D.1.14 CADaS - The Common Accident Data Set

Contract No: **TREN-04-FP6TR-S12.395465/506723**
Acronym: **SafetyNet**
Title: **Building the European Road Safety Observatory**

Integrated Project, Thematic Priority 6.2 "Sustainable Surface Transport"

Project Co-ordinator:
Professor Pete Thomas
Vehicle Safety Research Centre
Ergonomics and Safety Research Institute
Loughborough University
Holywell Building
Holywell Way
Loughborough
LE11 3UZ

Organisation name of lead contractor for this deliverable:
National Technical University of Athens (**NTUA**)

Due Date of Deliverable: **31/10/2008**

Submission Date:

Report Authors:
George Yannis, Petros Evgenikos, Antonis Chaziris (**NTUA**)
Jeremy Broughton, Brian Lawton, Louise Walter (**TRL**)
Stefan Hoeglinger, Thomas Leitner, Andrea Angermann (**KfV**)
Niels Bos (**SWOV**)
Stig Hemdorff (**DRD**)
Peter Hollo (**KTI**)
Jan Tecl (**CDV**)
Lucy Rackliff (**VSRC**)
Jaime Sanmartin, Jean-Francois Pace (**INTRAS**)

Project Start Date: **1st May 2004** **Duration:** **4,5 years**

Project co-funded by the European Commission within the Sixth Framework Programme (2002 -2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (inc. Commission Services)	
RE	Restricted to group specified by consortium (inc. Commission Services)	
CO	Confidential only for members of the consortium (inc. Commission Services)	

Project co-financed by the European Commission, Directorate-General Transport and Energy

Table of Contents

Appendices

Transport
Project co-financed by the European Commission, Directorate-General Transport and Energy

Index of Tables

Index of Figures

Transport
Project co-financed by the European Commission, Directorate-General Transport and Energy

Executive Summary

SafetyNet is an integrated project that aims to bring together a wide range of accident data and related information to support National and EU level road and vehicle safety policy. When completed, the combined set of information will be available to the public over the internet to form a core element of the EU Road Safety Observatory.

Within Work Package 1 (WP1) of the SafetyNet Integrated Project, one of the project's seven Work Packages that deals with the further enhancement and exploitation of the CARE accident data, the **improvement of accident data compatibility throughout Europe** was attempted. As harmonisation of accident data at national level (apart from the EC level) could be very beneficial for road accident analysis, using more common variables and values across the European countries, a Common Accident Data Set (CADaS) and methodology were established, to be used by any EU country that wishes to update their national road accident collection system.

In order to develop this recommendation for a Common Accident Data Set, three main activities were under way during the Task 1.4 duration: Initially, **information on existing national road accident data collection systems was gathered**, allowing for the exploitation of available experience, but also the identification of any specific requirements in some countries. National road accident data collection forms, methodologies and data definitions from all EU countries were gathered through the members of the CARE Experts Group, with the assistance of the European Commission, and also **a relevant questionnaire was developed** and filled-in by representatives from almost all EU countries. Processing and analysis of this information allowed for the recording of the current potential of all EU national data collection systems.

At the same time, **the needs of the main stakeholders in different EU countries were identified**, in order to define the necessary data for road accident analysis. According to the specific circumstances in each country and the specific needs of each stakeholder, different needs were expected to be recorded, thus this activity took place at both national and local level. The main interest groups were Public Services (Police, Hospitals etc), Central Governmental Authorities (Transport, Health), Local Governmental Authorities, Research Institutes and Industry (including transport associations). **An appropriate Grid was developed** to establish a list of various stakeholders by country and then identify their needs for accident data. By filling-in this Grid for several stakeholders, the maximum needs were defined for each country and these were further compared, in order to identify the minimum/common needs for all countries examined.

On the basis of the outcomes of the above mentioned activities, the national collection systems from all EC member states were analysed and through an iterative process, taking into account both data availability and usefulness, but also variables and values of CARE and the experience of other international data files (US - MMUCC, WHO) the recommendation for a Common Accident Data Set was

formulated, consisting of a **minimum set of standardised data elements**, which will allow for comparable road accident data to be available in Europe. In this way, more variables and values with a common definition will be added to those contained in the CARE database, maximising thus the potential of CARE database and allowing for more detailed and reliable analyses at European level.

This set of data to be **voluntarily transmitted by each country to the EU**, will be derived from the national road accident data collection system, it can be implemented on a voluntary basis at the national accident collection systems and will be gradually adopted by the EU countries. It is structured in a simple way, without levels of hierarchy, constituting in fact the record layout of the data set to be transferred to the EU. Moreover, the structure of the CADaS variables allows for various levels of detail to be selected for providing the requested data, by the use of alternative (aggregate) values.

CADaS consists of **73 variables and 471 values**. The selection of these variables and values resulted from the balanced co-consideration of some **basic criteria**, taking into account that variables and values must be comprehensive, concise and useful for road accident analysis at EU level, the level of detail of the variables and values should correspond to all data useful for macroscopic data analysis and that each country should have the possibility to choose alternative level of detail of the various variables and values. Data which are impossible or very difficult to be collected are not retained in the CADaS, however, the future perspective of using certain variables and values was also taken into account, even though those data are not currently collected by most of the countries. Existing CARE variables and values are of first priority within CADaS and additionally, CADaS variables and values refer to casualty road accidents.

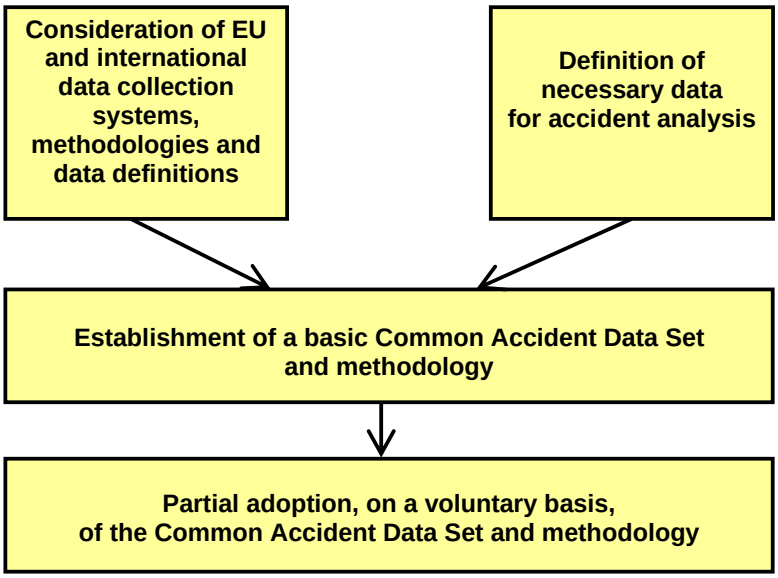
The CADaS variables are divided into **four basic categories**: **Accident** related variables, **Road** related variables, **Traffic unit** related variables and **Person** related variables. Several variables include two distinct types of values, referring to different level of detail: Detailed values, concerning information at the highest level of detail and alternative values, concerning information at a more aggregate level of detail, when more detailed values are not available.

The adoption of this recommendation by the European countries is very important to guarantee the success of this Task. CADaS can be adopted gradually by EU countries, but any part of it (variables, values, definitions and data formats) can be implemented within an existing national system, increasing the compatibility of national data with the respective CARE data. Subsequently, the **level of adoption of the CADaS can vary** according to any national needs and/or particularities and can be performed during any time in the future. Certainly, if one country wishes to enhance its national accident data or change its national system according to the CADaS protocol, it can adopt the CADaS proposal in pieces ("à la carte" system) or as a stand alone road accident data collection system.

1. Introduction

An important objective within SafetyNet WP1 is to **improve the compatibility of road accident data throughout Europe**. Currently, the CARE database contains a large number of road accident variables in disaggregate form, but it is acknowledged that more variables and values are necessary to better describe and analyse the road accident phenomenon at EU level. Additionally, due to differences in the collected data variables and values, their definitions, the differences of the accident data collection forms structures and the relevant data formats among the existing national databases, both accident data quality and availability are affected. Moreover, many data variables included in CARE lack reliability as the data are in many cases incomplete (few countries available or incomplete time series). Therefore the need for a common accident framework which would significantly enhance the CARE database with new and more compatible among the EU countries data and would allow for a comprehensive set of end products from all EU-27 member states to be progressively produced is considered essential. After elaborating this Common Accident Data Set, every EU country that wishes to update its data collection system could optionally and gradually adopt this proposed common data set.

A two-stage approach was adopted to achieve this, as it can be seen at the following diagram. On one hand, the data required for road accident analysis in several EU countries was identified and on the other hand, the current potential of the national data collection systems was recorded. The basic common accident data collection set and methodology were derived through an iterative process that took into account both data availability and usefulness, with the participation of experts and Governmental representatives.



Transport
Project co-financed by the European Commission, Directorate-General Transport and Energy

In order to establish a basic accident data collection set and methodology, **information concerning the existing national collection systems**, as well as the identification of the needs for road accident data are required. Within this framework, a questionnaire to collect information about the national accident collection forms, methodologies and data definitions in all EU countries was prepared. This questionnaire was initially developed by the Task 1.4 leader and subsequently all questions were thoroughly examined by all Task 1.4 partners, who contributed significantly to its further improvement. In the next phase, the recording and examination of national road accident data took place. Data elements, as well as the respective definitions used in each national system, were gathered and analysed in order to identify good practices in general, but also detailed variables and values for accident analysis. The results were exploited in the formulation of a recommendation for a Common Accident Data Set.

Moreover, the **identification of the needs for road accident data** was considered important for the establishment of a concrete proposal. On that purpose, the needs of the main stakeholders from several EU countries were recorded. According to the specific circumstances in each country and the specific needs of each stakeholder, different needs were expected to be recorded, thus this activity took place at both national and local level. The main interest groups were Public Services (Police, Hospitals etc), Central Governmental Authorities (Transport, Health), Local Governmental Authorities, Research Institutes and Industry (including transport associations). **An appropriate Grid was developed** to establish a list of various stakeholders by country and then identify their needs for accident data. By filling-in this Grid for several stakeholders, the maximum needs were defined for each country and these were further compared, in order to identify the minimum/common needs for all countries examined. Exceptional needs recorded, such as those of cyclists in the Netherlands could also be considered, but not for all countries. This Grid was distributed to all and was filled-in for the countries of the Task 1.4 partners (Greece, United Kingdom, Austria, Netherlands, Denmark, Hungary and Czech Republic).

After thorough co-examination of all information collected through the various activities of Task 1.4, the formulation of a complete recommendation for a **Common Accident Data Set (CADaS)** was carried out. This iterative process considered both data availability and usefulness, but also the currently used CARE variables and values and the experience of other international data files (US - MMUCC, WHO).

The recommendation for a **Common Accident Data Set consists of a minimum set of standardised data elements**, which will allow for comparable road accident data to be available in Europe. In this way, more variables and values with a common definition will be added to those currently included in the CARE database, maximising thus the potential of CARE and allowing for more detailed and reliable analyses at European level. CADaS is structured in a simple way, without levels of hierarchy, constituting in fact the record layout of the data set to be transferred to the EU. CADaS refers to the set of data to be voluntarily transmitted by each country to the EU, which should be derived from the national road accident data collection system. Moreover, the variables and values of

CADaS may be considered as recommendations for national police road accident data collection reports.

CADaS data elements were selected upon the basis of specific criteria. Existing variables and values of the CARE database were considered of first priority within the CADaS. Variables and values should be useful for road accident analysis and their level of detail should correspond to all data useful for macroscopic data analysis only. Each country should have the possibility to choose alternative level of detail of the various variables and values, while data which are very difficult to be collected were not retained in the CADaS. The future perspective of using certain variables and values was taken into account, even though those data might currently not be collected. The final version of the CADaS proposal includes 73 variables and 471 values which are sub-divided into specific categories according to their importance for road safety analysis and the level of detail.

The adoption of the CADaS recommendation by the European countries is a very important step towards the success of this Task. One of the CADaS advantages is that it can be **adopted gradually by EU countries**; however, any part of it (variables, values, definitions and data formats) can be implemented within an existing national collection system, increasing thus the compatibility of the national road accident data with the respective CARE data. If one country decides to start using the CADaS protocol, it can transform its national data into the CADaS data by using appropriate transformation rules and eventually transmit the transformed data to the EC. Consequently, the level of adoption of the CADaS can vary according to any national needs and/or particularities and can be performed during any time in the future.

Finally, the **adoption of CADaS should be supported by many different parties** in order to maximise its acceptance by the EU countries. National representatives at various levels (CARE Experts Group, High Level Group) could contribute by promoting the recommendation in their own countries. Additionally, a broader dissemination could be achieved through the ERSO website, but also through its presentation in scientific papers and national and international conferences. Finally, the assistance of the European Commission is important for promoting the CADaS and encouraging countries to implement pilot programmes for its implementation.

List of Task 1.4 Partners

Seven partners are mainly involved in Task 1.4 of SafetyNet Work Package 1 from seven different countries, as well as the Information Collection Coordinator (ICC) from the Loughborough University. Moreover the Research Institute on Traffic and Road Safety of the University of Valencia (INTRAS) participated on Task 1.4 as an affiliated member.

National Technical University of Athens	Greece	NTUA
SWOV Institute for Road Safety Research	Netherlands	SWOV
Kuratorium für Verkehrssicherheit	Austria	KfV
TRL Limited (Transport Research Laboratory)	United Kingdom	TRL
Centrum Dopravního Vvýzkumu (Transport Research Centre)	Czech Republic	CDV
Road Directorate - Ministry of Transport	Denmark	DRD
Közlekedéstudományi Intézet Rt (Institute for Transport Sciences Ltd)	Hungary	KTI
Vehicle Safety Research Centre - Loughborough University (ICC)	United Kingdom	VSRC
Instituto de Investigación en Tráfico y Seguridad	Spain	INTRAS

Transport
Project co-financed by the European Commission, Directorate-General Transport and Energy

2. National data collection systems

2.1 General

The establishment of a basic accident data collection set and methodology requires both the **recording of the existing national collection systems** and the identification of the needs for road accident data. Within this framework, it was decided to use both bottom-up and top-down approaches. One of the main activities of Task 1.4 consisted of preparing a questionnaire to collect information about the national accident collection forms, methodologies and data definitions in all EU countries. A draft version of this questionnaire was initially developed by the Task 1.4 leader (NTUA) and subsequently, all questions were thoroughly examined by all Task 1.4 partners, who made several recommendations for further improvement. Some general comments were:

a) The questionnaire should not get into details of national road accident databases but concentrate mainly on the national road accident collection forms and collection methodologies. Consequently, questions on technical issues (software used) should be omitted, also some more general questions on the national road accident collection systems.

b) Some questions concerning national accident databases could be combined with the respective ones on national accident collection forms in a single question.

c) Questions about stakeholders and generally accident data usage should be put in a separate section in the questionnaire.

d) The entire format of the questionnaire should be structured according to the information flow. Accordingly, questions on the general processing of national accident data should appear at the beginning of the questionnaire, followed by questions concerning data validation, national road accident data collection forms, stakeholders and national road accident databases.

e) It was decided that a single questionnaire would be prepared for Tasks 1.4 and 1.5 of WP1.

Moreover, the **recording and examination of the national road accident data** took place during the second phase of the project. Variables and values as well as the respective definitions used in each national system were collected and examined in order to identify good practices in data recording (how data are entered, whether several values may be attributed to the same variable etc) as well as identify important variables and values for accident analysis. The results from this study were exploited for the formulation of a recommendation for a common accident data set.

2.2 Analysis of questionnaires

The **common Task 1.4 and Task 1.5 questionnaire** (see Appendix II) was distributed through the European Commission to the members of the CARE Experts Group on February 7th 2005 and an initial deadline for feedback was set at the end of March 2005. The entire collection process of the questionnaires was facilitated by the ICC and feedback from 25 European countries was provided by the end of November (DE, EE, EL, ES, IE, CY, MT, AT, PT, SI, FI, CH, BE, CZ, DK, LV, SE, GB (not UK), NO, HU, PL, NL, SK, FR, LT). The responses from these countries were compiled and a preliminary analysis completed in June by the ICC. Further analyses were made by the Task 1.4 leader in early November, incorporating additional responses that were received late.

The questionnaire was divided into four different sections: National road accident collection system, National road accident data validation, Underreporting and Road accident data analysis. Some conclusions have been drawn about the current situation on accident data collection in European countries, based on the answers received.

Road accident collection system

The **main issues dealt within Task 1.4** of SafetyNet WP 1, were the potential of existing road accident collection systems, and the identification of the data needed for road accident analysis. This section of the questionnaire covers the current potential of national systems for accident data collection. More specifically, it relate to the length of time over which the collection systems have been developed, the content of national databases, the national procedures established for reviewing the system and the institutional arrangements which facilitate its successful operation.

The **original introduction of national road accident databases and collection forms** varied widely among the participating European countries. The first road accident collection form was introduced in Switzerland in 1930, and the first database in Austria in 1961, but no national road accident collection form and database existed in Malta before 1998. Moreover, in Hungary the first year that a road accident database was introduced was only recently, in 2002. The following Figure 1 displays the year when the national accident database and collection form were first introduced in each country.

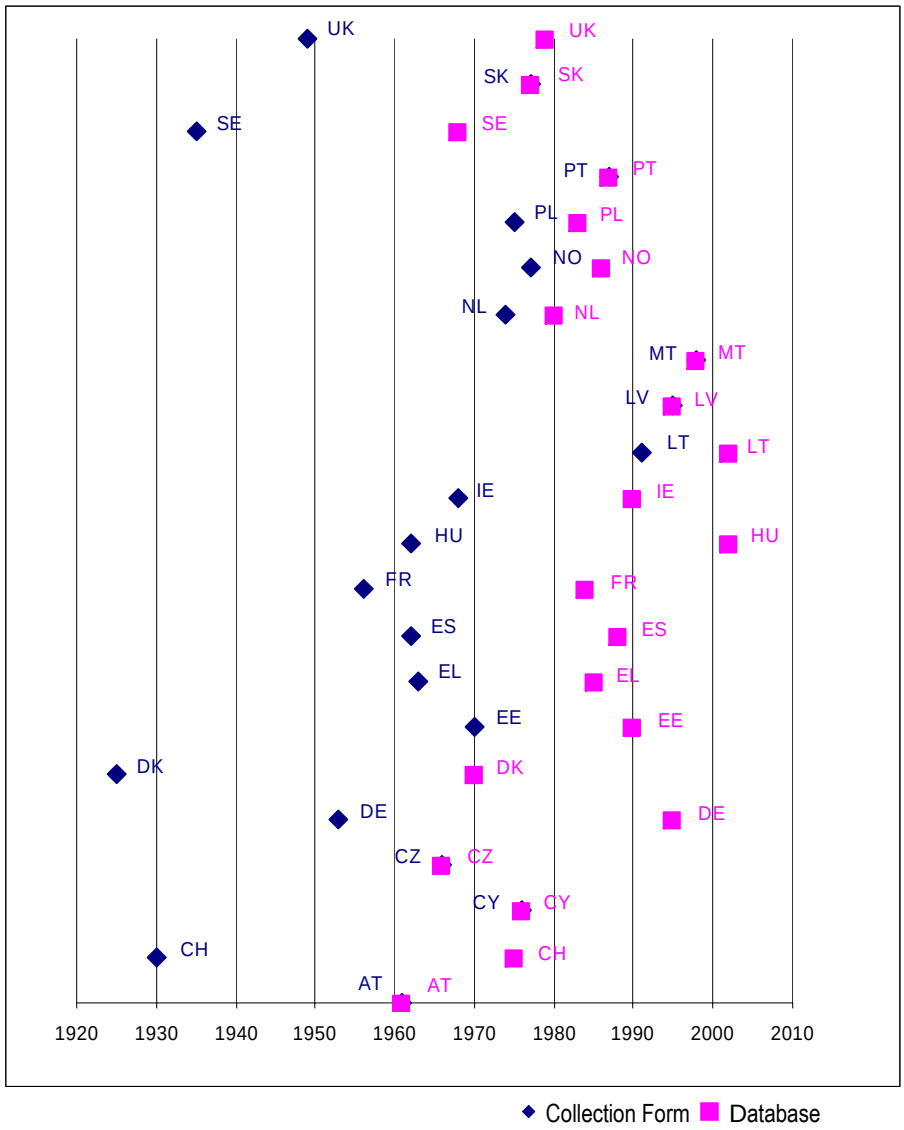


Figure 1. Road accident database and collection form first introduction (year)

Furthermore, **a number of changes are recorded in many European countries**, affecting both the collection form and the respective database. Some include definition changes (DE, ES, AT), incorporation of additional information, i.e. on alcohol consumption (EE), more detailed information (NO), usage of electronic collection form and linkage of databases (EE). Most of these changes occurred during the previous decade. It is surprising, however, that while the earliest data collection forms and databases were developed over 40 years ago, only in a few countries are the collection form and database revised regularly. A revision is rarely undertaken in most of the participating European countries, only when it is considered necessary; no review procedures are recorded in some countries. Although this does not indicate that collection systems are necessarily out of date, it might imply that institutional arrangements for review and change are poorly developed. Thus, the task of harmonising road accident collection forms and databases across Europe could be even more complicated.

Table 1. National Database and Collection Form Revision

	Revision of the National Database and Form			
	Yes			No
	10 yearly basis	5 yearly basis	Other	
AT				✓
BE				✓
CH			✓	
CY				✓
CZ			✓	
DE				✓
DK			✓	
EE				✓
EL	✓			
ES				✓
FI				✓
FR *	✓			
GB		✓		
HU			✓	
IE				✓
LT				✓
LV			✓	
MT				✓
NL			✓	
NO				✓
PL				✓
PT				✓
SE			✓	
SI				✓
SK				✓

* database only

In all participating countries, the authority responsible for entering national road accident data is the Police force, while National statistical offices, Police, Ministry of Interior / local authorities, and public authorities are ordinarily responsible for maintaining road accident databases. In most of the participating countries, however, **more than one authority is responsible for developing the respective database**. This may interfere with the potential harmonisation of the collection procedures among the European countries, as more than one agency is involved. Furthermore, more than one road accident database is maintained in many countries, mainly by hospitals, insurance companies, or Ministries; however, these are rarely linked.

Recording road accident data is an important task of the Police, so **the people involved should be properly trained**. The level of training could be a significant factor affecting the accuracy of accident data. In seven European countries (DE, SI, BE, LV, PL, FR and EE), however, no such specific training takes place. Furthermore, for the remaining countries there is some significant variation in the amount of training provided by the responsible authorities. For example, training in some countries simply includes the provision of guidelines, while in other countries a relevant manual is issued, including specific information on road safety issues.

Road safety analysis relies on accurate data, so this is an important task and it is surprising that in some countries very little or no training is provided.

Road safety data collected by the Police at National level are transferred to the responsible authority and entered in the respective database. Some countries apply a time limit to this data transfer, ranging from 24 hours to 8 weeks, while in BE and LV there is no such limit. However, it is not clear whether the existing time limits constitute a legal requirement rather than a guideline. Usually, data entered in the databases relate to casualties with fatal, serious or slight injuries, while additional data are collected in some countries and incorporated into the respective databases. The following Table 2 presents data that are included in the databases in each country. Finally, additional data are collected at local level in more than half of the countries, while CY, MT, NL, DE, SI, CZ, DK, LV, FR, LT, SK and PT have no additional local data collection.

Table 2. Data included in the National Road Accident Databases

	National Road Accident Database						
	Fatalities	Seriously Injured	Slightly Injured	Injured	Material Damages	Hospitalised	Other
AT	✓	✓	✓				✓
BE	✓	✓	✓				
CH	✓	✓	✓		✓		✓
CY	✓	✓	✓		✓		
CZ	✓	✓	✓		✓		
DE	✓	✓	✓		✓		
DK	✓	✓	✓		✓	✓	✓
EE	✓	✓	✓		✓		
EL	✓	✓	✓				
ES	✓	✓	✓		✓		
FI	✓			✓	✓		
FR	✓	✓	✓	✓		✓	
GB	✓	✓	✓				
HU	✓	✓	✓			✓	✓
IE	✓	✓	✓				
LT	✓	✓	✓	✓	✓		✓
LV	✓	✓	✓		✓		
MT	✓	✓	✓			✓	✓
NL	✓	✓	✓	✓	✓	✓	
NO	✓	✓	✓				
PL	✓	✓	✓				✓
PT	✓	✓	✓		✓		✓
SE	✓	✓	✓			✓	
SI	✓	✓	✓		✓		
SK	✓	✓	✓	✓	✓		✓

Based on the existing situation regarding road accident collection systems, it is commonly believed that **electronic completion of accident forms** would not only accelerate the entire collection procedure, and thus improve it, but also minimise the potential for data inconsistency. Electronic methods to collect data are already used in some European countries (CH, SE, SI, LV, IE, DE, FI, DK, CZ, SK, FR), while others are planning to use such methods in the near future (GB). Electronic

data entry across Europe could additionally facilitate harmonisation of the form and the related collection procedures at European level.

Furthermore, **the accident location** is recorded in all participating countries (apart from Lithuania), but in most countries these data lack precision. The use of satellite positioning systems (such as GPS, or Galileo in the future) in combination with Geographic Information Systems (GIS) could provide a very useful tool for ensuring the required accuracy in the recording of the accident location, thereby improving data collection.

Finally, regarding **the improvement of the road accident collection system** throughout Europe, all those involved should be better and more thoroughly trained to cope with this important task. Moreover, since the Police seem to be principally involved in data collection, their training and motivation should be improved.

National Road Accident Data Validation

This section of the questionnaire examines whether or not and to what extent the road accident data in the respective databases are **subject to validation**. Road safety experts were asked to assess the accuracy and reliability of their country's data.

Among the most important data limitations identified in many European countries is the fact that **not all accidents are reported to and recorded by the Police**. It was stated that the problem of underreporting mainly affects accidents in which vulnerable users are involved (pedestrians, two-wheelers, etc.) and accidents that occurred outside urban areas. Furthermore, in most countries certain items of road accident data are missing or are considered to be of low quality. These limitations interfere with reliable and constructive comparisons at both National and European level.

In several participating countries, the data that are most frequently missing relate to **material damage only accidents**, information on use of drugs and BAC level, or the existence of safety equipment (airbags, seatbelts, etc.), and data on speed. Nevertheless, eleven participating countries (DE, DK, EE, IE, CY, SI, GB, NL, LT, SK and SE) have reported that all necessary data are collected at national level.

In most of the participating countries the data collected at national level **are considered reliable, as far as road accident fatalities are concerned**, while the reliability of data for the generality of road accidents is mostly considered of medium or in some cases of low quality. The only nine countries that consider data for both accidents and for road accident fatalities to be of high quality are CY, DE, MT, AT, CZ, DK, LT, SK and GB. At the same time, in SIovenia the reliability of data for accidents is considered high, while the reliability of data for fatalities is considered medium.

Most of the participating countries **apply a validation methodology** that mainly consists of either internal validation checks in the database, checks for illogical or contradictory statements, cross-checking with other databases, or a combination

of all these. Only in DE, SI, SK, PL and MT no validation is applied. Furthermore, a number of consistency checks are carried out in almost all participating countries, with the exception of five countries (SI, SK, PL, IE, LV). These consistency checks usually include checking for double insertion of accidents to the database or random checks in order to ensure data coherence.

Underreporting

The only figures that could be considered comparable at international level are those **referring to fatal accidents**, as only these are considered reliable and have common definition among the European countries. Figures relating to accidents and injuries cannot be considered either reliable or comparable at international level, due to different definitions used as well as underreporting. Within this framework, the present section of the questionnaire considers definitions of road accident injuries, levels of underreporting, as well as relevant recent national studies.

With reference to comparability of road accident data, road safety experts from all participating countries were asked to **provide the definitions** used in their countries relating to serious and slight injuries. Table 3 summarises the answers and provides an insight as to whether these figures could be used for reliable comparison among European countries. More specifically, it seems that 10 countries use the same definitions regarding injury severity, which is the following: "*Seriously injured is a person who is hospitalised for at least 24 hours*", while "*Slightly injured is each non-fatally injured person who is not seriously injured*". The remaining countries either use different definitions (AT, GB, CH, IE, DK, MT, NO, HU, PL, SK), or do not distinguish between serious and slight injuries (CZ, FI, EE, SI, LT).

Table 3. Definitions of seriously and slightly injured persons used

Countries	Definitions of seriously injured									Definitions of slightly injured					Criteria of injury degree
	Disability>25 hours	Disability > 42 hours	Hospitalised>24 hours	Hospitalised	Hospitalised+serious injuries	Serious injuries	Disability>24 hours, or very serious injuries	Hospitalisation>7 days + serious injuries		Other than seriously injured	Disability < 42 hours	Slight injuries	Hospitalisation<7 days + slight injuries	Not hospitalised	
AT							✓	✓		✓			✓		Disability+injuries
BE			✓							✓					Hospitalisation
CH	✓									✓					Disability
CY			✓							✓					Hospitalisation
CZ															-
DE			✓							✓					Hospitalisation
DK						✓						✓			Injuries
EE															-
EL			✓							✓					Hospitalisation
ES			✓							✓					Hospitalisation
FI															-
FR			✓							✓					Hospitalisation
GB					✓							✓			Hospitalisation+injuries
HU					✓							✓			Hospitalisation+injuries
IE					✓							✓			Hospitalisation+injuries
LT															-
LV			✓							✓					Hospitalisation
MT						✓						✓			Injuries (Police judgement)
NL			✓							✓					Hospitalisation
NO				✓										✓	Hospitalisation
PL															Hospitalisation+injuries
PT			✓							✓					Hospitalisation
SE			✓							✓					Hospitalisation
SK		✓									✓				Disability
SL															(Doctor's judgement)
Total	1	1	10	1	3	2	1	1		12		5	1	1	

The **lack of comparability of accident data among European countries** is a major limitation that interferes with reliable and effective road safety analysis. Tackling this problem has been discussed in recent years and the use of the proposed term “hospitalised casualty” by all European countries has been examined. A common definition for hospitalised casualty has not yet been defined, as this term does not exist in official statistics. Thus, only data referring to fatal accidents can be currently considered for comparisons at European level.

Regarding accident data underreporting, even though it is considered a major limitation in most countries, it has **not been thoroughly investigated over the last decades**. Recent underreporting studies are available in very few participating countries, mostly regarding the number of fatalities. However, plans for underreporting studies are under way in several countries (DK, EE, EL, ES, FI, GB, PT, NL, PL, FR), in some cases consisting of linkage of or comparison between road accident and medical databases (DK, ES, GB).

Furthermore, it was reported that **in Spain and France a correction coefficient is calculated**, which is related to fatalities within 30 days. This is considered necessary, as the national road accident database records the number of fatalities as those who died at the scene of the accident, or within 24 hours, while in France up until 2004 the number of fatalities was recorded as those who died within 6 days from the day of the accident. Thus, a correction coefficient is applied on these data in order to obtain reliable and compatible at European level figures.

Databases containing medical data are maintained in most participating countries, with the exception of DE, CY, MT, NO, CH, SI and LV. However, **few of them are linked and compared to the national accident databases**. Table 4 summarises the existing situation regarding medical databases and whether these are either linked or compared at national level.

Table 4. Databases containing medical data

		Medical database			
		Linkage		Comparison	
	No medical database	Yes	No	Yes	No
AT		✓			✓
BE			✓		✓
CH	✓				
CY	✓				
CZ			✓	✓	
DE	✓				
DK			✓		✓
EE			✓	✓	
EL			✓		✓
ES			✓	✓	
FI			✓	✓	
FR			✓	✓	
GB				✓	
HU		✓			✓
IE			✓		✓
LT			✓		✓
LV	✓				
MT	✓				
NL		✓		✓	
NO	✓				
PL			✓		✓
PT			✓		✓
SE		✓		✓	
SI	✓				
SK		✓			✓

Based on the above Table, it seems that **a very small number of medical databases are linked to accident databases**, while more than half of the countries in which medical databases are available have not yet attempted to compare the respective data. Systematic comparison and linkage among medical and accident data could provide an important and reliable insight into the magnitude of underreporting in each country, but technical issues complicate the task. Furthermore, in some cases it is difficult to identify the people recorded in the accident database, thereby preventing a reliable and effective comparison with the respective medical data.

Road accident data analysis

Road accident analysis at European level is an indisputable need, in order to identify, assess and overcome road safety problems effectively. Based on reliable and compatible databases throughout Europe, road accident analysis can develop significant road safety measures, which will improve the existing situation. Within this framework, the present section of the questionnaire focuses on road accident analysis. More specifically, road safety experts from the participating European countries were asked to provide insight as to who uses road accident data in their country and at what extent the government takes into consideration the results of the analysis when developing national policies on road safety.

The main **national road accident statistics** are usually published annually in both electronic form and hardcopy. Furthermore, in almost all of the participating countries (apart from CY and LV) national road accident data are also available at the internet, while in many cases an English version of these data is available. Table 5 provides the website addresses where these data can be found.

Table 5. National road accident statistics at the internet.

Country	Internet application		Website	English version	
	Yes	No		Yes	No
AT	✓		http://www.kfv.at	✓	
BE	✓		http://www.statbel.fgov.be	✓	
BE			http://www.ibsr.be		✓
CH	✓		http://www.statistics.admin.ch	✓	
CY		✓			
CZ	✓		http://www.mvcr.cz		✓
DE			http://www.destatis.de/verkehr		
DK	✓		http://www.vd.dk	✓	
			http://www.vejsektoren.dk		✓
			http://www.statistikbanken.dk	✓	
EE	✓		http://www.mnt.ee		✓
EL	✓		http://www.statistics.gr	✓	
ES	✓		http://www.dgt.es		✓
FI	✓		http://www.stat.fi/til/ton/index.html	✓	
FR	✓		http://www.securiteroutiere.gouv.fr/observatoire	✓	
GB	✓		http://www.DfT.gov.uk/transtat/casualties	✓	
HU			http://www.ksh.hu	✓	
IE	✓		http://www.nra.ie/PublicationsResources/ListofPublications/RoadSafety/	✓	
LT	✓		http://www.policija.lt/viesoj		✓
			http://www.transp.lt	✓	
LV	✓				
MT	✓		http://www.nso.gov.mt	✓	
NL	✓		http://www.rws-avv.nl	✓	
NO	✓		http://www.ssb.no/vtu		
PL	✓		http://www.krbrd.gov.pl		✓
PT	✓		http://www.dgv.pt	✓	
SE	✓		http://www.sika-institute.se	✓	
SI	✓				
SK	✓		http://www.nehody.sk		✓

In most European countries, disaggregate data analysis is usually carried out by the competent authority, as well as other organisations that have access to the disaggregate data files. The most common organisations that carry out such

analysis are **universities** (EL, IE, PT, GB, BE), **local / regional authorities** (IE, AT, GB, SE), **road administrations** (FR, SK) **and research institutes** (CZ, BE, FR). The only countries that do not conduct disaggregate data analysis, other than that carried out by the competent authorities are EE, ES, MT, SI, DK, NL and LT. Finally, in some countries (FI, PT, CH) access to disaggregate data files is granted under special agreements.

Regarding the **main users of road accident data**, a variety of users, the most common of which are summarized in the following Figure 2 were identified by road safety experts. National public administration, research and scientific institutions, as well as accident involved bodies are ordinarily the main users of road accident data. At the same time, industries and professional associations are the least common accident data users. Furthermore, in all participating countries road accident data are used by local / regional authorities, in order to improve local / regional road safety. In most cases, local / regional authorities carry out high risk sites analyses, cost benefit analyses, or design educational campaigns. However, it is noted by many experts that in practice there are many limitations, regarding the access granted to these authorities.

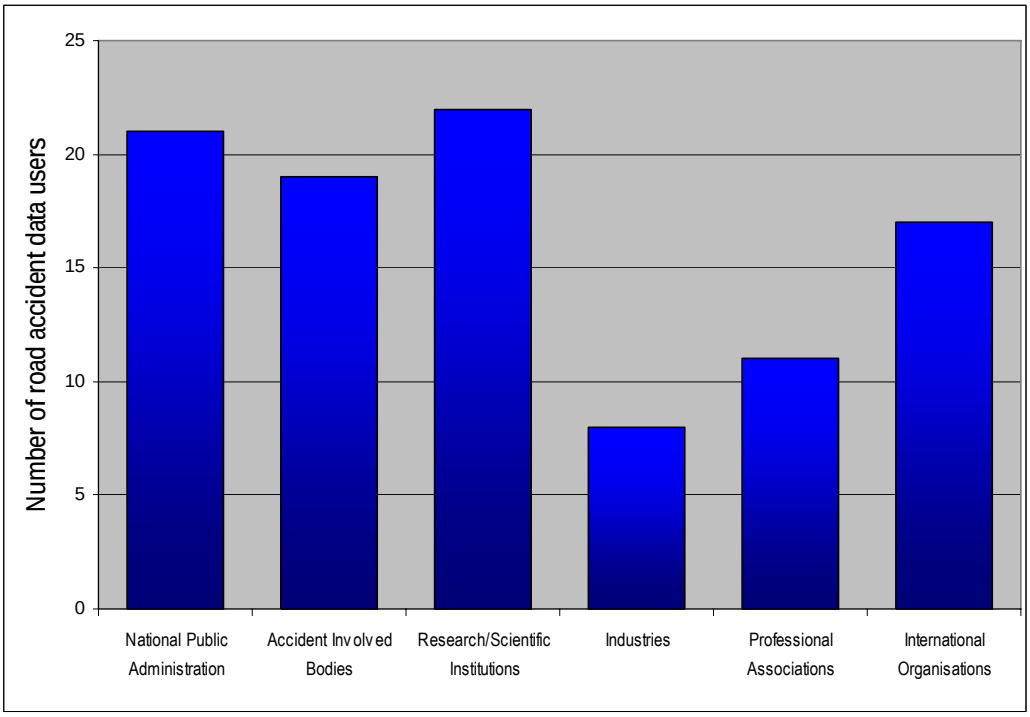


Figure 2. Road accident data users

It is commonly recognised by all countries that **road safety analysis and the derived results are seriously considered by each government** when developing national policies on road safety. However, in the case of Greece, Cyprus and Malta, this was suggested to be more in theory rather than practice, due to lack of systematic feedback. For the remaining countries, there are a number of ways in which road safety analysis can be used for policy and decision making, such as target setting (IE, EE, GB, DK) and introducing legislation regarding seat belt wearing or BAC limits (DE, IE, GB).

As far as the improvement of road safety analysis is concerned, experts from most of the participating countries believe that it is necessary to **improve road accident data quality** in order to improve analyses. Thus, common definitions should be used throughout Europe and exposure data should be collected at national level in each country. At the same time, the use of electronic collection forms could contribute to improving data quality. Furthermore, easier access to disaggregate data files should be provided, while accident databases should be linked with medical and other relevant ones. This will ensure better exploitation of road accident data. Figure 3 shows the most important actions, according to the national road safety experts, which could improve road accident analysis at local and national level.

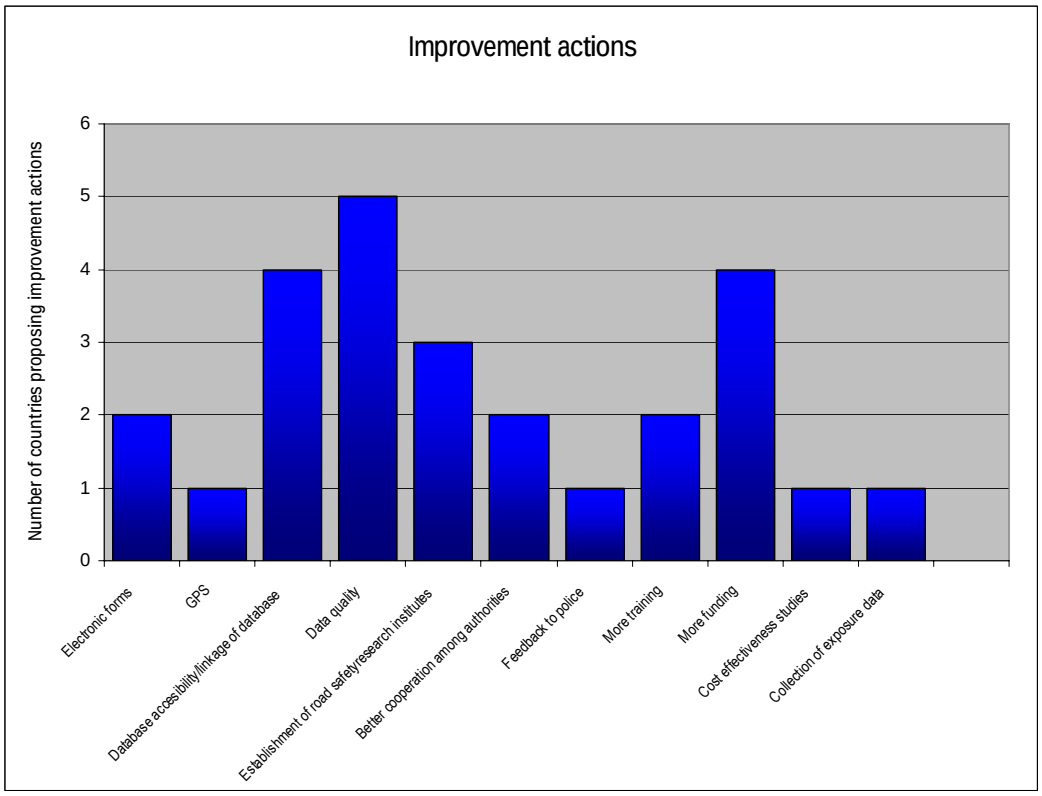


Figure 3. Actions to improve road accident analysis

Other important actions for improving road safety analysis include **better cooperation among all involved authorities and better funding** for studies of road safety issues.

2.3 National accident data collection forms

In order to **explore the current potential of road accident data collection systems** across Europe and identify the most necessary data for road accident analysis, a thorough examination of national road accident data collection systems is required. On that purpose, an analysis, including analytical information on each national collection system was carried out.

National road accident data collection systems differ significantly as regards the type of information collected. Some countries collect more information than others, but even the common accident data are not always compatible among the countries, due to different definitions of the collected variables and values. Moreover, the type and the level of detail of the collected data are different in each national system.

An analysis that would **provide sufficient information and enable comparisons between the national collection systems**, required specific information on the collected data (national road accident data collection forms), the way these data are recorded (instructions for the completion of the collection forms), detailed description of the data elements (list of variables and values) and detailed data definitions. Additionally, this information was requested in both electronic form and in hardcopy, enabling the further processing. Finally, all information was received in both native and English language for completeness reasons. If some of the documents were not fully available in English language, any available parts translated into English were requested.

The amount and type of the information required for this task, presupposed a well coordinated effort in order to establish the necessary links with the respective national experts. On that purpose a request was addressed to the members of the **CARE Experts Group** by the EC.

As long as the first set of information was received, **a thorough examination of the national accident data collection systems started**. However, in order to provide sufficient evidence and identify as accurately as possible the necessary road accident data at European level, information from as many countries as possible was necessary. Finally, information from 26 countries was gathered, as presented at the following Table 6.

Table 6. Overview of information collection

	Hard Copy								Electronic Form							
	Collection form		Instructions		Variables/Values		Definitions		Collection form		Instructions		Variables/Values		Definitions	
	Native language	English	Native language	English	Native language	English	Native language	English	Native language	English	Native language	English	Native language	English	Native language	English
AT	✓	✓	✓	✓					✓	✓	✓	✓				
BE	✓		✓		✓	✓	✓		✓		✓		✓	✓	✓	
CH											✓		✓	✓	✓	
CY	✓	✓	✓		✓	✓							✓	✓	✓	✓
CZ	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	
DE	✓		✓		✓	✓	✓	✓	✓		✓		✓	✓	✓	✓
DK									✓		✓	✓	✓	✓	✓	✓
EE													✓	✓		✓
EL	✓	✓	✓		✓	✓							✓	✓		
ES	✓	✓	✓	✓	✓											
FI									✓		✓		✓	✓	✓	
FR									✓							
HU	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓		
IE									✓	✓			✓	✓		
IT										✓				✓		✓
LT	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
LU																
LV									✓				✓	✓		✓
MT	✓	✓			✓	✓	✓	✓	✓	✓			✓	✓	✓	✓
NL									✓	✓						
NO														✓		
PL						✓								✓		
PT	✓		✓		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
SE									✓		✓		✓	✓	✓	
SI													✓	✓	✓	✓
SK	✓		✓						✓		✓	✓	✓	✓		✓
UK	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Transport
Project co-financed by the European Commission, Directorate-General Transport and Energy

According to Table 6, **almost all EU countries sent information concerning their national road accident data collection system**. However, the amount as well as the quality of information differed significantly between countries. More specifically, although many countries sent their national accident data collection forms, the documents were not available in English language and as a result the interpretation of the information was in many cases difficult. Moreover in many cases, the availability of the documents only in hard copies (and not in electronic form) required increased processing time.

Additionally, **not all countries transmitted all documents required**. For example, for a number of countries, only the data collection forms were gathered, without the respective variables and values definitions. As a result, these data could not be compared with data from the other countries or with the CARE variables and values. In some cases where at least partial information could be exploited (i.e. available data set and fill-in instructions), the information was taken into account in the formulation of the proposal for a Common Accident Data Set.

In total, **fourteen countries sent information in hard copies**. However, only in nine of these countries the collection forms were available in English language. Moreover, only eight countries sent the respective data definitions out of which, six sent a copy of these definitions in English. Finally, twelve countries sent data fill-in instructions five of which sent a copy of these instructions in English.

Concerning the electronic version of the documents, more information was gathered. **Eighteen countries sent the national collection forms** (nine in English) and thirteen countries sent data fill-in instructions (seven in English). Twenty two countries sent their list of variables and values in English (twenty also in native language) but only twelve of these could send an English version of the respective data definitions.

The amount of information gathered allowed for the **identification of the most common data elements used throughout the EU** for road accident data collection. These data were taken into account on the next step of the identification of necessary data for accident analysis (presented in the following Chapter 3) and further on the formulation of the recommendation for a Common Accident Data Set.

3 Identification of necessary data

3.1 General

Within the framework of Task 1.4, all partners agreed that the **needs of the main stakeholders in different EU countries should be identified** (top-down approach) in order to define the data needed for road accident analysis. As different needs were expected to be recorded, according to the circumstances of each country and of each stakeholder, this activity should take place at both national and local level. The main interest groups could be Services (Police, Hospitals etc), Central Government (Transport, Health), Local Government, Scientists and Industry (including transport associations).

An appropriate Grid was developed to establish a list of various stakeholders by country and then identify their needs for accident data. By completing this Grid for several stakeholders, the maximum needs were defined for each country and these were further compared in order to identify the minimum/common needs for all countries. Exceptional needs recorded, such as those of cyclists in the Netherlands could also be considered, but not for all countries.

Furthermore, it was agreed that the collection of needs from the various stakeholders could be accomplished in three ways:

a) By exploiting the reviews of national collection systems

b) By exploiting the existing experience of the partners' organisations from own consultant services to several national stakeholders

c) By exploiting any accident publication

It was also decided that the needs of each stakeholder would not just be identified, but also **ranked** according to a three level scale (low/high/ not used). Additionally, frequency of accident data requirements in each country was also be taken into account when deciding on the list of the accident data needs.

As there was no time in the process for the CARE experts to be thoroughly consulted, **an original matrix of needs** based on feedback alone was produced and an iterative process was further used to further improve the outcomes, based on expert judgement. The Grid was initially developed by the Task 1.4 leader and was further finalised according to the remarks of all Task 1.4 partners.

In general, the following three-step methodology was developed within the project:

Step 1 The original needs for accident data should be identified through several stakeholders in various EU countries, by creating and completing a Grid. The maximum needs were defined for each Task 1.4 partner's country, allowing for further comparison in order to identify the minimum/common needs for all countries.

Consequently, we came up with results mainly from six EU countries, as well as some general comments for Denmark.

Step 2 The original Grids were compiled and processed, in order to define the main accident data needs in the six countries. A list with all common requirements was prepared.

Step 3 Further discussion among all SafetyNet partners will be organised before the finalisation of the necessary data for accident analysis.

The Grids were compiled and processed and **three main Tables were developed**: one with all road accident variables used by several stakeholders, ranked by categories (Road user category, Vehicle category, etc), another with all road accident variables used by several stakeholders, absolutely ranked according to the level of usage (high/low/not used) and one Table where the main national stakeholders of road accident data in each country are identified, according to the level of usage of the various road accident variables. Several conclusions were derived by each of these Tables and were presented in the Work Package 1 meetings in Barcelona and Myconos. At the final stage of the Grids' processing, the most important road accident stakeholders were identified and their needs for data have been defined.

The **further collection of the national road accident data collection forms** enabled for the identification of the most common variables and values used at EU level, as well as the comparison of these data elements with the respective CARE variables and values. An indicative analysis performed by the Task 1.4 leader concerning accident related variables is presented in section 3.3. The information attained through the national collection forms, together with the analysis of the Grids, was exploited for the identification of necessary data for accident analysis at EU level, but also for the formulation of a proposal for a Common Accident Data Set. This proposal was formulated through an iterative process (each version was receiving several comments and a new version was prepared by the Task 1.4 leader) in which all Task 1.4 partners as well as the members of the CARE Experts group have contributed. This recommendation for a **C**ommon **A**ccident **Da**ta **S**et (CADaS) is presented in Chapter 4.

3.2 Analysis of Grids

Improving road accident comparability throughout Europe is the first step towards reliable and meaningful road accident analysis. Within this framework, the needs of the main stakeholders for accident data were identified.

This section presents the main results of the filled-in Grid, but also the analysis of these results concerning both the importance of stakeholders and the degree of use (high/low/not used) of road accident data.

The Grid (see Appendix III) **was distributed to all partners** in December 2004 and was filled-in by March 2005 for the countries of the Task 1.4 partners (Greece, United Kingdom, Austria, The Netherlands, Denmark, Hungary and Czech Republic). The Grid examines the needs of several existing main stakeholders for road accident data, focusing on which data they currently use and to what extent. The variables incorporated in the Grid were the ones commonly used in most EU countries and also included in the CARE system. The partners who filled-in the Grids could further include any other variables used in their countries, which are considered important. The scope of this task was to provide an insight to the needs of the main road accident data users, as these were identified by the Task 1.4 partners and not an extensive list of variables and values. The variables selected were divided into four main categories: variables referring to road user, variables referring to vehicle, variables referring to road environment, and variables referring to accident type. Furthermore, the stakeholders examined were grouped into six broad categories, covering most of the possible cases: National Public Administration (Ministries, etc.), Accident directly involved bodies (Police, etc.), Industry (vehicle industry, etc.), Research and Scientific Institutions (Universities, etc.), Professional Associations (freight transport associations, etc.), Other (touring assistance, etc.). These categories were decided by all Task 1.4 partners who filled-in the Grids, but again they could include in the list any other national road accident data users.

Based on the answers received by six of the participating European countries, **a ranking of the identified variables** was performed, by calculating the number of stakeholders (percentage of the stakeholders) in all European countries who considerably use these variables (Tables 7 and 8). Feedback from Denmark included only general information on the issues treated in the Grid, as it was mentioned that the groups of road accident stakeholders, as these were described in the Grid, are very different to the ones existing in the relevant Danish structure. At the same time, all identified stakeholders were also ranked according to the average number of road accident variables that they seem to use frequently (High use) (Table 9). The numbers in Tables 7 and 8 correspond to stakeholders that extensively use the respective variables, whereas the numbers in Table 9 correspond to variables used by the several stakeholders. Various conclusions were drawn from these Tables.

In Table 7 **the Ratio column for each road accident variable in each country is calculated** by dividing the number of users that systematically use the variable to the total number of national road accident users. This ratio is an

Table 7. Number of road accident data stakeholders per road accident variable and country

		Austria					Greece					Netherlands					Hungary					Czech Republic					United Kingdom					
	Variables	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	Av. Ratio
Road user	Age	13	1	9	23	57%	16	6	3	25	64%	12	0	0	12	100%	10	1	0	11	91%	23	1	0	24	96%	3	0	0	3	100%	**79%**
	Sex	13	1	9	23	57%	14	6	5	25	56%	12	0	0	12	100%	7	4	0	11	64%	23	1	0	24	96%	3	0	0	3	100%	**73%**
	Injury severity	13	1	9	23	57%	15	7	3	25	60%	6	0	6	12	50%	10	0	1	11	91%	23	1	0	24	96%	3	0	0	3	100%	**71%**
	Person class	13	1	9	23	57%	16	6	3	25	64%	10	0	2	12	83%	3	1	7	11	27%	23	1	0	24	96%	3	0	0	3	100%	**69%**
	Alcohol/drug test	13	2	8	23	57%	14	8	3	25	56%	9	0	3	12	75%	10	0	1	11	91%	14	9	1	24	58%	3	0	0	3	100%	**64%**
	Psychophysical circumstances	13	1	9	23	57%	13	9	3	25	52%	2	0	10	12	17%	6	3	2	11	55%	12	11	1	24	50%	0	0	3	3	0%	**47%**
	Car passenger type	12	1	10	23	52%	11	8	6	25	44%	2	1	9	12	17%	3	6	2	11	27%	12	12	0	24	50%	3	0	0	3	100%	**44%**
	Driving license age	13	1	9	23	57%	11	9	5	25	44%	5	1	6	12	42%	5	4	2	11	45%	8	7	9	24	33%	0	0	3	3	0%	**43%**
	Nationality	12	2	9	23	52%	11	10	4	25	44%	6	1	5	12	50%	4	5	2	11	36%	7	14	3	24	29%	0	0	3	3	0%	**41%**
	Hit and run	12	1	10	23	52%	11	9	5	25	44%	2	0	10	12	17%	3	5	2	11	27%	5	6	13	24	21%	1	1	1	3	33%	**35%**
	Average					55%					53%					55%					55%					63%					63%	**57%**
Vehicle	Vehicle type	13	1	9	23	57%	13	7	5	25	52%	11	0	1	12	92%	11	0	0	11	100%	16	8	0	24	67%	3	0	0	3	100%	**68%**
	Security equipment	13	1	9	23	57%	15	6	4	25	60%	8	1	3	12	67%	9	0	2	11	82%	9	9	6	24	38%	0	0	3	3	0%	**55%**
	Vehicle age	13	1	9	23	57%	13	6	6	25	52%	2	0	10	12	17%	7	1	3	11	64%	9	12	3	24	38%	0	2	1	3	0%	**45%**
	Average					57%					55%					58%					82%					47%					33%	**56%**
Road environment	Speed limits	12	2	9	23	52%	12	4	9	25	48%	10	0	2	12	83%	9	0	2	11	82%	20	0	4	24	83%	3	0	0	3	100%	**67%**
	Road type	11	2	10	23	48%	12	6	7	25	48%	10	0	2	12	83%	9	0	2	11	82%	20	4	0	24	83%	3	0	0	3	100%	**66%**
	Area type	11	2	10	23	48%	16	5	4	25	64%	10	0	2	12	83%	7	2	2	11	64%	20	4	0	24	83%	1	0	2	3	33%	**66%**
	Road surface conditions	12	1	10	23	52%	9	8	8	25	36%	10	0	2	12	83%	8	1	2	11	73%	17	3	4	24	71%	2	0	1	3	67%	**59%**
	Region	11	2	10	23	48%	15	4	6	25	60%	11	0	1	12	92%	1	6	4	11	9%	16	6	2	24	67%	2	1	0	3	67%	**57%**
	Junction control	11	2	10	23	48%	10	4	11	25	40%	9	1	2	12	75%	8	1	2	11	73%	15	5	4	24	63%	2	0	1	3	67%	**56%**
	Road markings	11	2	10	23	48%	10	3	12	25	40%	9	1	2	12	75%	6	3	2	11	55%	18	2	4	24	75%	0	0	3	3	0%	**55%**
	Junction type	11	2	10	23	48%	10	6	9	25	40%	9	1	2	12	75%	6	2	3	11	55%	15	5	4	24	63%	2	0	1	3	67%	**54%**
	Number of lanes	11	2	9	23	48%	9	4	12	25	36%	9	1	2	12	75%	6	3	2	11	55%	16	4	4	24	67%	2	0	1	3	67%	**54%**
	Carriageway type	11	2	10	23	48%	9	4	12	25	36%	8	2	2	12	67%	6	3	2	11	55%	15	5	4	24	63%	2	0	1	3	67%	**52%**
	Lighting conditions	12	1	10	23	52%	12	6	7	25	48%	10	0	2	12	83%	8	1	2	11	73%	6	13	5	24	25%	2	0	1	3	67%	**51%**
	Weather conditions	12	1	10	23	52%	9	7	9	25	36%	10	0	2	12	83%	7	2	2	11	64%	6	13	5	24	25%	2	0	1	3	67%	**47%**
	Average					49%					41%					79%					57%					64%					64%	**57%**
Accident type	Accident type	12	1	10	23	52%	15	7	3	25	60%	11	0	1	12	92%	10	0	1	11	91%	19	4	1	24	79%	3	0	0	3	100%	**71%**
	Collision type	12	1	10	23	52%	15	7	3	25	60%	11	0	1	12	92%	9	1	1	11	82%	19	4	1	24	79%	3	0	0	3	100%	**70%**
	Vehicle manoeuvre	12	1	10	23	52%	14	8	3	25	56%	9	1	2	12	75%	6	3	2	11	55%	9	10	5	24	38%	2	0	1	3	67%	**53%**
	Pedestrian movement	11	2	10	23	48%	14	8	3	25	56%	10	0	2	12	83%	8	1	2	11	73%	5	14	5	24	21%	2	0	1	3	67%	**51%**
	Average					51%					57%					83%					70%					54%					83%	**61%**

* Total number of national road accident data users

Scale of significance: H high importance L low importance - Not used

Transport

Project co-financed by the European Commission, Directorate-General Transport and Energy

indicator for the use of each variable. The average ratio is calculated as the weighted ratio of all ratios and provides an insight of the level of use of each main variable category in the several countries. Finally, in the last column of Table 7, the calculated average ratio for each variable is a weighted ratio deriving from the answers of the six countries and indicates the degree and frequency of use and in a way the significance of each variable.

Table 7 shows that road accident variables related to **the category of accident type** seem to be more important for stakeholders in general, as on average they are highly used more frequently (61%), followed by the respective of the road environment and road user category (57%).

Furthermore, variables regarding **the vehicle category** are not considered highly important, with the exception of the vehicle type variable which is widely used by most of the stakeholders (68%), followed by information on the existence of security equipment in the vehicle (55%).

Among the variables related to **the road user** category, age (79%), gender (73%), injury severity (76%), and person class (69%), are frequently used by the stakeholders, unlike nationality (41%) and hit and run accidents (35%).

As far as **the road environment** variables are concerned, speed limits (67%) and road and area type (66%) are among the most highly used road accident data, while lighting conditions (51%) and weather conditions (47%) seem to be considered less important by most of the stakeholders.

Variables regarding **accident type** (accident type, collision type, pedestrian movement, vehicle manoeuvre) seem to be considered as important, as they are all highly used by more than 50% of stakeholders.

In the next Table 8, **the road accident variables were absolutely ranked**, regardless their category, according to their average ratio, calculated as described above. Such ranking also allows to identify which variables are more frequently used by the various stakeholders, independently of the country. From this Table 8, it can be seen that generally, 22 out of 29 variables are highly used by half or more of the stakeholders in the six countries, while it seems that the variables mostly used by the stakeholders (over 70% of the stakeholders use them) are those related to the road user's age (79%), gender (73%), injury severity and accident type (71%), as well as to the collision type (70%). Other important road accident variables for the various stakeholders (over 50% of the stakeholders use them) concern: person class, vehicle type, speed limits, road and area type, alcohol/drug test, road surface conditions, region, junction control, security equipment, road markings, junction type, number of lanes, vehicle manoeuvre, carriageway type, lighting conditions and pedestrian movement.

At the same time, the least used variables when conducting road accident data analyses, are related to driving licence age (43%), nationality (41%) and hit and run accidents (35%).

Table 8. Number of road accident data stakeholders per road accident variable and country - ranked according to average ratio

	Austria					Greece					Netherlands					Hungary					Czech Republic					United Kingdom					
Variables	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	Av. Ratio
Age	13	1	9	23	57%	16	6	3	25	64%	12	0	0	12	100%	10	1	0	11	91%	23	1	0	24	96%	3	0	0	3	100%	**79%**
Sex	13	1	9	23	57%	14	6	5	25	56%	12	0	0	12	100%	7	4	0	11	64%	23	1	0	24	96%	3	0	0	3	100%	**73%**
Injury severity	13	1	9	23	57%	15	7	3	25	60%	6	0	6	12	50%	10	0	1	11	91%	23	1	0	24	96%	3	0	0	3	100%	**71%**
Accident type	12	1	10	23	52%	15	7	3	25	60%	11	0	1	12	92%	10	0	1	11	91%	19	4	1	24	79%	3	0	0	3	100%	**71%**
Collision type	12	1	10	23	52%	15	7	3	25	60%	11	0	1	12	92%	9	1	1	11	82%	19	4	1	24	79%	3	0	0	3	100%	**70%**
Person class	13	1	9	23	57%	16	6	3	25	64%	10	0	2	12	83%	3	1	7	11	27%	23	1	0	24	96%	3	0	0	3	100%	**69%**
Vehicle type	13	1	9	23	57%	13	7	5	25	52%	11	0	1	12	92%	11	0	0	11	100%	16	8	0	24	67%	3	0	0	3	100%	**68%**
Speed limits	12	2	9	23	52%	12	4	9	25	48%	10	0	2	12	83%	9	0	2	11	82%	20	0	4	24	83%	3	0	0	3	100%	**67%**
Road type	11	2	10	23	48%	12	6	7	25	48%	10	0	2	12	83%	9	0	2	11	82%	20	4	0	24	83%	3	0	0	3	100%	**66%**
Area type	11	2	10	23	48%	16	5	4	25	64%	10	0	2	12	83%	7	2	2	11	64%	20	4	0	24	83%	1	0	2	3	33%	**66%**
Alcohol/drug test	13	2	8	23	57%	14	8	3	25	56%	9	0	3	12	75%	10	0	1	11	91%	14	9	1	24	58%	3	0	0	3	100%	**64%**
Road surface conditions	12	1	10	23	52%	9	8	8	25	36%	10	0	2	12	83%	8	1	2	11	73%	17	3	4	24	71%	2	0	1	3	67%	**59%**
Region	11	2	10	23	48%	15	4	6	25	60%	11	0	1	12	92%	1	6	4	11	9%	16	6	2	24	67%	2	1	0	3	67%	**57%**
Junction control	11	2	10	23	48%	10	4	11	25	40%	9	1	2	12	75%	8	1	2	11	73%	15	5	4	24	63%	2	0	1	3	67%	**56%**
Security equipment	13	1	9	23	57%	15	6	4	25	60%	8	1	3	12	67%	9	0	2	11	82%	9	9	6	24	38%	0	0	3	3	0%	**55%**
Road markings	11	2	10	23	48%	10	3	12	25	40%	9	1	2	12	75%	6	3	2	11	55%	18	2	4	24	75%	0	0	3	3	0%	**55%**
Junction type	11	2	10	23	48%	10	6	9	25	40%	9	1	2	12	75%	6	2	3	11	55%	15	5	4	24	63%	2	0	1	3	67%	**54%**
Number of lanes	11	2	9	23	48%	9	4	12	25	36%	9	1	2	12	75%	6	3	2	11	55%	16	4	4	24	67%	2	0	1	3	67%	**54%**
Vehicle manoeuvre	12	1	10	23	52%	14	8	3	25	56%	9	1	2	12	75%	6	3	2	11	55%	9	10	5	24	38%	2	0	1	3	67%	**53%**
Carriageway type	11	2	10	23	48%	9	4	12	25	36%	8	2	2	12	67%	6	3	2	11	55%	15	5	4	24	63%	2	0	1	3	67%	**52%**
Lighting conditions	12	1	10	23	52%	12	6	7	25	48%	10	0	2	12	83%	8	1	2	11	73%	6	13	5	24	25%	2	0	1	3	67%	**51%**
Pedestrian movement	11	2	10	23	48%	14	8	3	25	56%	10	0	2	12	83%	8	1	2	11	73%	5	14	5	24	21%	2	0	1	3	67%	**51%**
Psychophysical circumstances	13	1	9	23	57%	13	9	3	25	52%	2	0	10	12	17%	6	3	2	11	55%	12	11	1	24	50%	0	0	3	3	0%	**47%**
Weather conditions	12	1	10	23	52%	9	7	9	25	36%	10	0	2	12	83%	7	2	2	11	64%	6	13	5	24	25%	2	0	1	3	67%	**47%**
Vehicle age	13	1	9	23	57%	13	6	6	25	52%	2	0	10	12	17%	7	1	3	11	64%	9	12	3	24	38%	0	2	1	3	0%	**45%**
Car passenger type	12	1	10	23	52%	11	8	6	25	44%	2	1	9	12	17%	3	6	2	11	27%	12	12	0	24	50%	3	0	0	3	100%	**44%**
Driving license age	13	1	9	23	57%	11	9	5	25	44%	5	1	6	12	42%	5	4	2	11	45%	8	7	9	24	33%	0	0	3	3	0%	**43%**
Nationality	12	2	9	23	52%	11	10	4	25	44%	6	1	5	12	50%	4	5	2	11	36%	7	14	3	24	29%	0	0	3	3	0%	**41%**
Hit and run	12	1	10	23	52%	11	9	5	25	44%	2	0	10	12	17%	3	5	2	11	27%	5	6	13	24	21%	1	1	1	3	33%	**35%**

* Total number of national road accident data users

Scale of significance: H high importance L low importance - Not used

Transport

Project co-financed by the European Commission, Directorate-General Transport and Energy

Table 9. Number of road accident variables per stakeholder and country

	Austria				Greece				Netherlands				Hungary				Czech Republic				United Kingdom				Average
1. National Public Administration	High	Low	Total	Ratio	High	Low	Total	Ratio	High	Low	Total	Ratio	High	Low	Total	Ratio	High	Low	Total	Ratio	High	Low	Total	Ratio	
Ministry of Transport	29	0	29	100%	20	10	30	67%	30	11	41	67%	18	10	28	62%	29	0	29	100%	29	1	30	97%	31
Ministry of Public Works					17	13	30	57%	35	3	37	78%					17	10	27	59%					31
Ministry of Public Order					29	0	29	97%																	29
Ministry of Justice	0	2	2		4	13	17	13%	32	0	32	71%					29	0	29	100%					20
Ministry of Interior (local authorities)	29	0	29	100%	17	13	30	57%	32	0	32	71%					29	0	29	100%					30
Ministry of Health	0	13	13	0%	13	3	16	43%	36	1	37	80%	19	10	29	66%	8	7	15	28%					22
Statistical Office	29	0	29	100%	30	0	30	100%	6	0	6	13%	26	0	26	90%	9	13	22	31%					23
Local Authorities	29	0	29	100%	17	13	30	57%	27	3	30	60%	20	6	26	69%	16	10	26	55%	25	5	30	83%	29
2. Accident Directly Involved Bodies																									
Police	29	0	29	100%	29	0	29	97%	41	0	41	91%	13	15	28	45%	29	0	29	100%	14	3	17	47%	29
Insurances	29	0	29	100%	16	14	30	53%	35	0	35	78%	3	0	3	10%	29	0	29	100%					25
Hospitals	0	0	0	0%	13	4	17	43%	8	0	8	18%	7	2	9	24%	8	7	15	28%					10
3. Industry																									
Vehicle Industry	20	9	29	69%	17	7	24	57%									16	11	27	55%					27
Road Construction Industry	0	16	16	0%	17	13	30	57%									20	7	27	69%					23
Other Industries	0	0	0	0%	0	0	0	0%									0	10	10	0%					0
4. Research and Scientific Institutions																									
Public Institutes	29	0	29	100%	30	0	30	100%	43	2	45	96%	26	1	27	90%	17	11	28	59%					32
Private Institutes	29	0	29	100%	30	0	30	100%									16	12	28	55%					29
Universities	29	0	29	100%	30	0	30	100%	35	0	35	78%	22	6	28	76%	16	12	28	55%					30
5. Professional Associations																									
Freight Transport Associations	0	0	0	0%	16	8	24	53%									24	4	28	83%					17
Passenger Transport Associations	0	0	0	0%	8	16	24	27%									24	4	28	83%					17
Truck Drivers' Associations	0	0	0	0%	16	8	24	53%									24	2	28	83%					17
Other Associations	0	0	0	0%	0	0	0	0%									9	17	26	31%					9
6. Other																									
Touring Assistance	29	0	29	100%	8	16	24	27%					24	4	28	83%	14	14	28	48%					27
Disabled People Associations	0	0	0	0%	20	6	26	67%									12	15	27	41%					18
Institutes for Alcoholics	7	1	8	24%	10	10	20	33%									6	10	16	21%					15
Road Users' Associations	29	0	29	100%	0	0	0	0%					24	4	28	83%	19	10	29	66%					22

Transport
Project co-financed by the European Commission, Directorate-General Transport and Energy

In Table 9 the **number of road accident variables used by the several road accident data stakeholders in six European countries** are presented. Additionally, a distinction between high and low frequency of use of these variables takes place, allowing to identify which stakeholders require more variables than others. A relevant ratio is calculated by dividing the number of road accident variables, which are highly used by an authority and the total number of accident variables that the same authority uses. Based on these ratios for each country, as well as on the average number of road accident variables, which are used by each stakeholder, an overall picture on the correlation between the several stakeholders and the road accident data can be drawn. Consequently, stakeholders who extensively use road accident variables can be identified and their specific needs can be recorded. Several interesting conclusions can be derived from this Table 9.

As far as the several stakeholders are concerned, it seems that **breakdown of road accident data stakeholders is not applicable in all countries**. For example 25 stakeholders were identified in Greece, whereas in the United Kingdom there were only 3 important ones (Ministry of Transport, Police, and Local Authorities). However, a preliminary ranking of the importance of all stakeholders, according to the frequency and the extent of accident data use, is attempted.

Moreover, in Table 9 it can be seen that the total number of accident variables used by each stakeholder in the same country varies significantly, according to their domain of interest and also the extent of the data analyses performed.

The "National Public Administration" **common stakeholders most often involved in road safety issues** are: Ministry of Transport, Ministry of Public Works, Ministry of Interior, Ministry of Public Order and Local Authorities. These stakeholders frequently use more than 30 variables related to road accidents. At the same time, statistical offices and Ministries of Health seem to use approximately 20 road accident variables.

Furthermore, the **bodies directly involved in accidents** that use most of the variables examined are the Police forces, followed by insurance companies. Hospitals seem to use very few (10) road accident variables regularly. Public research institutes and universities are highly interested in all types of road accident data, whereas in the industry sector, only vehicle and road construction industries can be considered as road accident data stakeholders (in Austria, Greece and Czech Republic).

These results show that **the most important road accident stakeholders** are the Ministry of Transport, Police and Local Authorities, which are identified by all participating countries. Furthermore, statistical offices are considered to be among the most important stakeholders, as their role in processing road accident data seems to be essential in most European countries. Thus, even though the grid results indicate that other stakeholders use more accident data than statistical offices (Table 9), it is nevertheless necessary to include them among the most important stakeholders.

Other important stakeholders are Universities, as well as both private and public Research institutes. In most countries, Universities and Research institutes carry out road accident analysis, using approximately over 30 variables of road accident data, including also the additional variables, which are used in some of the countries (apart from the ones initially included in the Grid). Their needs for accident data, thus, should be carefully considered, as they produce significant results regarding road safety. At the same time, it seems that Ministries of Public Works, Interior and Public Order use more than 30 road accident data variables and should also be considered as significant stakeholders.

Less important stakeholders seem to be professional associations, such as freight and passenger transport association, which regularly use less than 20 variables. Furthermore, the Ministry of Health and hospitals do not seem to use more than 20 variables of road accident data and are not considered as important stakeholders by the participating European countries. Even though it is important for hospitals to collect road accident data in order to estimate the level of underreporting, they do, not carry out accident analysis. Thus, it was predictable that these stakeholders do not emerge as significant data users.

Table 10. Most important road accident data stakeholders according to the extent and frequency of use of road accident variables

National Road Accident Data Users	ROAD USER										VEHICLE			ROAD ENVIRONMENT												ACCIDENT TYPE				ROAD USER CHARACTERISTICS			
	Person class	Age	Sex	Injury severity	Nationality	Driving license age	Alcohol/drug test	Psychophysical circumstances	Car passenger type	Hit and run	Vehicle type	Vehicle age	Security equipment	Area type	Road type	Junction type	Junction control	Carriageway type	Lighting conditions	Weather conditions	Speed limits	Number of lanes	Road surface conditions	Road markings	Region	Accident location (kilometric)	Collision type	Pedestrian movement	Accident type	Vehicle manoeuvre	Data base importance for the user	User Role for Road Safety	Intensity of use of the data base	Potential data source for the data base
1. National Public Administration																																	
Ministry of Transport	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H		H	H	H	H	H	H	H	H
Ministry of Public Works	H	H	H	H	H	H	H				H		H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
Ministry of Public Order	H	H	H	H	H	H	H	H	H		H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H
Ministry of Interior	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H		H	H	H	H	H	H	H	H
Statistical Office	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H		H	H	H	H	H	H	H	H
Local Authorities	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H		H	H	H	H	H	H	H	H
2. Accident Directly Involved Bodies																																	
Police	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H		H	H	H	H	H	H	H	H
4. Research and Scientific Institutions																																	
Public Institutes	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H		H	H	H	H	H	H	H	H
Private Institutes	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H		H	H	H	H	H	H	H	
Universities	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H		H	H	H	H	H	H	H	

Scale of significance:
H high importance
L low importance
- Not used

Transport
Project co-financed by the European Commission, Directorate-General Transport and Energy

The importance of the stakeholders, according to the frequency and the extent of accident data use, was further considered and Table 10 summarises which road accident variables are frequently used by the most important stakeholders, as these were identified in relevant literature and according to the experience of road accident data experts involved in Task 1.4.

As most important "National Public Administration" stakeholders are considered the Ministry of Transport, the Ministry of Public Works, the Ministry of Public Order, the Ministry of Interior, Statistical Offices and Local Authorities.

The most important "Accident Directly Involved Bodies" stakeholder is the Police, whereas public and private research institutes, also Universities.

In order to improve the road accident analysis that is carried out throughout Europe, **the needs for accident data of the most important stakeholders** need to be identified. From Table 10, it seems that the majority of variables examined by the Grid are frequently used by all the important stakeholders. As mentioned in the Grids, the least frequently used variable is the one referring to accident location, which is only used by the Ministries of Public Works and Public Order. This variable is, however, essential when designing remedial measures, for example by improving the design of a junction. Other variables that are not used by all important stakeholders are those referring to hit and run, car passenger type, and psychophysical circumstances of the road user, as well as the one referring to vehicle age. At the same time, the stakeholder using less frequently road accident data is the Ministry of Public Works.

3.3 Analysis of the national accident data collection forms

The analysis presented in the previous section provided a **better insight on data availability and needs** at a European level. As a further step, the national accident data collection forms were examined in order to identify the most commonly used variables and values for accident data collection and analyses throughout Europe. Any useful features identified in national collection systems, including data elements, methodologies, fill-in systems etc. were identified and exploited for the formulation of the proposal for a Common Accident Data Set. Finally, all variables and values that are currently included in the CARE database were also used for the identification of necessary data.

After considering the **structure and the elements of each national collection form**, as well as the results of the analysis of the Grids, four main groups of data elements were further examined:

- Accident related information (General information concerning the accident, the road and the road environment)
- Accident type and manoeuvre information (collision types, manoeuvres, subsequent events etc)
- Vehicle related information
- Person related information

Based on which variables / values are included in the several data collection forms, as well as the results of the analyses of the relevant questionnaire and the Grids, describing the needs of various accident data users for road accident data, each partner concluded to the **most appropriate / necessary variables** and values of the specific group for which he was responsible. These variables would be incorporated into the proposed common accident data set. Moreover, each partner should also consider and describe the necessity of these proposed variables / values and additionally suggest how these should be filled-in (data format). On that purpose, all data collection forms, element definitions and fill-in instructions were sent to the Task 1.4 partners both in electronic and hard copies. Additionally, a document was developed by the Task 1.4 leader describing various structures concerning the filling-in for different types of variables / values (see Appendix IV).

A 3-step methodology was developed for performing the analysis of the national accident data collection sets:

Step 1
Identify which variables/values of the specific group are used in the various countries and which are their definitions. Examine in how many countries each variable/value identified is used.

Step 2
Examine the system for data filling-in for each variable in different countries.

Step 3
Choose the most appropriate variables/values, as well as the respective system for data filling-in for each group of variables.

The variables/values of CARE and the respective definitions were the basis for the analysis of the national collection systems by the four partners, but other international data files were also considered (US- MMUCC, WHO).

The four parts were eventually **compiled together by the Task 1.4 leader in a concise proposal**, which would be a first complete draft. Four Task 1.4 partners were responsible for the four accident information groups. NTUA was responsible for accident information, SWOV for accident type and manoeuvre information, TRL for vehicle related information and KfV for person related information.

Within the framework of the study of the national collection systems, the Task 1.4 leader thoroughly examined all **accident, road and environment variables** used in road accident data collection throughout Europe. The number of common accident related variables between CARE and each country was identified, as well as the compatibility between definitions and structures. Some indicative results from this study are shown in Figure 4.

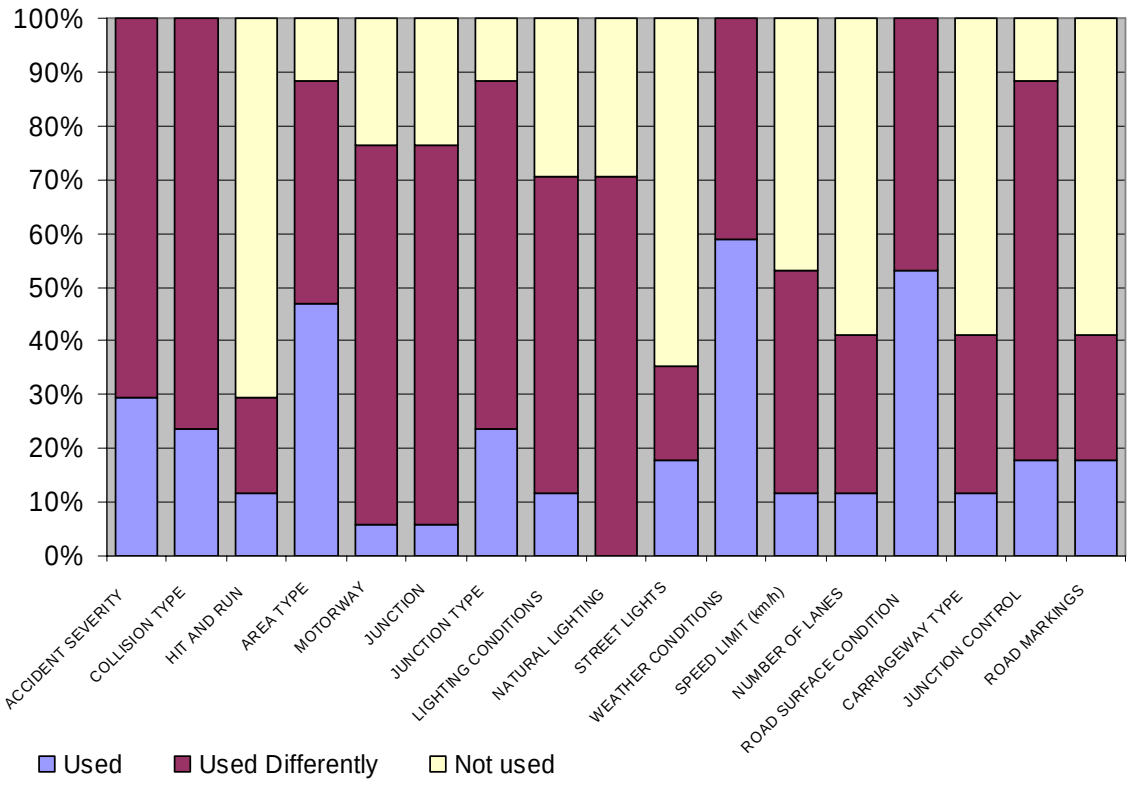


Figure 4.CARE accident related variables used per country

Figure 4 presents **the usage of accident related variables included in the CARE database** by several EU countries. The percentage of usage of each variable (number of countries using the variable divided by the total number of countries) is shown. More specifically, for each variable the percentage of countries using the variable as it is in CARE are highlighted in blue color, the percentage of countries using the variables differently (different structure or definition or both) are shown in red and the percentage of countries not using the variable at all are shown in yellow.

The analysis of accident related data elements was also performed **on a value level**. The set of values of each accident related variable included in CARE was compared to the respective value set of each national variable. The percentage of usage of each value (number of countries using the value divided by the total number of countries) shows the availability of the value in the European accident data collection systems and provides an indication of the importance of the value at European level.

The results of this analysis allowed for a **first indication of data availability and importance** at EU level and together with the input from the Grids presented in the previous section, were exploited in the formulation of the (accident related) first draft of the proposal.

Similar analyses were also performed by the other three Task 1.4 partners, responsible for the groups of variables concerning accident type and manoeuvre, vehicle and person information, who formulated the respective first draft parts and transmitted them to the Task 1.4 leader in order to be included into an integrated recommendation. Their contributions were harmonized (as far as the formatting and presentation of the documents were concerned) and the four documents were compiled into a first draft proposal for a Common Accident Data Set, which was circulated to all Task 1.4 partners.

Primarily, the first complete draft of the **C**ommon **A**ccident **Da**ta **S**et (CADaS) proposal was examined by all Task 1.4 partners and was presented during the WP1 meetings. Many partners provided comments / suggestions which resulted to improved versions of the proposal. Moreover, subsequent versions were also presented at the CARE Experts meetings, during which valuable contributions were received by the national Experts.

Each time **new feedback from National Experts or Task 1.4 partners** was received, the Task 1.4 leader was responsible for the incorporation the of new comments / suggestions into the proposal and for the formulation of a new complete draft version. Since the first version of CADaS was established, several major and minor amendments were proposed by the Task 1.4 partners, the EC CARE administration and the national representatives. Many of these suggestions proved very important for the fine tuning of the proposal, as others led to significant changes on its content and structure.

The initially proposed structure comprised from the **four main groups of variables** (accident, accident type and manoeuvre, vehicle and person), as well as a proposal for a common accident data collection form were gradually enhanced by the feedback received until the final proposal was established. A more flexible structure comprising from four variable groups (accident, road, traffic unit, person related) as well as the record layout approach (for more details see Chapter 4) was adopted.

The CADaS proposal was finalized in September 2008 after the incorporation of all comments received. The rational behind this recommendation, as well as its structure are presented in greater detail in the following Chapter 4. Moreover, the complete list of variables, values and the respective definitions as well as the data formats and the subsequent Annexes of CADaS are presented on a separate document (CADaS Reference Guide).

4. A recommendation for a Common Accident Data Set (CADaS)

4.1 General

The **first version of the Common Accident Data Set (CADaS)** was presented in March 2007. Since then, several new versions of the document were circulated including comments and suggestions from the Task 1.4 partners which contributed to the overall improvement of the proposal. One of the primary goals of Task 1.4 was to provide a recommendation for a pan-European road accident data collection form. However, since the formulation of the first draft of this proposal, it was decided that efforts within Task 1.4 should be concentrated on providing more general guidelines on the data needed at a European level, rather than providing an accident data collection form, allowing for a more flexible usage of the proposed data set and enabling countries to adopt the CADaS without introducing major changes in their current national systems.

Moreover, the first complete draft of the recommendation was based on a **hierarchical approach**. Variables were structured at several levels and the values at the lower levels could only be recorded if the respective values of the upper levels were applicable. However, it was decided that such an approach would be binding for the countries, limiting the potential of the proposal. In order to ensure flexibility in usage, a new approach was adopted by introducing a record layout format. In this way, each level of each variable would constitute a new variable itself, avoiding any hierarchy within variables. This approach maximizes the potential of existing national data to be transformed using the CADaS protocol and allows countries to collect any type of additional road accident data, in any way they wish (meeting also any national / regional / local needs or particularities).

Transport
Project co-financed by the European Commission, Directorate-General Transport and Energy

4.2. Establishment of a basic common accident data collection set and methodology

4.2.1. Background

Since long, road accident data are collected in the European Union (EU) countries by the use of **their own national collection systems**. At European level, road accident data are also available since 1991 in disaggregate level in CARE, the Community database on road accidents resulting in death or injury. CARE comprises detailed data on individual accidents as collected by the Member States, using a structure which allows for maximum flexibility and potential with regard to analysing the information contained in the system.

The purpose of CARE system is to provide a powerful tool which would make it possible to identify and quantify road safety problems throughout the European roads, evaluate the efficiency of road safety measures, determine the relevance of Community actions and facilitate the exchange of experience in this field. Parts of the national data sets are integrated into the CARE database in their original national structure and definitions, however, as existing national accident data collection systems are not always compatible and comparable among the countries, the Commission provides and applies a framework of transformation rules to the national data sets, allowing CARE to have compatible data.

CARE database currently contains 55 common road accident variables. However, it is acknowledged that more variables and values are necessary to better describe and analyse the road accident phenomenon at EU level. Due to differences in the collected data variables and values, their definitions, the differences of the accident data collection forms structures and the relevant data formats among the existing national databases, both accident data quality and availability are affected. Consequently, lack of accident data uniformity among and within EU countries hinders the exploitation of CARE potential and limits data analyses and comparisons at EU level.

Under this perspective, the recommendation for a **C**ommon **A**ccident **Da**ta **S**et (**CADaS**) has been developed consisting of a minimum set of standardised data elements, which will allow for comparable road accident data to be available in Europe. In this way, more variables and values with a common definition will be added to those contained in the CARE database, maximising thus the potential of CARE database and allowing for more detailed and reliable analyses at European level.

This recommendation for the CADaS was based on the identification of the data required for accident analysis through the **input of Experts from the CARE and SafetyNet**, but also on the analysis of the currently available national accident data collection systems in Europe. The variables and values included in the CARE database were used as a basis of CADaS but other international data files were also considered (US - MMUCC, WHO).

The current version of CADaS was finalised after 54 months of elaboration within the SafetyNet integrated project in close cooperation with the CARE team of the European Commission and after 30 months of elaboration within the CARE Governmental Experts Group. Both the structure and the contents of CADaS were **improved after several iterations** taking into account the suggestions (sometimes conflicting) of the SafetyNet Experts, the CARE Experts and the EC CARE team. It is obvious that not all suggestions of the CARE Experts could be adopted; however the current version of CADaS received the maximum possible acceptance by the CARE Experts Group, as a result of its voluntary "a la carte" way of implementation. It is noted that some National Administrations continue to not agree with some of the structure and contents' choices and even with some fundamental design principles, as is the case of the Danish Road Directorate.

4.2.2. Scope and purpose

The recommendation for a Common Accident Data Set (CADaS) consists of a minimum set of standardised data elements, which will allow for comparable road accident data to be available in Europe. **The CADaS can be implemented on a voluntary basis** at the national accident collection systems and be gradually adopted by the EU countries. Thus, progressively, more and more common road accident data from the various countries will be available in a uniform format. In this way CARE, the European data base with disaggregate data on road accidents, will gradually contain more and more compatible and comparable data, allowing for more reliable analyses and comparisons across the EU countries.

The recommendation for a Common Accident Data Set (CADaS) refers to **the set of data to be voluntarily transmitted by each country to the EU**, which should be derived from the national road accident data collection system. This means, that the EU countries will not be legally obliged to adopt the CADaS and can continue using their national systems, however, if they wish they can enhance them in order to be able to provide the CADaS data to the EU. In case the countries do not wish to adopt the CADaS they should continue transmitting national road accident data to the EU in the current format. Moreover, if Member States wish to adopt the CADaS, some variables might need to be collected under a different structure, in order to meet local/regional/national needs or particularities. Countries can continue using the particular variables and values for collecting national data and use appropriate tranformations when these data are transmitted to the EC in the CADaS format.

The CADaS is structured in a simple way, without levels of hierarchy, constituting in fact the **record layout of the data set to be transferred to the EU**. The structure of the CADaS variables allows for various levels of detail to be selected for providing the requested data, by the use of alternative (aggregate) values. In this way, the data to be transferred can be more easily produced at national level. However, the variables and values of CADaS may also be considered as recommendations for national police road accident data collection reports. Moreover, CADaS variables and values can be further

enhanced (derived variables to be added) inside the CARE database allowing for a wide range of analysis reports.

4.2.3. Selection criteria for CADaS data

The selection of the data variables and values incorporated in **CADaS** resulted from the balanced co-consideration of the following basic criteria:

- **Variables and values must be useful for road accident analysis**, especially at EU level.
- **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, which is of local interest.
- **Each country should have the possibility to choose alternative level of detail** of the various variables and values.
- **Variables and values must be comprehensive and concise**. Each variable must include description and scope (importance to road safety) attribute values and their definitions and the data format.
- **Data which are impossible or very difficult to be collected are not retained in the CADaS**, independently of their value for road accident analysis; as such data might be of low quality.
- **The future perspective of using certain variables and values** was taken into account, even though those data are not currently collected by most of the countries due to current technical difficulties (i.e. latitude and longitude of the accident location, etc.).
- **Existing variables and values of CAREPLUS are of first priority within CADaS.**
- **CADaS variables and values refer to casualty road accidents**, 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.

4.2.4. CADaS structure

The CADaS variables presented in the Reference Guide are divided into **four basic categories**. The category in which each variable is included can be identified by a unique letter (code) at the beginning of the name of the respective variable. The categories and the relevant codes used to describe each category are the following:

A, for **Accident** related variables,
R, for **Road** related variables,
U, for **Traffic Unit** related variables,
P, for **Person** related variables.

In the following Figure 5 the interrelation among the four basic categories is presented, clearly indicating the links of the various road accident variables included in CADaS.

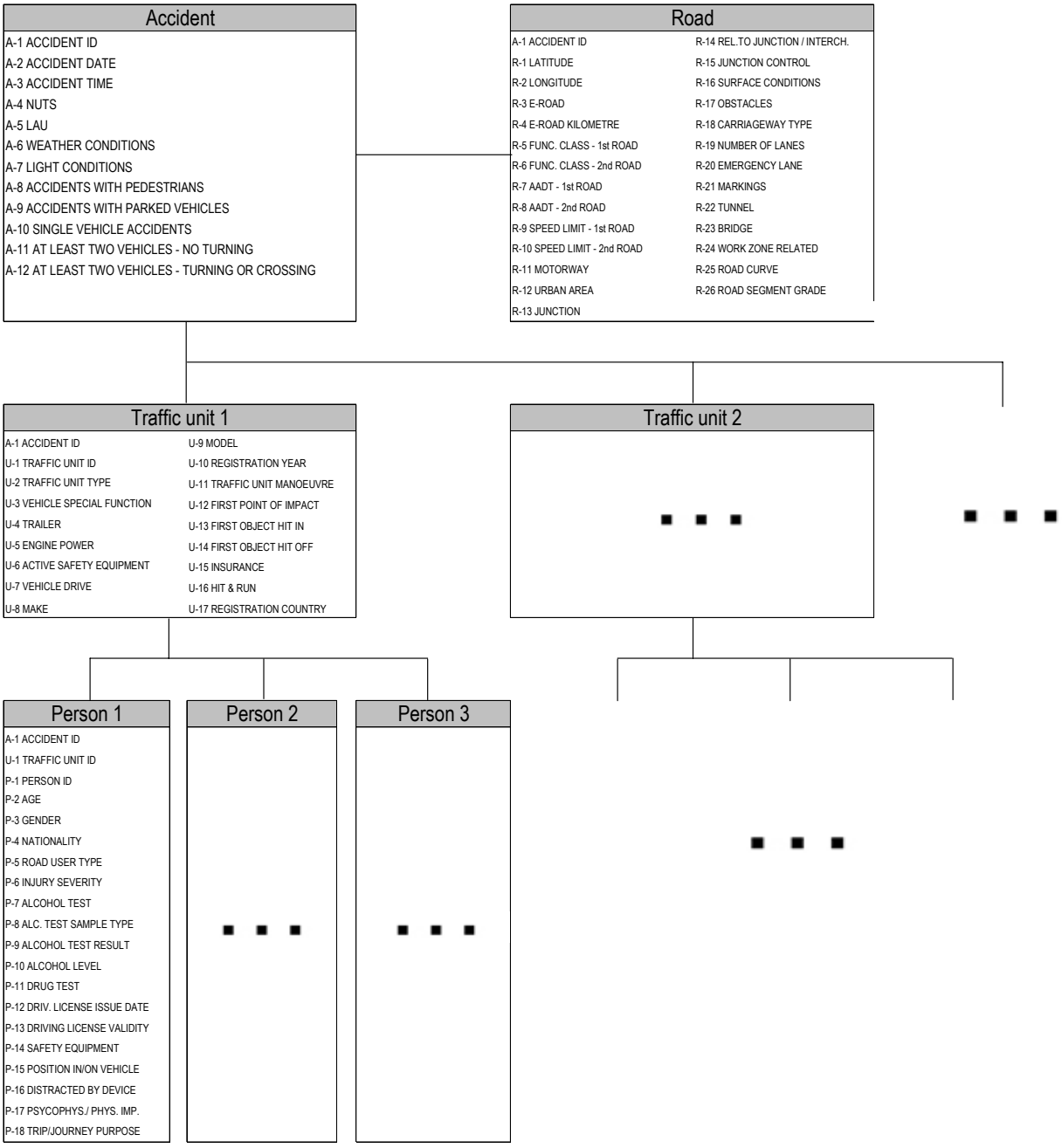


Figure 5 CADaS structure

Several variables include two distinct types of values, referring to different level of detail:

1. **Detailed values:** concern information at the highest level of detail.
2. **Alternative values:** concern information at a more aggregate level of detail, when more detailed values are not available.

Alternative values do not differ from detailed values apart from their level of detail. These values are complementary and can be used when more detailed data are not available (for example concerning the “Traffic Unit type” variable, if a country does not collect the values “car” and “taxi” separately, it can provide this information through the “car or taxi” alternative value). Both detailed and alternative values have the same number of digits and occupy the same columns on the record layout. For alternative values, the **(A)** coding identifier is used next to the category code (e.g. AA, RA, VA, PA).

Due to the fact that the recommendation of CADaS is designed to be adopted gradually and on a voluntary basis by the EU countries, over the coming years, the recommended variables were separated into two broad categories, according to their importance for road accident analysis: **variables of high importance (H) and variables of lower importance (L)**. Apart from their importance for road safety analysis, CADaS variables are separated according to the current reliability the collected data and the related collection feasibility.

It should be clear that **all EU countries should continue using their national systems** and collect accident data in any way they find most appropriate and if they wish, they can make the necessary adjustments allowing to provide the CADaS data to the EU. These data can be directly collected, or derived from collected data, or obtained through linkage to other national data bases (driving licenses, social security, road network, vehicle registry, etc.). They can certainly opt for the level of detail of CADaS data to collect and transmit, according to the various proposed CADaS alternative values.

The structure of CADaS allows for **maximum flexibility**. The proposed record layout allows for the data to be provided in a simple way without levels of hierarchy, irrespectively how these data are collected in each national system. Flexibility is also ensured by the use of the alternative values, which allow for the provision of information at different levels of detail, according to the existing national data collection system.

The proposed **value coding** is indicative and refers to the format, in which the CADaS data should be transmitted to the EU by the countries that wish to fully or partially adopt them. With reference to the proposed measurement units of various values (kms/miles, etc.), the most commonly used measurement units by the EU countries are retained. However, necessary adjustments of these measurement units can be adopted in case some countries use different metric systems (i.e. miles per hour instead of kilometres per hour for speed measurement).

For each of the variables included in the **CADaS**, the following information is presented:

Variable Label: The label of the proposed variable, consisting from the category identifier (A, R, U or P), 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 and (L) for variables of lower importance.

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).

4.2.5. Summary of variables and values

The number of variable and values contained in the CADaS are presented at the following Table 11:

Table 11. CADaS variables and values in numbers

Variable category	Code	Number of Variables			Number of Values		
		High (H) importance	Lower (L) importance	Total	Detailed values	Alternative values (A)	Total
Accident	A	7	5	12	86	13	98
Road	R	11	15	26	106	13	119
Traffic Unit	U	7	10	17	137	15	152
Person	P	11	7	18	91	10	102
Total		36	37	73	420	51	471

5. Adoption of the common data collection set and methodology

The establishment of the Common Accident Data Set provides the opportunity to dispose more comparable among the EU countries data, derived from national road accident data collection systems. European countries can adopt the CADaS whenever they wish and can continue using their national data collection systems. If one country decides to start using the CADaS protocol it can transform its national data into the CADaS data by using appropriate transformation rules and eventually transmit the transformed data to the EC.

The **adoption of the CADaS does not presuppose any changes** in a country's national data collection system. Certainly, if one country wishes to enhance its national data collection system or change it according to the CADaS protocol, it can adopt the CADaS proposal in "pieces" (à la carte) or as a stand alone road accident data collection system.

CADaS can be adopted gradually by EU countries, but any part of it (variables, values, definitions and data formats) can be implemented within an existing national system, increasing thus the compatibility of national data with the respective CARE data. Consequently, **the level of adoption of the CADaS may vary according to any national needs and/or particularities** and can be performed during any time in the future.

In Figure 6, the **current, intermediate and future** (based on the CADaS adoption) **processes** of the national road accident data files are presented. Using both (current and future) approaches ensures compatibility of the accident data among EU countries and the main difference of these two approaches is related to the degree of involvement of the country in the process.

Taking into account that many Member States may partially adopt CADaS, an intermediate phase is also necessary. During this phase, countries may use a part of the CADaS in order to transform specific variables and values at national level and transmit the rest of the data in the current format in order to be transformed using the CAREPLUS protocol. According to the proposed future process, transformation of the national accident data (based on the CADaS protocol) will be performed at the national level and the derived CADaS variables and values will be transmitted to the EC, where they will be included in a more automatic way into the CARE database. This process will allow for more common variables and values but also for higher quality, given that the **national authorities better perceive any particularities related to national data collection**, and subsequently can better identify the interrelation between the collected and the CADaS variables.

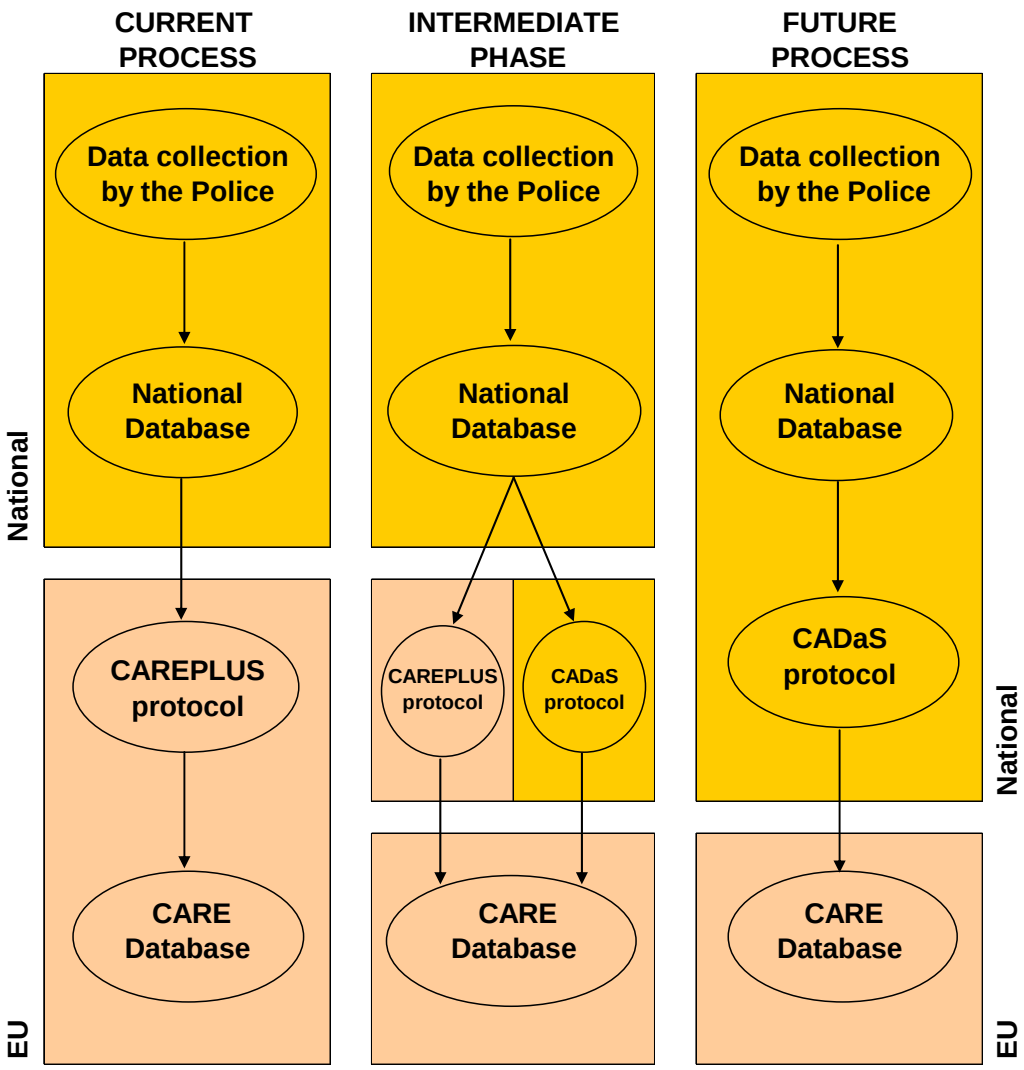


Figure 6. Accident data transformation processes

Many European countries have been using their national accident data collection systems since long and have gained significant experience in the field. These national systems provide complete and concise data, which are extremely useful for research and policy making purposes. However, other countries have only recently introduced a road accident data collection system or consider revisions and improvements to their existing national system. The implementation of **CADaS in such cases can provide an invaluable tool for benchmarking road accident data collection** and subsequently improve the efficiency of a national data collection system.

As a further step, **a pilot phase for the implementation of the CADaS** could be considered. Such a pilot can enable the identification of possible weaknesses in the recommendation that could only be tracked through an actual application of the system; as a result, it would significantly strengthen the content of the Common Accident Data Set. This pilot phase could be

implemented in countries wishing to revise their national systems or countries with less experience in road safety wishing to exploit the experience of other countries through the CADaS recommendation.

Finally, the **adoption of CADaS should be supported by many different parties** in order to maximise its acceptance by the EU countries. National representatives at various levels (CARE Experts Group, High Level Group) could contribute by promoting the recommendation in their own countries. Additionally, a broader dissemination could be achieved through the ERSO website, but also through its presentation in scientific papers and national and international conferences. Finally, the assistance of the European Commission is important for promoting the CADaS and encouraging countries to implement pilot programmes for its implementation.