions, not mentioned above.

9 Unknown: It was not possible to record a special function.

Comments: Important to evaluate the accident involvement of vehicles used for special uses.

V8. Vehicle manoeuvre

Definition: The controlled manoeuvre for this motor vehicle prior to the accident.

Obligation: Mandatory

Data type: Numeric
Data values:
1 Reversing: The vehicle was reversing.
2 Parked: Vehicle was parked and stationary.
3 Entering or leaving a parking position: The vehicle was entering or leaving a parking position
4 Slowing or stopping: The vehicle was slowing or stopping
5 Moving off: The vehicle was still and started moving. Does not include vehicle leaving or entering a parking position.
6 Waiting to turn: The vehicle was stationary, waiting to turn.
7 Turning: The vehicle was turning (includes U-turns).
8 Changing lane: The vehicle was changing lane.
9 Avoidance manoeuvre: The vehicle changed its course in order to avoid an object on the carriageway (including another vehicle or pedestrian).
10 Overtaking vehicle: The vehicle was overtaking another vehicle.
11 Straight forward / normal driving: The vehicle was moving ahead away from any bend.
12 Other
13 Unknown

2.2.5 Person data elements

The person data elements describe the characteristics, actions, and consequences relating to the people involved in the accident. These elements are to be completed for every person injured in the accident, and also for the drivers of all vehicles (motorized and non-motorized) involved in the accident.

P1. Person number
Definition: Number assigned to uniquely identify each person involved in the accident.
Obligation: Mandatory
Data type: Numeric (two-digit number, nn)
Comments: The persons related to the first (presumed liable) vehicle will be recorded first. Within a specific vehicle, the driver will be recorded first, followed by the passengers. Allows the person record to be cross-referenced to accident, road and vehicle records to establish a unique linkage with the Accident ID (A1) and the Vehicle number (V1).

P2. Occupant's vehicle number
Definition: The unique number assigned for this accident to the motor vehicle in which the person was an occupant (V1).
Obligation: Mandatory
Data type: Numeric (two-digit number, nn)
Comments: Allows the person record to be cross-referenced to the vehicle records, linking the person to the motor vehicle in which they were travelling.

P3. Pedestrian's linked vehicle number
Definition: The unique number assigned for this accident to the motor vehicle which collided with this person (V1). The vehicle number assigned under (V1) to the motor vehicle which collided with this person.
Obligation: Mandatory

Data type: Numeric (two-digit number, nn, from V1)
Comments: Allows the person record to be cross-referenced to the vehicle records, linking the person to the motor vehicle that struck them.

P4. Date of birth
Definition: Indicates the date of birth of the person involved in the accident.
Obligation: Mandatory
Data type: Numeric (date format – dd/mm/yyyy, 99/99/9999 if birth date unknown)
Comments: Allows calculation of person's age. Important for analysis of accident risk by age group, and assessing effectiveness of occupant protection systems by age group. Key variable for linkage with records in other databases.

P5. Gender
Definition: Indicates the gender of the person involved in the accident.
Obligation: Mandatory
Data type: Numeric
Data values:
1 Male: On the basis of identification documents / personal ID number or determined by the police.
2 Female: On the basis of identification documents / personal ID number or determined by the police.
9 Unknown: Sex could not be determined (police unable to trace person, not specified).
Comments: Important for analysis of accident risk by sex. Important for evaluation of the effect of sex of the person involved on occupant protection systems and motor vehicle design characteristics.

P6. Type of road user
Definition: This variable indicates the role of each person at the time of the accident.
Obligation: Mandatory
Data type: Numeric
Data values:
1 Driver: Driver or operator of motorized or non-motorized vehicle. Includes cyclists, persons pulling a rickshaw or riding an animal.
2 Passenger: Person riding on or in a vehicle, who is not the driver. Includes person in the act of boarding, alighting from a vehicle or sitting/stranding.
3 Pedestrian: Person on foot, pushing or holding a bicycle, pram or a pushchair, leading or herding an animal, riding a toy cycle, on roller skates, skateboard or skis. Excludes persons in the act of boarding or alighting from a vehicle.
8 Other: Person involved in the accident who is not of any type listed above.
9 Unknown: It is not known what role the person played in the accident.
Comments: Allows for analysis of accident risk by road user type (in combination with Vehicle type, V2). Important for evaluation of countermeasures designed to protect specific road users.

P7. Seating position
Definition: The location of the person in the vehicle at the time of the accident.
Obligation: Mandatory for all vehicle occupants
Data type: Numeric

Subfield: Row
Data values:
1 Front
2 Rear
3 Not applicable (e.g. riding on motor vehicle exterior)
8 Other
9 Unknown

Subfield: Seat
Data values:
1 Left
2 Middle
3 Right
4 Not applicable (e.g. riding on motor vehicle exterior)
8 Other
9 Unknown

Comments: Important for full evaluation of occupant protection programmes.

P8. Injury severity

Definition: The injury severity level for a person involved in the accident.
Obligation: Mandatory
Data type: Numeric
Data values:
1 Fatal injury: Person was killed immediately or died within 30 days, as a result of the accident.
2 Serious/severe injury: Person was hospitalized for at least 24 hours because of injuries sustained in the accident.
3 Slight/minor injury: Person was injured and hospitalized for less than 24 hours or not hospitalized.
4 No injury: Person was not injured.
9 Unknown: Injury severity was not recorded or is unknown.

Comment: Important for injury outcome analysis and evaluation and appropriate classification of accident severity (PD1). Important element for linkage with records in other databases.

P9. Safety equipment

Definition: Describes the use of occupant restraints, or helmet use by a motorcyclist or bicyclist.
Obligation: Mandatory
Data type: Numeric
Subfield: Occupant restraints
Data values:
1 Seat-belt available, used
2 Seat-belt available, not used
3 Seat-belt not available
4 Child restraint system available, used
5 Child restraint system available, not used
6 Child restraint system not available
7 Not applicable: No occupant restraints could be used on the specific vehicle (e.g. agricultural tractors).

8 Other restraints used
9 Unknown: Not known if occupant restraints were in use at the time of the accident.
10 No restraints used

Subfield: Helmet use

Data values:

1 Helmet worn
2 Helmet not worn
3 Not applicable (e.g. person was pedestrian or car occupant)
9 Unknown

Comments: Information on the availability and use of occupant restraint systems and helmets is important for evaluating the effect of such safety equipment on injury outcomes.

P10. Pedestrian manoeuvre

Definition: The action of the pedestrian immediately prior to the accident.

Obligation: Mandatory

Data type: Numeric

Data values

1 Crossing: The pedestrian was crossing the road.
2 Walking on the carriageway: The pedestrian was walking across the carriageway facing or not facing traffic.
3 Standing on the carriageway: The pedestrian was on the carriageway and was stationary (standing, sitting, lying etc).
4 Not on the carriageway: The pedestrian was standing or moving on the sidewalk or at any point beside the carriageway.
8 Other: The vehicle or the pedestrian was performing a manoeuvre not included in the list of the previous values.
9 Unknown: The manoeuvre performed by the vehicle or the pedestrian was not recorded or it was unknown.

Comments: Provides useful information for the development of effective road design and operation, education and enforcement measures to accommodate pedestrians.

P11. Alcohol use suspected

Definition: Law enforcement officer suspects that person involved in the accident has consumed alcohol.

Obligation: Mandatory for all drivers of motorized vehicles, recommended for all non-motorists (pedestrians and cyclists).

Data type: Numeric

Data values:

1 No
2 Yes
3 Not applicable (e.g. if person is not driver of motorized vehicle)
9 Unknown

P12. Alcohol test

Definition: Describes alcohol test status, type and result.

Obligation: Conditional (mandatory if alcohol use suspected, P25)

Data type: Numeric
Subfield: Test status
Data values:
1 Test not given
2 Test refused
3 Test given
9 Unknown if tested
Subfield: Test type
Data values:
1 Blood
2 Breath
3 Urine
8 Other
9 Test type unknown
Subfield: Test result
Data values
Value
Pending
Result unknown
Comments: Alcohol-related accidents are a major road safety problem. Information on alcohol involvement in accident facilitates evaluation of programmes to reduce drink-driving.

P13. Drug use
Definition: Indication of suspicion or evidence that person involved in the accident has consumed illicit drugs.
Obligation: Mandatory for all drivers of motorized vehicles, recommended for all non-motorists (pedestrians and cyclists).
Data type: Numeric
Data values:
1 No suspicion or evidence of drug use
2 Suspicion of drug use
3 Evidence of drug use (further subfields can specify test type and values)
4 Not applicable (e.g. if person is not driver of motorized vehicle)
9 Unknown

P14. Driving licence issue date
Definition: Indicates the date (month and year) of issue of the person's first driving licence, provisional or full, pertaining to the vehicle they were driving.
Obligation: Mandatory for all drivers of motorized vehicles
Data type: Numeric (MMYYYY)
Data values:
Value (MMYYYY)
Never issued a driving licence
Date of issue of first licence unknown
Comments: Allows calculation of number of years' driving experience at the time of accident.

2.2.6 Person data elements derived from collected data

PD1. Age
Definition: The age in years of the person involved in the accident.
Data type: Numeric
Comments: Derived from Date of birth (P4) and Accident date (A2). Important for analysis of accident risk by age group, and assessing effectiveness of countermeasures by age group.

2.3 Data collection and storage process

There are three primary methods by which accident data can be collected; police reports, hospital reports and in-depth investigations.

2.3.1 Police reports

In most countries, the Police play a key role in the accident data collection process since they are the first to arrive at the accident scene and record the needed data and are the last to update the related data. The Police are also responsible for providing the authorities with the collected data. Relevant authorities such as the police, ministries or governmental departments are then responsible for maintaining the National accident data files and publishing related statistics.

When called to an accident with casualties, the Police have to carry out an on-site investigation and sometimes fill in an autopsy report as well as a part of the accident data collection form. This form will be completed later at the police headquarters. When the 30-days definition of fatalities is in place, the accident data forms have to be kept in the police headquarters for at least one month and be finalised with the necessary updates for any killed road users.

When the national road accident data are finalised, the Police are in charge of forwarding the data to the body responsible for the national accident data file, e.g. the National Statistical Office, the Ministry of Transport etc.

The main tool for accident data collection is the data collection form, hence the central national authority responsible for the national accident file has to carry out the initial development and the revisions later on, with the aim to cover not only the national needs but also the international requirements.

The accident data collection form has to be coupled with clear instructions for filling in, as well as for the data transmission process to the national data file. The national road accident data form has to be revised regularly (at least once every ten years) in order to better cope with the new needs of road accident analysis at national and international level, while attention should be given to compatibility issues before and after the modifications.

The road accident data collection form should also include detailed information on the accident type and conditions, the road infrastructure and the road and traffic environment. Moreover, it should include detailed information on each vehicle involved in the accident and on each road user (driver, passenger or pedestrian) affected by the accident.

Consequently, the national accident data collection form should be simple and self-explaining in its structure. Moreover, the related instructions should be precise and detailed, in order to provide clear and complete data definitions. It is also recommended that all existing standardized international definitions of variables and values are adopted by the national authorities when developing or revising their accident data collection forms.

Once the road accident data collection form is finalised by the Police, the form is forwarded to the national authority responsible for maintaining the national road accident data file. The necessary data quality control should then be undertaken within

Then, the data should be coded and introduced in the electronic national road accident data file. Data coding includes the attribution of identification numbers to all accidents, vehicles and persons involved, as well as the attribution of numerical codes to all data values. It is also suggested to use different coding (i.e. groups of values) for the same variable, in order to allow for different levels of detail to be directly available for the data users. For example, it is common to code person age both in years and in age group classifications.

The structure of the national data file should be in accordance with the structure of the accident data collection form. The use of sub-files, with each of them referring to the accident, person and vehicle, would be efficient due to the hierarchical relationships of the accident components. The different sub-files should be linked by means of the accident, vehicle, road and person identification numbers, so that combined information on all accident components can be easily retrieved. Thus, the national accident data file will include disaggregate data for all accidents components, which can be retrieved by means of queries.

2.3.2 Hospital data

Data can be collected concerning road accident casualties who attend / are admitted to hospital as a consequence of their accident. This provides the potential for the formation of a database relating to Hospital Episodes.

For example, information on casualties admitted to hospital as in-patients in England is contained in the Hospital Episodes Statistics (HES) database owned by the Information Centre of the National Health Service (NHS) (www.hesonline.nhs.uk/). It is compiled by the Information Centre (IC) from over 300 NHS Trusts in England. Casualties treated in Accident and Emergency departments who are not subsequently admitted to a hospital are not included in the HES database. However, all casualties admitted to a bed in a hospital in England should be recorded in the data even if the admission did not require an overnight stay. International standard diagnostic classifications are used in the health records (ICD-10). These include transport accident codes which allow for the identification of road transport accident casualties. More specifically, they allow the identification of road user type and casualty class (e.g. casualty being a passenger of a motorcycle.

For this method, the hospital admissions records are based on periods of care (episodes) under a particular consultant. So a single patient may have more than one episode of care arising from a single accident (e.g. if they transfer to another consultant). Therefore, some data cleaning (de-duplication) needs to be carried out to identify records relating to the same patient and same accident.

As with the Police data, clear guidelines for the collection and coding of variables to be included in Hospital data are required. Identifiers should be put in place that allow matching of hospital and police data in the event that both sources are collected within a country. This enables a rich database to be developed that benefits from both the on-scene report from the police and also the detailed injury outcome from the hospital.

2.3.3 In-depth accident investigations

In-depth accident data, sometimes termed microscopic data, is an ideal method to identify and evaluate human factor issues related to real world accidents and potential Human Machine Interface (HMI) issues faced by road users. The advantage of this data source is the high level of detail known about each accident and how this can be related to a number of outcomes. Microscopic data is usually collected by independent research teams with a strict methodology collecting key variables pertaining to the accident, vehicle, road user, injury data, interview information, road infrastructure and scene information, accident reconstructions and accident causation analysis all of which is collected and analysed by experienced investigators.

The data collected by the in-depth collection activities is independent and transparent, as opposed to the national reporting systems which are generally based on judicial investigations, although these will be impartial investigations they will often be collected with 'vehicle to blame' in mind. In-depth accident data collected by the researchers is aimed at the cause of the accident, not who was to blame. (Hagstroem, 2010)

Accident investigations are undertaken in two ways; at the scene or retrospectively. These are achieved by collecting data from accidents wither within minutes of their occurrence, where a specialist investigation team attend the scene along with the emergency services; or by retrospectively undertaking in-depth examinations of the vehicles and recording their damage characteristics and assessing their crashworthiness.

The information gathered at the scene or retrospectively is enhanced with follow up data including injury outcomes and causes for casualties who attend hospital and via questionnaires sent to those involved in the accident along with any available witness statements.

The data from in-depth accident investigations, whilst generally funded by a governmental body, tend to be managed, stored and analysed by research institutes contracted by the government.

2.3.4 Representivity of accident data

When setting up accident data collection protocols at a country level, it is essential that consideration be given to harmonisation of these protocols across countries so that cross-country comparative analyses can be made as robustly as possible. This has been considered at a European level within several projects including DaCoTA where a common protocol for European in-depth investigations was established (Atalar, D 2012).

Once common national methods are in place, accident data from Police and Hospital sources potentially provide the national picture in terms of the accident population and resulting injury outcomes and therefore also have the potential to be fully representative of the accident constellation.

For in-depth accident investigations, requiring specialist teams, sampling needs to be taken into consideration in order to build a data base that is fit for the required analysis purpose. To establish true representivity an ideal sampling plan would involve randomly sampling accidents 24-7 all year round from a region / regions that are nationally representative. This however is not generally feasible due to practical and financial implications.

The DaCoTA project outlined the following principles for achieving a pan-European representative accident sample for in-depth accidents (Hargstroem 2010); these can be generalised to a pan-African database or achieving national representivity from regions within a country.

- Determine a sampling area which is representative of the national picture
- Within the sampling area, random sampling is considered a necessary precondition to have broadly representative results.
- Stratification reduces the sample variance and still guarantees representativeness of the sample
- Multiple selection criteria (e.g stratification according to different variables such as road user type, accident severity) are possible provided the source of information is reliable.
- Different strategies for sampling across regions / countries can be accommodated provided they are undertaken consistently and transparently and as long as no (large) biases in the sample are introduced.

2.4 Data collection priorities

The common collection system, that in every case should be gradually implemented by the African countries, will deliver a common road accident database in a uniform format. By this mean, the common database will be continuously updated with compatible and comparable data, allowing for more reliable analyses and assessments across the African countries.

The variables and values suggested in this section may also be considered as recommendations for national road accident data collection reports.

As an initial approach, the following **selection criteria** apply for defining the minimum data elements included in the data system:
a) Data elements and values must be useful for road accident analysis at both national and international level. These elements should be routinely collected when a road accident occurs.
b) The level of detail of the variables and values corresponds to all data useful for macroscopic data analysis and not for detailed reconstruction of the scene of the accident (in-depth analysis).
c) Data elements and values should be comprehensive and concise. Each variable must include description and scope (importance to road safety) attribute values, their definitions and the data format.
d) Data which are very difficult to collect should not be included, regardless of their value for road accident analysis.

e) All variables and values refer to casualty road accident, i.e. all road accidents involving at least one moving vehicle and one person injured or killed as a consequence of this accident. Not injured participants within an injury accident can optionally be recorded. Material damage-only accidents are not considered.

The **data structure** is proposed to follow the structure proposed in the WHO manual (2011), which is separated in four categories:

Accident related variables
Road related variables
Vehicle related variables
Person related variables

For each of the variables included in the data set, the following information can be presented:
Variable Label: The label of the proposed variable, consisting of the category identifier (Accident, Road, Vehicle or Person), the numbering and the name of the variable. The importance of the variable for road safety analysis is also added: (H) for variables of high importance (1st priority) and (L) for variables of lower importance (2nd priority).
Variable definition and scope: A brief description of the variable is provided, followed by the importance and usefulness of the variable, explaining the rational lying behind its selection.
List of values: The attribute values to each variable are listed.
Value labels: Each value is identified by the code of the variable, followed by a number which corresponds to each value and its name. The (A) code is added next to the variable category code for the alternative value, when is the case.
Value definitions: The definition of each value of the variable is provided, indicating also any particularities of the value and any relevant assumptions regarding its collection.
Data Format: The way in which each variable has to be provided. Data formats concern:
- the possibility to attribute one or more values to a variable,
- the format of the value (code, number, text).

However, the data structure is further proposed to be established on a 2-fold priorities scenario based on a combination of usefulness and convenience to collect.

An overview of the structure of the four categories is presented in Table 2.1:

Accident related variables		Road related variables		Vehicle related variables		Person related variables	
1st priority	2nd priority	1st priority	2nd priority	1st priority	2nd priority	1st priority	2nd priority
Accident ID	Impact type	Type of roadway	Speed limit	Vehicle number	Engine size	Date of birth	Person ID
Accident date		Road functional class	Road obstacles	Vehicle type	Vehicle special function	Gender	Occupant's vehicle number
Accident time		Junction	Road surface conditions	Vehicle make		Type of road user	Pedestrian's linked vehicle number
Accident region - municipality			Traffic control at junction	Vehicle model		Seating position	Safety equipment
Accident location			Road curve	Vehicle model year		Injury severity	Pedestrian manoeuvre
Accident type			Road segment grade	Vehicle manoeuvre		Driving licence issue date	Alcohol use suspected
Weather conditions						Age	Alcohol test
Light conditions							Drug use
Accident severity							

Table 2.1: Overview of the proposed data structure of the common road accident data set

2.5 The example of Cameroon

A reference for the data collection process in African Countries is the project for the design and implementation of traffic accident databases and of an information system on road safety realised by CTL, SWOV and IBSR for Cameroon.

Before the implementation of the project, Cameroon showed a lack of data and tools available to decision makers to support them in identifying road safety problems, assess the potential effectiveness of the selected measures and to actually evaluate the effectiveness of those measures.

In particular, there was neither a reliable database of traffic accidents or an information system centralizing all accident data or a National Road Accident Collection Form. Each institution (National Police, Gendarmerie, hospitals) set up its own system for collecting traffic accident revealing shortcomings and errors (omissions, lack of accuracy or misinterpretation).

Thus, the aim of the project is to improve the whole accident data collection process in Cameroon enhancing the timeliness, the accuracy and the completeness of data. A quality database on road safety, included in a centralized and integrated information system for accidents data collection, management and analysis has been implemented, in order to drop paper based data collection methods. However, their adoption is not expected by all the actors involved in road safety data collection within the project duration. Especially for Police and Gendarmerie the implementation and dissemination of these tools for accident data collection is a gradual process.

Two information systems, developed by the CTL, have been adapted to the needs and conditions of Cameroon:

- SFINGE, structured on "primary" databases addressed to National Police, National Gendarmerie, Ministere de Transports (MINT) and Observatoire National de Santé Publique of the Ministère Santé Publique, (ONSP)
- SAFETY MANAGER addressed to the Analysis Centre of Ecole Nationale Supérieure des Travaux Publics (ENSTP).

SFINGE has been integrated and adapted in order to process also data on road traffic injuries collected in hospitals. It allows to:

- Collect traffic accident data directly at the accident site (for example, using a laptop computer) and immediately computerize the data;
- Manage and process data and its computerization (that is, the manual transfer of data from the data sheet to the software);
- Analyse data included in the database in order to create automatically statistics and reports, according to user-defined queries (for example filtering data by date, by user, by road or zone, etc.);
- Transfer the data from the "primary" database to the central one (at the Analysis Center of ENSTP);
- Geo-refer data on a map (Google Maps and / or other) to allow accurate identification of accident location.

The SAFETY MANAGER is an information system organized in two parts:

- the "private" part for data acquisition, management and analysis and for the safety measures planning and selection, available only to
- the "public" part, which is available to all citizens (in the form of a web portal), to carry out communication activities on road safety.

The *private* part is composed of different functions:

- Acquisition and management of data, such as creating new accident files, exporting and importing data from different sources (Police, Gendarmerie, ONSP, etc.);
- Management of databases;
- Road safety analysis: definition of subsets of accident data, accidents mapping , report preparation, descriptive analysis of traffic accidents.
- Selection of road safety interventions: creation of projects for selecting interventions, identification of critical road infrastructure elements, identification and classification of accidents, identification of accidents causation, identification and economic evaluation of measures.

The purpose of the *public* part is to provide a tool of communication of accident data and road safety results. This part is composed of different elements:

- Statistics on traffic accidents in Cameroon and in CEMAC zone, including maps and diagrams;
- Crowdsourcing tool to give citizens the opportunity to express opinions for proposed interventions or to indicate specific problems directly on a map;
- Information on various aspects of road safety, e.g. policies, projects, technical documents, communication and training tools.

The two information systems are integrated each other in order to facilitate data flows between the actors responsible for data collection and management and those responsible for specialized road safety analyses.

3 Exposure data

3.1 Introduction

Exposure indicators are considered indispensable in risk studies and international comparisons. Multiple linkages of databases as well as systematic surveys of road user behaviour could facilitate the identification of relevant exposure data. However, for the purposes of international comparisons and priority settings, efforts should be targeted in defining exposure indicators as well as their compatibility to the accident data.

The exposure measures can be classified into two groups:
- Road traffic estimates: road length, vehicle kilometres and vehicle fleet.
- Road user at risk estimates: person kilometres, population, number of trips, time in traffic and driver population.

Among these measures, vehicle fleet, driver population and road length are useful alternative exposure measures in many countries worldwide, since the related data are recorded systematically by most countries. However, the definitions used for the variables and values are often not compatible.

Some basic requirements for the collection of such exposure measures are the following:
- Travel/mobility surveys for the collection of vehicle- or persons kilometres data should be in the form required for accident risk analysis.
- Traffic counts systems have to be established on the national and main interurban road network and at a later stage urban and rural areas to be included.
- A common vehicle classification should be considered by all countries.
- A common method for calculating vehicle-kilometres from the traffic counts should be adopted.

As far as Africa is concerned, only seven countries were found to have collected exposure data. In every case the collection of these data should be performed under a common framework in order to obtain comparable indicators at international level. Therefore, the two-step methodology developed in the EU funded research project "SafetyNet" for the EU countries is recommended to be used also in the case of Africa. The methodology includes:
1) harmonisation of existing data and methods, including common transformation rules for all countries and all exposure indicators, in order to improve their national collection methods
2) collection of new harmonised data, including data collection at African level with common definitions and methods.

In order to develop preliminary guidelines regarding the harmonisation of the data, an assessment between the existing data and transformation rules from the countries collecting the respective data should be carried out followed by a pilot study. This procedure will lead to the identification of the data needs and the common definitions of variables and values for the second step. Then, for the harmonised data, data collection should be carried out by means of a common collection form (e.g. standard Tables).

The data needed for the estimation of the exposure indicators are the following:
- Road length data by road type, area type and region
- Vehicle fleet data by vehicle type and vehicle age
- Driver population data by driver age and gender
- Vehicle-kilometres by vehicle type, age, road type, area type
- Person-kilometres by person class, age and gender

Once these indicators have been harmonised and collected, additional data needs may be tackled, such as:
- Vehicle fleet by engine type
- Driver population by nationality and experience
- Vehicle-kilometres by engine size
- Person-kilometres by nationality and experience
- Number of trips by person class, age, gender and vehicle type
- Time spent in traffic by person class, age, gender and vehicle type

3.2 Definitions and description

3.2.1 Population

Population is a common exposure indicator used in road safety analyses due to the availability of the related data. Three variables are useful when assessing accident risk at a population level: person age, gender and nationality. In addition, population at regional level would be important for calculating respective risks.

All variables and values (in particular person age, gender and nationality) included in the population registers should have a straightforward meaning. Therefore, their definitions and their compatibility should be assessed and used for any risk calculation in matching with population based road safety variables and values in the accident data base.

All countries have to collect population data in national registers and update them on a regular basis by conducting nation-wide censuses. Considering that most censuses are carried out on a regular basis (e.g. every 10 years), data for the intermediate years are estimations, which are used for the annual updates of the registers.

Attention should be given to the character of population data. In general, international databases provide average population data or population as of the 1st of January of every year. To avoid misleading results, population data with the same characteristic should be used.

However, for international comparisons risk calculations based on population data are not sufficient, especially in the case of large differences of motorisation level, traffic density etc. among the countries. Therefore, additional exposure data have to be collected for risk assessment.

3.2.2 Driver population

The best source for driver population data is usually the national driver licences databases. However, differences may exist among the countries concerning the registration of licenced drivers in total or for specific vehicle types. In addition, errors or failure to update systematically the register may lead to wrong estimations of the number of drivers. For example, when individuals who have died or who are no longer licensed are not marked or removed from the register there is an overestimation of the number of drivers.

Consequently, the following information should be available in the national registers on an annual basis:

- the total number of active driver licences
- the number of driver licences by licence group and by age group of the driver.

3.2.3 Road length

Road length data is a practical exposure variable for the estimation of traffic risk at the network level. The variables selected have to be compatible with the respective accident data concerning road. Thus, type of road, area type and region/municipality are regarded as useful variables.

Information on road length by area type or region may be available in local authorities, while for the main road network data may be available in national authorities. In order to aggregate the existing information, the cooperation of several authorities responsible for the operation and maintenance of road network is needed, while procedures such as national questionnaires could be developed on that purpose.

If relevant data are not available, national authorities should carry out the required activities for collecting this information. Road length data may be collected on-site, using vehicles equipped with odometers, or with maps. In both cases, care must be taken in order to adequately handle intersection areas and avoid double measuring their length.

3.2.4 Vehicle fleet

While the best estimation of exposure can be given by the number of vehicle-kilometres, such data are not always available and are very expensive to collect. In the case that these data are available, they are not always reliable. Therefore, the second best exposure indicator is considered to be the vehicle fleet, due to its correlation with the level of motorisation.

Taking into account that the fatality risk is entirely different depending on the type of the vehicle (e.g. bus, car, or bike) it is necessary to make the comparisons in the respect of different vehicle categories. Consequently, the following information should be available in the national registers on an annual basis:

- total number of registered vehicles
- number of vehicles by vehicle type and by age group of the vehicle.

3.2.5 Vehicle kilometres

As mentioned before, the number of vehicle-kilometres is probably the most appropriate exposure indicator for the estimation of accident risk. Vehicle kilometres are a direct measure of traffic volume and can be available in a significant level of disaggregation, i.e. time, vehicle type, road type, driver characteristics etc.

However, in practice, the availability and the level of disaggregation of vehicle kilometres varies significantly and is strongly dependent on the type and features of the collection method used in each country. Moreover, the calculation of the exposure estimate is not consistent throughout countries resulting in a low overall compatibility. Vehicle kilometres are estimated by several methods, most of which include data collection by surveys and traffic counts. Furthermore, estimations are also carried out by the use of statistical models and combinations of methods.

In order to obtain a common and compatible risk exposure measurement unit, the definition of the indicator should be uniform between all countries. In the Glossary of Transport Statistics (Eurostat, 2003) a definition of vehicle kilometre is proposed, which could form the basis for a common definition:
"*Vehicle kilometre - Unit of measurement representing the movement of a road motor vehicle over one kilometre. The distance to be considered is the distance actually run. It includes movements of empty road motor vehicles. Units made up of a tractor and a semi-trailer or a lorry and a trailer are counted as one vehicle*".

Vehicle kilometre data are most useful for traffic risk analyses related to the vehicle and the road network. For the estimation of traffic risk at vehicle level, the vehicle type, vehicle age, vehicle engine size and road type are the most important variables, while the vehicle type, area type, road type and region variables are most important for the estimation of traffic risk at network level.

3.2.6 Person kilometres

Person kilometres can be collected either by travel surveys or by traffic counts and occupancy rate estimates. Travel surveys provide more detailed data than other methods. Moreover, data on person kilometres for non-motorized road users (bicycles and pedestrians) as well as cross tabulated data for age/gender groups of road users (both motorized and non-motorized) can be obtained only through surveys.

Person-kilometre data estimated by surveys are more usable for the variables: person class, person age and person gender and less usable for the vehicle type and the year. However, data are collected through surveys based on all these indicators.

Travel surveys are currently the most promising method available in order to have adequate data on person kilometres distributed by age/gender/road user. Thus, it is important to design the surveys in ways that allow for relevant risk calculations to be made. It is therefore recommended that travel surveys are conducted as follows:

- For risk exposure purposes travel surveys ought to be nationwide. Travel surveys in particular areas are less suitable because it is difficult to know how representative the area is, what the exact area covered is and it may be difficult to have precise correspondence between exposure data and accident data.
- Travel surveys ought to have sub samples distributed over a whole year (for instance sub samples every day) in order to account for seasonal travel variations.

- Travel surveys ought to include data also for professional drivers and travels conducted as part of work in addition to private travels.
- Travel surveys based on person samples often lack data for children. A possible way to obtain some data for children is to ask car drivers about age and gender of passengers.
- It is important to distinguish between travel made in a road traffic environment and travel made outside the road network. For pedestrians and cyclists this is particularly relevant.
- In order to reduce the problems with inaccurate reporting of distances and time, one should adopt tests of logic and reason to check answers.
- In addition to distance travelled one ought to try to register travel time as well.

3.3 Data collection methods

Since only seven African countries have available exposure data, a system could be established at African level in order to collect comparable exposure data. This collection system should focus on the collection of disaggregate time-series exposure data, by road user, transport mode and network characteristics, organised in order to provide data in a consistent and systematic way.

Different data collection processes should be included, which are:
1) Travel surveys, which may provide a higher level of disaggregation by using both vehicles and persons as units
2) Traffic count systems, which can provide continuous exposure measurements over time.

In order to obtain comparable data, a common exposure data collection framework should be developed, including both data collection processes.

Travel surveys
A travel survey at African level, would allow for the collection of exposure data cross-tabulated per person, vehicle and road network characteristics. Existing efforts concerning other countries, e.g. at European level, may provide useful insight regarding the type of data required and the recommended collection methods.

Then, methodological and practical issues should be tackled, such as the type of survey, the sample size, the target population, the duration etc. The creation of a common survey questionnaire and the development of a common methodology for calculating exposure indicators from the survey data, together with their confidence intervals, would also be very demanding tasks.

Therefore, the following steps are required for a travel survey at African level:
- Review of the travel surveys internationally.
- A common travel survey design for all African countries.
- Implementation of the survey design in each country.

As an example, a National Travel Survey (NTS) is carried out in England (https://www.gov.uk/government/collections/national-travel-survey-statistics). This NTS data collection consists of a face-to-face interview and a 7 day self-completed written travel diary, allowing travel patterns to be linked with individual characteristics. The NTS covers travel by people in all age groups, including children. Approximately 16,000 individuals, in 7,000 households in England, participate in the NTS each year. This method replies upon a sampling frame (list of households) from which a random sample of participants can be drawn in which satisfies pre-defined strata. This would present a challenge in parts of Africa where such a sampling frame would be less available.

Count systems

As mentioned before, traffic count systems can provide with continuous measurements over time, which would be useful for monitoring exposure. The first step is the implementation and operation of traffic count system in the African countries which will cover initially the main road network and will be expanded progressively to lower level of roads. It is also important that such a system would ensure vehicle classification in an accurate way (i.e. including two-wheelers). On that purpose, guidelines should be elaborated for traffic count systems.

There are two general methods used to collect traffic data: manual and automatic (US DoT 2013).

Manual refers to visually observing number, classification, vehicle occupancy, turning movement counts, or direction of traffic. Methods include using tally sheets or electronic counting boards and these are considered accurate and cost effective.

Automatic refers to the collection of traffic data with automatic equipment designed to continuously record the distribution and variation of traffic flow in discrete time periods (e.g. by 5 min., 15 min., hour of the day, day of the week, and month of the year from year to year). Automatic methods may include both permanent and portable counters. Permanent, continually operating traffic monitoring equipment is used to provide both current measures of traffic flow and to provide a time series record of traffic flow attributes that describe how traffic flow changes over time at that location. Permanent traffic monitoring locations should have:

- Long lived sensors that can withstand the harsh roadway environment;
- Power sources (either electrical power or solar power with battery backup);
- Communications (land lines or cellular communications); and
- Environmental protection (temperature, moisture, dirt, electrical surge protection on power and communications lines, and protection against animal and insect infestation).

Permanent sensors represent both a large financial investment and a large data resource.

Driver and vehicle registers

The improvement of national registers and the development of a common African register would be valuable to the improvement of vehicle fleet and driver population data. Both national and African registers should include disaggregate data, with transformation rules, updates and other improvements being implemented centrally. With the introduction of vehicle standards in Africa such as the MOT (annual basic safety check logging for every vehicle), opportunities to record annual odometer data (such as undertaken but the UK Department for Transport) become feasible.

3.4 Data collection priorities

The exposure data collection process in Africa is poor. As already mentioned a common framework should be established for collecting exposure data in order such indicators to be consistent but also comparable at both continent and international level.

However, the main methodologies to collect such data (travel surveys or traffic count systems), besides being expensive and difficult to manage from the organisational point of view, need time to show results. On the other hand, certain exposure indicators are more available, since their collection process is managed on a systematic basis from national governmental bodies.

Therefore, a 2-stage structure for exposure data collection is proposed (Table 3.1).

1st priority	**2nd priority**
Population	Road length
Driver population	Vehicle kilometres
Vehicle fleet	Person kilometres

Table 3.1: Overview of the proposed exposure data collection structure

4 Road Safety Performance Indicators

4.1 Introduction

Safety performance indicators (SPIs) are measures (indicators), reflecting those operational conditions of the road traffic system, which influence the system's safety performance. SPIs are aimed to serve as tools in assessing the current safety conditions of a road traffic system, monitoring the progress, measuring impacts of various safety interventions and making comparisons.

The performance indicators can be divided into four pillars - problem areas: road, vehicle, road user and post-accident care. Indicative indicators on these four pillars consist of:

- road: number and length of road safety audits conducted, number of identified high risk sites and related interventions
- vehicles: mean age of vehicle fleet, number of technical inspections
- road user: seat-belt use rates, helmet use rates, speeding, drink-driving and use of mobile phone while driving
- post accident care: number of staff working on it, number of ambulances.

In Africa, SPIs are focused mainly on behavioural aspects. However, such data, although highly prioritised by the questioned experts, are rather limited.

The present report aims to deliver recommendations for establishing the necessary systems of data collection, provide the definitions of variables and values for producing national SPIs in certain areas of the aforementioned pillars, and make them comparable at African level. These areas were defined based on the survey results as well minimum requirements based on the international practice. On that purpose, practices and manuals referring to international comparisons and harmonization were exploited, and a common framework to collect SPIs is proposed.

4.2 Definitions and description

4.2.1 SPIs on drink-driving

Alcohol use by road users and especially by drivers of motor vehicles increases the road accident risk considerably. Consequently, most countries ban the use of alcohol among drivers, or set low legal limits for blood alcohol concentrations. Nevertheless, a high proportion of fatal accidents involve drink-driving in most countries. Road safety policy makers need information about the state of this problem in their countries.

A SPI reflecting the alcohol related road toll is the **percentage of drivers under the influence of alcohol**.

Another more comparable indicator, which, however, seems to be out of line with the basic idea of SPIs, is suggested in the SafetyNet project and is based on accident data. The proposed SPI is the

percentage of severe and fatal injuries resulting from road accidents involving at least one active road user under the influence of alcohol.

In order to estimate the first indicator a sampling frame has to be defined, while for the second one a national system has to be set up. Medically trained persons should take the blood specimen and provide the respective results. It is also noted that amendments of the road traffic law may be needed in countries where alcohol testing of drivers involved in fatal accidents is not mandatory. The police should ensure that blood or breath samples are taken from all drivers involved in road accidents and should report the results to the agency responsible for national road accident statistics. As a minimum requirement, the SPIs should be produced and reported for each country and for each examined year.

4.2.2 SPIs on speeding

Speed is one of the main accident causes and has a direct influence on accident severity. Due to the massive character of speeding and inappropriate speeds, managing drivers' speeds has a high safety potential. Therefore, representative, reliable and valid speed data are needed in order to support policy decision makers.

Speed data usually contain a large amount of different information, such as information related to vehicles, roads, seasons and time that can be disaggregated and analysed in various ways. Measurement systems typically produce both individual and aggregated data, but there are also systems in use, which primarily measure speed data on an individual basis.

4.2.3 SPIs on the use of protection systems

The non-use of protection systems is associated with severe injuries and fatalities. Such systems are the seat-belts for vehicle occupants, the helmets for riders of powered two-wheelers and cyclists and the child restraint systems. The assessment of the use of protection systems in traffic allows for identifying the magnitude of the problem and preventing fatal injuries in road traffic.

The SPIs examined in this section are the following:

- daytime wearing rates of seat belts, in front seats (passenger cars + vans /under 3.5 tons), in rear seats (passenger cars + vans /under 3.5 tons), by children under 12 years old (restraint systems use in passenger cars), and in front seats (HGV + coaches /above 3.5 tons)
- daytime usage rates of safety helmets by cyclists, moped riders and motorcyclists.

The SPIs are estimated by conducting a national observational survey, where the measurements should be classified by type of road, such as motorways, rural roads and urban roads. The values for major road types are then aggregated into one indicator (of each type) for the country. It is important that the assessment is conducted on a regular basis (preferably annual).

4.2.4 SPIs on vehicles

The SPIs on vehicles are related to the level of protection afforded by the vehicles which constitute the fleet in a country. When accidents occur, the potential of the vehicle itself to prevent injuries can determine whether the outcome is a fatality or something less serious. Thus, improvements in passive safety do not affect the occurrence of accidents, but help to minimise the consequences

when accidents happen. Unsafe operational conditions could be defined as the presence within the fleet of a number of vehicles:
1) that will not protect the occupant well in a collision (accident worthiness)
2) with an increased capacity to inflict injury (compatibility).

The vehicles (passive safety) area differs from the other SPI areas, since the estimation of the indicators is not based on surveys, but the necessary data are taken from national databases. The minimum information which is required to produce some calculations of vehicle age (as a proxy for vehicle accident worthiness) and fleet composition (as a measure of compatibility), are total number of vehicles listed by:
• year of manufacture (or year of first registration)
• vehicle type (using definitions compatible with accident database).

4.3 Data collection and storage methods

Concerning the data collection for estimating SPIs, two main methods exist: the first one requires observational techniques on which results can be based and a sampling frame has to be defined, while the second ones needs mostly national statistics and data collected centrally by national registers.

The first approach concerns mainly the road related indicators, the use of protective systems and mobile phone while driving, while the estimation of the remaining indicators is based mostly on the second collection method. Consequently, different approaches are required for sampling the data.

4.3.1 Setting up a survey

The first step when setting up a survey is to define the sampling procedure in order to obtain a national sample. The main components to be determined are:
- the **survey population**, which are road users, vehicles, the total road network etc.
- the **sampling unit**, which may be individual, section of road etc.
- the **sampling design**, with the most frequently used being the following:
    1. Simple random sample: Each element has an equal chance of being selected. In order to carry this out, a list of all elements in the survey population is required, such that a sample can be selected at random.
    2. Stratified sample: the population is first divided in non-overlapping strata, after which a simple random sample will be determined in each stratum. Advantages, if the stratification is carried out properly, are: the variance is lower, while the pureness from random sampling is achieved; costs are often lower because of administrative reasons. Correct results are to be found in the general sample if one respects the proportions within the strata.
    3. Clustered sample: this is a simple random sample of groups (clusters) of elements. The advantage is that one does not need to have a complete list of elements of the population, only a complete list of the groups. It is often cheaper to carry out if one takes clusters e.g. based on geographical location.
    4. Multiple stage design: draw a simple random sample of clusters, first, and afterwards a simple random sample of elements within that cluster. This is typically used when the clusters are too large to question all elements. The advantages are the same as in a clustered sample. The

clustered and the multiple stage design sample also share some disadvantages. If the clusters are too intra-homogenous, if there are too few clusters or if they are of very unequal sizes the variance might increase rapidly.
- the **survey instrument**, which may be paper and pencil, interview, observational study, etc.

Then, the **sample size** has to be defined which is based on several factors, like the accuracy or the sampling error with which conclusions will be drawn. Thus, depending on the type of sampling design and the respective statistical background the sampling size can be defined.

In addition, if a nationally representative survey is not possible and only partial results are available, their transferability across time and place has to be examined. In order to obtain aggregate results that reflect the behavioural characteristics of the whole national population by examining only a few locations, the respective transformational rules have to be determined. However, this process may involve a considerable amount of time and is not necessarily the same for every country.

4.3.2 SPI collection methods

4.3.2.1 Drink-driving

Percentage of drivers under the influence of alcohol

Roadside surveys (alcotests) have to be performed, in which the blood alcohol concentration (BAC) of drivers will be measured. However, some methodological issues have to be clarified, so that the respective indicator is comparable among African countries, e.g. a common sampling and testing protocol has to be agreed.

Percentage of severe and fatal injuries resulting from road accidents involving at least one active road user under the influence of alcohol

The specific SPI can be implemented gradually, starting with the BAC of fatally injured drivers and then extending to all active road users involved in severe injury accidents. The requirements for such a task to be successful are:

1. Mandatory blood testing of all fatally injured drivers.
2. Mandatory breath/blood testing of all drivers involved in fatal accidents (whether or not the drivers are killed or injured).
3. Mandatory breath/blood testing of all active road users involved in fatal accidents.
4. Extension of procedures mentioned under 1-3 to severe injury accidents, starting with testing severely injured drivers and resulting in testing all active road users involved in severe injury accidents.

4.3.2.2 Speeding

There are three types of devices for collecting speed data:
- hand-held devices, such as radar and laser guns, that are handled by a human operator.
- in-road devices, loop detectors and axle detectors (pneumatic tubes, piezo-electric detectors, quartz-electric detectors)
- out-of-road devices, such as Doppler-based microwave radars, LIDAR devices (light detection and ranging), passive acoustic devices, active infrared devices and cameras.

The basic recommendations concerning speed data collection formed within the SafetyNet project are as follows:

1. Selection of measurement sites

- Should be based on a random procedure
- Procedure: random selection of road segments across the whole road network (first stage) and identification of appropriate locations on them (second stage)
- Appropriate location means:
  - Straight and uniform section of road
  - Section where it is possible to drive at a higher speed than the speed limit
  - Section with a small gradient (<5% on at least 500 meters preceding)
  - Away from junctions (>500 meters)
  - Away from any speed calming device (> 500 meters)
  - Away from road works (> 500 meters)
  - Away from pedestrian crossings (> 500 meters)
  - Away from any speed limit change or sign (> 1000 meters)
  - Away from work zones, parking zones, important roadside developments
  - Pavement surface in good condition
  - Away from the sections where the speed is used to be enforced by the Police
- Separate samples should be drawn for different road types
- Speed should be at least monitored on motorways, single carriageway rural roads and single carriageway urban distributor roads.

2. Period of measurement

- Avoid measuring speed during congestion-prone periods: peak hours, local events
- Avoid measuring speed under bad weather conditions
- Preferably measure speed during late spring or early autumn
- Concentrate measurements on typical working days
- For day measurements, measure between 09h30 and 15h30
- For night measurements, measure between 22h00 and 06h00

3. Practical considerations for the measurements

- Measure at least 200 vehicles
- Measure traffic count during the period of measurement
- Carefully follow the usage instructions of the device
- Document the maximum information on the measuring sites (including speed limit)

4. Data analysis

- Start with an error control of the data
- Exclude from the analysis all hour periods with more than 600 vehicles per hour and per lane
- Split the data for day and night periods
- Split the data per vehicle type
- Compute at least the following indicators:
  - Average speed for light vehicles during day
  - Average speed for light vehicles during night
  - Standard deviation of speed for light vehicles during day
  - Standard deviation of speed for light vehicles during night
  - V85 of speed for light vehicles during day
  - V85 of speed for light vehicles during night

- Percentage of light vehicles over the speed limit during day
- Percentage of light vehicles over the speed limit during night
- Percentage of light vehicles 10 km/h over the speed limit during day
- Percentage of light vehicles 10 km/h over the speed limit during night

5. Documentation and reporting

- Document all steps of the survey carefully
- Publish results at least annually

4.3.2.3 Use of protective systems (seat belts and helmets)

The requirements suggested in the SafetyNet project that should be fulfilled by the national system for producing SPIs on protection systems are presented below:

1. Design requirements

- Probability-based requirement
- Other sampling requirements (population, demographic, time/day requirements)
- The population is divided into the following types of road users:
  - A: Passenger car and vans occupants over 12 years old in front seats
  - B: Passenger car and vans occupants over 12 years old in rear seats
  - C: Children in passenger cars under 12 years old in front and rear seats
  - D: Occupants of coaches and heavy-duty vehicles over 12 years old in front seats
  - E: Occupants of coaches over 12 years old in rear seats
  - F: Pedal cyclists
  - G: Moped riders
  - H: Motorcyclists

In addition, the SPI values should be assessed by type of road user. Pedal cyclists and moped riders are excluded when examining motorways, while occupants of coaches and heavy-duty vehicles over 12 years old in front seats and occupants of coaches over 12 years old in rear seats are excluded when examining roads inside urban areas.

Moreover:

- Demographics concern regional units that can be included in the sample.
- Concerning time/day, all daylight hours for all days of the week can be included in the sample.
- The sample data should be collected through direct observations of protective systems' use by road users on roadways.
- Protection systems' use shall be determined by observation of the use or non-use of the protection systems.
- Observation is performed on road profiles, intersections, petrol stations or other eligible locations such as in the vicinity of shopping centres. Automatic video devices can be used as well.
- The observations should be performed by independent observers (not uniformed police or other officers).
- Instructions to observers should specify which road section and which direction of traffic on that road are to be observed.
- Observers should follow clear instructions on how to start and end an observation period and how to stop and start observations if traffic flow is too heavy to observe all concerned individuals or if they begin moving too quickly for observation.

2. Documentation requirements
All sample design, data collection, and estimation procedures used in country surveys must be well documented.

3. Requirements for the presentation of results
The values of all SPIs and the conversion rates have to be presented both numerically and graphically on an annual basis.

4.3.2.4 Vehicles

The "best practice" for vehicle fleet databases is to ensure that systems exist for maintaining the accuracy of the database. This means that procedures should be in place to:
1) Remove scrapped vehicles from the database.
2) Ensure that vehicles that are not taxed and/or licensed still appear on the database if they are still being used on the roads.

One way of ensuring that data meets these requirements is that vehicle fleet data is the responsibility of a national governmental body.

Additional requirements include:
1) Provide detailed and accurate descriptions of vehicle makes and models.
2) Classify vehicles according to vehicle-types compatible with CARE definitions.
3) Distinguish between smaller (less than 3.5 tonnes) and larger goods vehicles, since these are significantly different when assessing their compatibility in collisions with passenger cars or vulnerable road users.
4) Register all motorised vehicles, including public service vehicles and mopeds.

4.4 Data collection priorities

Based on the results of the extensive questionnaire as well as the exploited international reports (WHO, IRF), information on road SPIs are limited in African countries.

Concerning the data collection for estimating SPIs, two main methods exist: the first one requires observational techniques on which results can be based and a sampling frame has to be defined, while the second one needs mostly national statistics and data collected centrally by national registers. Since the second method is more easy to implement and far more available at present in many African countries, the related SPI data consist a priority in the collection process. Therefore, the following 2-stage structure for SPIs collection is proposed (Table 4.1).

1st priority	2nd priority
Number of vehicles by year of manufacture (or registration year)	% of drivers over legal limits
Number of vehicles by vehicle type	% of severe or fatally injuries attributed to alcohol
	Speeding
	Daytime wearing rates of seat-belts
	Front seats (passenger cars+vans)
	Rear seats (passenger cars+vans)
	Child restraint systems (children <12 y.o.)
	Front seats (hgvs)
	Daytime wearing rates of helmets
	Motorcyclists
	Moped riders
	Cyclists

Table 4.1: Overview of the proposed road SPI collection structure

5 Recommendations

5.1 Overall synthesis

The present deliverable aims to produce recommendations for a minimum set of harmonised data collection procedures and definitions that could be applied in the short- to medium term to improve African data collection systems. On that purpose, relative manuals from European and international projects were exploited by giving emphasis on the collection systems and definitions of three types of data: accident data, exposure data and road safety performance indicators. The recommendations for all types of data consist of a **minimum set of data elements and a common collection system**.

When developing a **common accident data system**, the minimum data elements should be defined based on selection criteria, concerning the usefulness of the selected variables and values, the level of disaggregation and the difficulty of their collection. All variables and values should refer to casualty road accidents. Additionally, the accident data structure is suggested to comprise four categories of variables, which are related to crash, road, vehicle and road user characteristics.

As far as **road accident data** are concerned, the police plays the major role in the data collection process, since they are the first who record the needed data, finalise them after the period of 30 days and forward them to the responsible national authority. The data collection form is recommended to be revised frequently, include detailed information on the vehicles and road users involved in the accident, as well as adopt all existing standardized international definitions of variables and values.

Concerning **road fatalities**, the international 30-days definition is recommended to be adopted by the African countries. On that purpose, the countries that are not currently utilizing such a definition should modify the data collection process and develop appropriate conversion factors. Underreporting is also an issue that should be tackled, so that the databases are further improved and comparability of the data among the countries is reached. The data are recommended to be adjusted by means of linking Police data with hospital data.

Regarding the exposure and performance indicators, the respective variables and values are recommended to be defined in such a way that they will be compatible to the accident data. The **exposure measures** concern two groups of data, the road traffic estimates and the road user at risk estimates. The recommendations of the present report include a list of primary data that should be collected in order to calculate the exposure indicators, as well as additional information that could be collected at a next step. The collection processes examined concern travel surveys and traffic count systems, while national registers may also provide with useful and commonly used exposure data, such as population, drivers' population, vehicle fleet etc.

Two mainly data collection methods exist for estimating the **road safety performance indicators**: the first one concerns observational techniques, while the second needs national statistics and data collected by national registers. Specific recommendations are given for each of the examined core areas; namely drink-driving, speed, use of protective systems and vehicles safety. In general, these recommendations concern the survey requirements (design requirements, measurement requirements, period of surveys etc.), data analysis and documentation and reporting of the final results.

The present report delivers specific recommendations on three types of data: accident data, exposure data and road safety performance indicators. However, due to limited experience, unavailability and lack of standardization in the collection process of such data for most African countries, a 2-fold priorities scenario is proposed on each data type, based on a combination of usefulness and ease to collect.

5.2 General implementation roadmap

In order to implement the above mentioned recommendations for a common data collection system and definitions certain prerequisites need to be met:

- **Establishment of capacity at the authorities to collect, process, analyse data and support decision making**

  It is very important for the authorities to be trained in basic practices of data management in terms of collection, storage and forwarding the files for further assessment. The overall intention is to develop a culture of substantiated decision making on all the organizations involved. The bodies involved in this capacity building should be the Police, the hospitals, and the public organizations involved in surveys for collecting exposure and SPI data. Special emphasis should be given in the underreporting of road accident data, which could be tackled by linking Police and hospital data.

- **Summary sampling and costing**

  The data elements should be comprehensive, concise, and refer to casualty road accidents. Demanding data in terms of time, cost, and collection barriers in general, should be avoided regardless of their value for road accident analysis. In Africa, not all the countries have the same performance level in terms of road safety definitions and data collection systems. Therefore, a 2-stage priorities scenario on accident data, exposure data and road safety performance indicators is proposed.

  As far as their cost is concerned, the collection of 1st priority data (accident, exposure and SPIs) concerns national data that are expected to be available in national databases, and therefore, no significant costs are involved. The upgrade of data collection forms, data definitions, adoption of the recommended international protocols etc. can be typically undertaken by the relevant authorities within their ongoing relevant activities. However, in several cases this also implies administrative restructuring and related significant time and effort, the cost of which requires a thorough country-specific investigation.

  Regarding the 2nd priority data, the cost of surveys depends on the sample size, which in turn depends on the country size. A travel survey for a country of population higher than 10 million inhabitants requires a minimum sample of 5,000 of survey respondents, in order to be able to record exposure per the basis road user, vehicle and road type characteristics. For a computer-assisted phone survey, and taking into account country variations of unit costs, the cost of a travel survey is estimated to range between 45,000 and 90,000 euros. A respective sample for a country of 5 to 10 million inhabitants would be 2,500 respondents, and the respective cost of the survey would range from 22,500 to 45,000 euros.

Three types of roadside surveys are required for the 2nd priority exposure and SPIs:

- alcohol survey: for 8 sampling points per country (covering different road and area types) for each survey, it is estimated that 6 hours of daytime and nighttime roadside controls for 7 days can provide a sample of ~6,800 road users (per road user, vehicle and road type), with the costs ranging from 33,500 to 67,000 euros.
- speed survey: for 8 sampling points per country (covering different road and area types) for each survey, it is estimated that 3 hours of daytime roadside observations for 7 days can provide a sample of ~33,500 road users (per road type and vehicle type), with the costs ranging from 8,500 to 16,500 euros.
- use of protection systems survey: for 8 sampling points per country (covering different road and area types) for each survey, it is estimated that 6 hours of daytime roadside observations for 7 days can provide a sample of ~16,800 road users (per road user, vehicle and road type), with the costs ranging from 8,500 to 16,500 euros.

It is noted that for countries with very high population, it is recommended to implement the surveys for a representative region.

- Adopt standard data definitions and standard data collection processes

Data elements and values must be useful for road accident analysis at both national and international level. Depending on the data elements, the collection process should be standardised and performed either when a road accident occurs (accident data) or on a periodic basis (exposure data – SPI surveys).

- Dedicated budget

Based on worldwide experience, countries with dedicated road safety budget, have a higher operational level of road safety.

- Formation of Pan-African coordinative organization

In order to assess the standardization level of the data collection process in the whole African continent and define data collection priority areas for further improvement, there is a need for structuring an organization comprised by African public bodies, research institutions and NGOs. Among the core assignments of this organization is to coordinate the data collection management as well as the support, from a technical point of view, the monitoring, analysis and publishing process of the data.

5.3 SaferAfrica implementation roadmap

Within the context of the SaferAfrica project, the recommendations for a common data collection system and definitions need to be rapidly conversed to the involved local authorities of each African country. Therefore, a **network of national experts** should be defined and spread out geographically to cover Africa. This network, managed by a SaferAfrica coordinator, will be in charge not only of distributing the relevant recommendations for data, but also of addressing the needs of the otherproject activities and tasks as well.

Finally, in order to assess the reliability of the proposed recommendations as well as the efficiency of the network of African experts, certain case studies need to be defined and examined.

The implementation roadmap consists of the following steps:

I. Identify data set needed as well as costs
II. Secure funding
III. Carry out regular data collection
IV. Process (data base) and analyse

6 References

Atalar, D., Talbot, R., Hill, J. (Eds) Traiing Package including training manuals and draft protocols, Deliverable 2.3 of the EC FP7 project DaCoTA

Duchamp G. and Treny V. (Eds.) (2008) *Risk Exposure Data – Recommendations for collection and exploitation.* Deliverable D2.5 of the EU FP6 project SafetyNet.

Eurostat (2003) *Glossary for transport statistics*. Document prepared by the Inter-secretariat Working Group on Transport Statistics, Third Edition.

Hakkert, A.S and V. Gitelman (Eds.) (2007) *Road Safety Performance Indicators: Manual*. Deliverable D3.8 of the EU FP6 project SafetyNet.

Hagstroem, L., Fagerlind, H., Danton, R., Reed, S., Hill, J., Martensen, H., Margaritis, D., Jahi, H, Morris, A. & Thomas P. (2010). Report on purpose of in-depth data and the shape of the new EU-infrastructure, Deliverable 2.1 of the EC FP7 project DaCoTA contract no: 233659.

US DoT Traffic Monitoring guide 2013:
https://www.fhwa.dot.gov/policyinformation/tmguide/tmg_fhwa_pl_13_015.pdf

WHO (2011) *Data Systems. A road safety manual for decision makers and practitioners*, World Health Organisation.

Yannis G., Evgenikos P. and Chaziris A. (Eds.) (2008) *CADaS - The Common Accident Data Set.* D1.14 of the EU FP6 project SafetyNet.

Yannis G., Papadimitriou E. and Chaziris A. (Eds.) (2008) *Risk Exposure Data Common Framework*. D2.3 of the EU FP6 project SafetyNet.

List of Abbreviations

EC	European Commission
EU	European Union
SPIs	Safety performance indicators