CEDR Transnational Road Research Programme
Call 2012: Safety

Funded by Belgium/ Flanders, Germany, Ireland, Norway, Sweden, United Kingdom

BRoWSER: Base-lining Road Works Safety on European Roads

Database specification and design

Visualisation extension

	**Transport Research Laboratory** www.trl.co.uk
	**Karlsruhe Institute of Technology** www.ise.kit.edu
	**Belgian Road Research Centre** www.brrc.be
	**Trinity College Dublin** www.tcd.ie
	**Slovenian National Building and Civil Engineering Institute** www.zag.si

February 2015.

CEDR Call2012: Safety
BRoWSER: Base-lining Road Works Safety on European Roads

Database specification and design
Visualisation extension

Submission date: February 2015

Start date of project: 01/02/2013 End date of project: 31/10/2015

Authors of this deliverable:
Brian Lawton, TRL, United Kingdom
Jill Weekley, TRL, United Kingdom
Xavier Cocu, BRRC, Belgium
Alan O'Connor, TCD, Republic of Ireland
Mojca Ravnikar Turk, ZAG, Slovenia
Matthias Zimmermann, KIT, Germany
Dr Iain Rillie, TRL, United Kingdom (Quality Review)

Version:1.0

Contents

1 Introduction

1.1 The BRoWSER project

The project Base-lining Road Works Safety on European Roads (BRoWSER) was initiated as a response to the Description of Research Need (DoRN) for the CEDR Transnational Road Research Programme Call 2012 on Safety.

The aim of the CEDR Transnational Research Programme (2012 call) seeks "to significantly reduce risks to road workers with an objective of Zero Harm". BRoWSER addresses two of the topics within the 2012 Call under the heading of "Safety of road workers and interaction with road users". These are:

- Collect data on worker injuries and near misses by country, road administration and employer
- Understand the optimum road works layouts that enable road users to approach, travel through and exit works without causing injury to workers and others

The aim of the BRoWSER project is to help National Road Authorities (NRAs) enable a data-led approach to be taken to managing road worker safety. This knowledge of how road workers are exposed to risk from accidents and road user error is essential for effective safety management as it allows the real risks to be managed rather than those perceived to be the problem. The BRoWSER project focuses on the interaction between road workers and traffic and will allow consideration of road worker accidents, incidents and near misses (where available) alongside data for road works practices, network characteristics and road user accident data at road works.

1.2 Background

The benefits of a European Road Worker Casualty Database (EuRoWCas) were identified in the project deliverable D1.1 (BRoWSER Benefits Case). In order to realise these benefits, and by extension to achieve the overall objectives of the research programme, road worker incident data must be collected in a clearly defined and standardised format.

Initial requirements for the data fields were specified in project deliverable D2.1 (Input data definition document for EuRoWCas). The collection of these data is not possible with (or without some adaptation to) the existing data collecting processes in the individual countries. Therefore it was agreed that a three-month data collection trial would be carried out to demonstrate the feasibility of data collection. The results of the trial were reported in the project deliverable 'Final Trial Report' and this showed that such data collection is practicable and that there is an appetite for collecting data. An additional outcome of the trial was the opportunity to carry out end-user testing on the data format specification, identify modifications to the data fields and associated options and perform a high-level gap analysis between the complete dataset and the data that could be collected in practice in each of the trial countries.

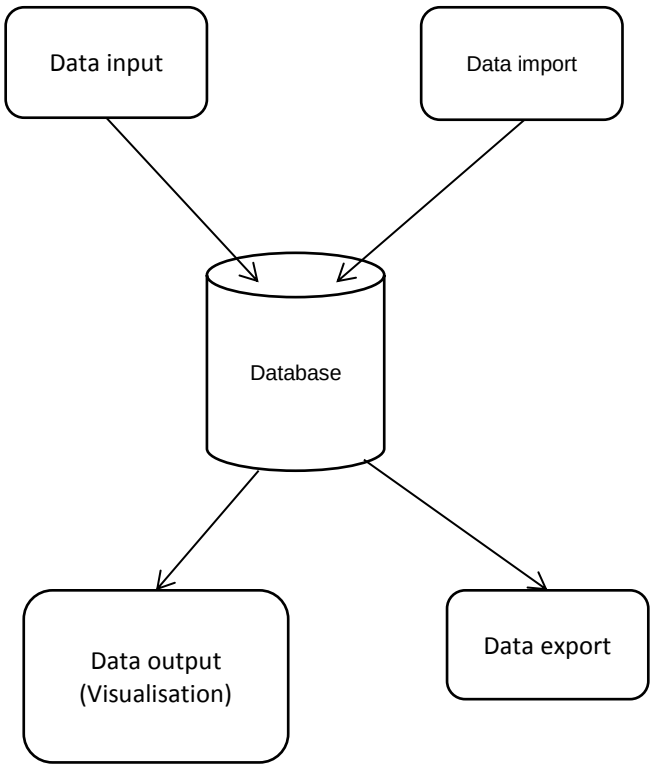
1.3 This document

This document presents an updated technical specification for the data fields and associated values, resulting from the findings and experiences of the data collection trial. This covers the format of the data fields, how they should be coded, relationships between the fields and associated options for each field. This document also provides an overall concept for the EuRoWCas database system, and discusses specifications for the format of data import, data export, data input interface and data output in support of data visualisation. Implementation of a database is also discussed, along with consideration of the implications of different levels of data collection and availability.

This extended version of the report (first issued in December 2014) includes further consideration of the data visualisation output and provides some options for potential future implementation.

2 System concept and structure

The proposed structure for the EuRoWCas database system is illustrated in the diagram below.



The overall concept is such that, if this technical specification is followed, the software or operating system used to host a EuRoWCas database does not matter, as the data imported into, stored within and exported from the database will be to a common standard.

The implication of this common standard approach is that there can be databases at different levels within the EU and Member States. For example:

- The EU-level database can take imports from country-level or region-level databases.
- Country-level databases can take data from region-level or contractor-level databases (or a combination of both).
- Region-level or contractor-level databases can take data from individual databases hosted by contractors, sub-contractors, enforcement authorities etc.

This will be possible due to the harmonisation of the data format via a publically available specification for EuRoWCas rather than via a specific product.

The system structure comprises five elements:

Database: The central data storage structure, which forms the main focus of this specification.

Data input: How data are input into the database directly by contractors or equivalent.

Data import: How data are imported from other EuRoWCas data sources, for example a sub-contractor's database. The standard data format would make data import trivial.

Data output / visualisation: How data can be extracted directly from the database and viewed / analysed through a suitable user interface. (Producing options and a demonstration of this is the focus of Work Package 5).

Data export: How data can be downloaded from the database in a raw format for import into another EuRoWCas database or into an analysis package. The data format is the same as that for the data import – this allows different levels of the database structure to interact.

3 Data specification

3.1 General

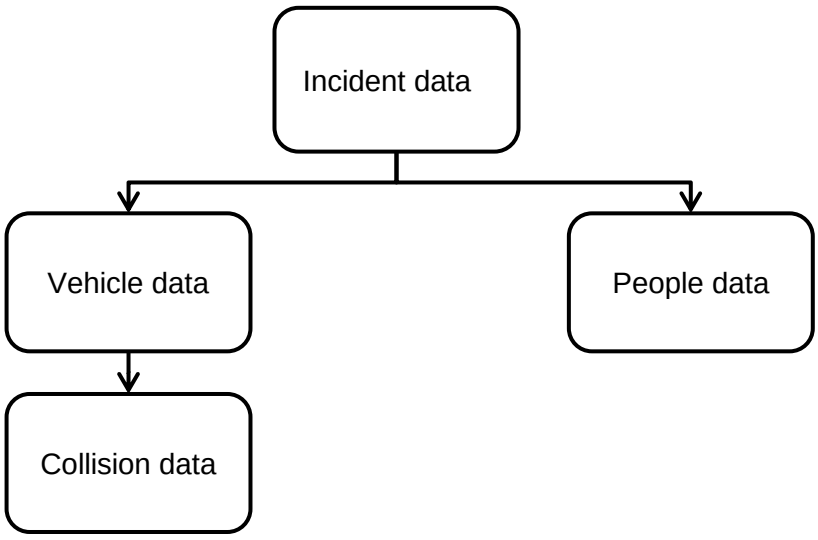
For each incident recorded in the database, the user must input data on three aspects: the incident circumstances, the vehicles involved and the people involved.

The incident circumstances data are divided in several elements:

Base data – basic information about the incident, such as time, date and location and the involvement of people and vehicles.
Site data – information on the carriageway and lane configuration at the time of the incident.
Road works data – information on the road works and associated traffic management in place at the time of the incident.
Environmental data – information on environmental factors such as weather, visibility and lighting conditions.

Part of the base data input by the user is the number of vehicles and people involved in the incident. For each vehicle and person involved, an individual vehicle or person record is required, linked to the incident record. These records should be populated automatically from the information provided in the base data input, with the user then adding additional detail to each vehicle or person record.

For each person recorded as involved in the incident, the user is asked to provide further details such as person type, whether they were involved in a collision (as opposed to a near miss), and the level of any injury that occurred. For each vehicle involved in the incident, further details are also requested. In order to allow for a vehicle to be involved in multiple collisions within the same incident (for example, a vehicle collides with another vehicle and subsequently with a restraint barrier), the user inputs the number of collisions for each vehicle. For each collision, further details are then also collected. If vehicles and/or people are known to be involved in the incident, but no further details are known about them, records are still created in the database. This produces the following hierarchy for the data:



The goal would be to collect all fields in the database for each incident. However, in practice and particularly during initial implementation of the system, this is unlikely to be possible in all countries. There are four critical fields – country, date, time and location – which must be completed in order to create an incident record. If one of the critical fields is missing, the incident cannot be recorded. These fields combine to form the unique incident identifier, or ID.

All other fields are strongly recommended but optional, which allows road authorities to carry out a staged implementation if desired. More discussion of this concept and the impact of different levels of data collection can be found in Section 4.

Regardless of which fields are collected in an individual implementation of the system, all fields must be included in in the data output format in order for the standardised format to be maintained. This means that for fields that are not collected in an individual implementation, a null value must be entered automatically by the system.

3.2 Data field specification structure

For each data field in the EuRoWCas database, a number of key parameters must be defined. These are:

- *Data field ID*: Identifier for each data field
- *Data field description**: Description or associated question for the data field..
- *User interface format*: Format through which the user should input the data. In most cases this is a drop-down list of options from which the user must select.
- *Internal database format*: This is the format in which the data should be stored within the database. In most cases the data should be encoded as integers corresponding to the text options presented to the user. This allows different implementations of the system to present the options in different languages whilst keeping the raw data standardised.
- *Precedents*: Associated data field(s) that define whether the particular field is enabled (relevant) or disabled (not relevant). Note that whether 'disabled' fields are visible or not is an individual choice for each implementation.
- *Dependents*: Associated data field(s) that depend on the option selected by the user for this particular field.
- *Options**: For fields presented in the format of drop-down lists, this provides the associated options that the user can select.
- *Null value*: The default value that should be taken by the field if no data are entered (or if the field is not presented to the user as discussed in Section 3.1). In most cases, the null value is 'unknown' – note that this is the 'Unknown' option in the drop-down list and therefore the value to be recorded in the database is the integer value associated with the 'Unknown' option.

Fields shown with an asterisk (*) will include text that will need to be presented in the native language of the user and so would be translated into different languages in different implementations.

3.3 Data fields

3.3.1 Base data

Data field ID	B1
Data field description	**Incident ID**
User interface format	Automated
Internal database format	Text string
Precedents	ID is automatically created based on country, date/time and location
Dependents	Forms part of V1 and P1 for each vehicle and person record
Options	N/A
Null value	N/A

This is a unique compound text string created automatically from country / date / time / location, e.g. UK_YYYYMMDDHHMM_Location. The incident ID should not be numbered due to potential duplication when aggregating datasets. If any of these fields are missing, a record cannot be created for this incident.

Data field ID	B2
Data field description	**Country**
User interface format	Drop-down list
Internal database format	Text string (two letters)
Precedents	None
Dependents	Forms part of B1. Critical field.
Options	BE; BG; CZ; DK; DE; EE; IE; EL; ES; FR; HR; IT:CY: LV; LT; LU; HU; MT; NL; AT; PL; PT; RO; SI; SK; FI; SE; UK
Null value	N/A. Critical field.

These are the EU country codes for the 28 EU Member States.

Data field ID	B3
Data field description	**Road authority**
User interface format	Drop-down list
Internal database format	Text string
Precedents	List populated automatically depending on B2
Dependents	None
Options	Options will need to be defined and encoded on a national level
Null value	Unknown

This should note the road authority responsible for managing the road on which the incident took place.

Data field ID	B4
Data field description	**Date**
User interface format	Calendar (clickable)
Internal database format	Date YYYY/MM/DD
Precedents	None
Dependents	Forms part of B1. Critical field.
Options	N/A.
Null value	N/A. Critical field

Data field ID	B5
Data field description	**Time**
User interface format	Time (selectable) – drop-down hours and minutes
Internal database format	Time HH:MM (24hr)
Precedents	None
Dependents	Forms part of B1. Critical field.
Options	N/A.
Null value	N/A. Critical field.

If exact time is not known, the user should enter an approximate time.

Data field ID	B6
Data field description	**Incident description**
User interface format	Free text
Internal database format	Text
Precedents	None
Dependents	None
Options	N/A.
Null value	Unknown

The incident description is a brief description of the circumstances and will most likely be in the national language. As such, in a pan-European database it will be included just for reference.

Data field ID	B7
Data field description	**Location and direction identifier**
User interface format	Country-specific
Internal database format	Country-specific
Precedents	None
Dependents	Forms part of B1. Critical field.
Options	N/A.
Null value	N/A. Critical field.

The location information will need to be country-specific and may be road number and marker post, coordinates, latitude and longitude etc. The location information should include indication of the carriageway direction. Each country will need to define their own format, which uniquely identifies the location concerned. For pan-European mapping purposes the system will need to convert this into standard geographical coordinates.

Data field ID	B8
Data field description	**How many vehicles were involved in total?**
User interface format	Integer input
Internal database format	Integer (inc. zero)
Precedents	None
Dependents	Creates vehicle records for each vehicle and pre-fills vehicle V1 in each.
Options	Integers (including zero); unknown
Null value	Unknown

Users should enter the number of vehicles known to be involved in the incident (including road user vehicles and road works vehicles).

Data field ID	B9
Data field description	**How many road users were involved in total?**
User interface format	Integer input

Internal database format	Integer (exc. zero)
Precedents	None.
Dependents	Creates person records for each road user and pre-fills P1 and P2=road user in each.
Options	Integers (exc. zero); unknown
Null value	Unknown

Road users should be included regardless of whether they sustained an injury. This field cannot be zero since a road user must be involved in order for the incident to be of relevance to the EuRoWCas database.

Data field ID	B10
Data field description	**How many road workers were injured?**
User interface format	Integer input
Internal database format	Integer (inc zero)
Precedents	None.
Dependents	Creates person records for each road worker and pre-fills P1 and P2=road worker, P6=yes, P8=yes in each.
Options	Integers (exc. zero); unknown
Null value	Unknown

Only injured road workers are recorded individually. The presence of non-injured road workers is recorded in field R7 as either 'road workers present' or 'road workers not present'; this is because (in the absence of road worker injury) the focus is on the potential risk to road workers, which exists independently of the number of workers present.

3.3.2 Site data

Data field ID	S1
Data field description	**Carriageway type**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	Motorway; other dual carriageway; primary single carriageway; other single carriageway; unknown
Null value	Unknown

Data field ID	S2
Data field description	**Is there a hard shoulder?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If yes, enable S3. If no, disable S3
Options	Yes; no; unknown
Null value	Unknown

Hard shoulder in this context is a paved strip beside a motorway for vehicles stopping in emergencies to be off the main carriageway. If the paved strip is not sufficiently wide to allow a traffic management vehicle to pull out of the live traffic lane then the user should select 'No'

Data field ID	S3
Data field description	**Was the hard shoulder open to traffic?**

User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled / disabled by S2
Dependents	None
Options	Hard shoulder open for emergency use only; hard shoulder open as a running lane; hard shoulder closed; unknown
Null value	Unknown. If disabled, 'not relevant'

Data field ID	S4
Data field description	**Standard number of lanes on carriageway**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enable the correct number of lanes in S5
Dependents	For S4=n, S5 should be enabled n times. If unknown, disable S5.
Options	Unknown;1;2;3;4;5;6
Null value	Unknown.

This is the number of lanes on the carriageway (excluding any hard shoulder) in the absence of works. Note that the definition of a carriageway is "a width of road on which a vehicle is not restricted by any physical barriers or separation to move laterally". This means that for dual carriageways (where there is a barrier or separation) S4 should be the number of lanes per direction; for single carriageways, S4 should be the total number of lanes on the road. .

Data field ID	S5
Data field description	**Lane configuration**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Field created for each of n lanes in S4.
Dependents	None
Options	Unknown; Open to traffic; Open as contraflow; Closed
Null value	Unknown. If disabled, 'not relevant'.

Data field ID	S6
Data field description	**Junction detail**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	Unknown; not at or within 20 metres of a junction; at/on a slip road; at/on a roundabout; T junction; crossroads; other junction
Null value	Unknown

Data field ID	S7
Data field description	**Permanent speed limit**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None

Options	130kph; 120kph; 110kph (70mph); 100kph (60mph); 90kph; 80kph (50mph); 70 kph; 60kph (40mph); 50kph (30mph); 40kph; 30kph; unknown
Null value	Unknown

For individual countries the options could be customised to show only kph or mph. In this case, the underlying encoding of the options must remain unchanged regardless of which options are displayed.

3.3.3 Road works data

Data field ID	R1
Data field description	**Was a temporary speed limit in place?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If yes, enable R2 and R3. If no or unknown, disable R2 and R3
Options	Yes; No; Unknown
Null value	Unknown

Data field ID	R2
Data field description	**Was this temporary speed limit advisory or mandatory?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by R1 = yes.
Dependents	None
Options	Advisory; Mandatory; Unknown
Null value	Unknown. If disabled, 'not relevant'

Data field ID	R3
Data field description	**Temporary speed limit**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by R1=yes
Dependents	None
Options	130kph; 120kph; 110kph (70mph); 100kph (60mph); 90kph; 80kph (50mph); 70kph; 60kph (40mph); 50kph (30mph); 40kph; 30kph; unknown
Null value	Unknown. If disabled, 'not relevant'.

For individual countries the options could be customised to show only kph or mph. In this case, the underlying encoding of the options must remain unchanged regardless of which options are displayed. Since different speed limits often apply across a work zone, the user should select the speed limit where the incident occurred.

Data field ID	R4
Data field description	**Road works classification type**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	Static - short daytime (i.e. up to 8 hours); static - short nighttime (i.e. up to 8 hours); static – medium (more than 8

	hours and up to 24 hours; static – long (more than 24 hours); mobile works; emergency / incident support; unknown
Null value	Unknown

Data field ID	R5
Data field description	**Country-specific road works description**
User interface format	Free text
Internal database format	Text string
Precedents	None
Dependents	None
Options	N/A
Null value	Unknown

Free text field so that users can reference standard layouts or similar (within country this could potentially pre-fill fields)

Data field ID	R6
Data field description	**Road works activity period**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	During installation of works; durings works period; during removal of works; unknown
Null value	Unknown

Data field ID	R7
Data field description	**Were there road workers present at the time of the incident?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	Yes; no; unknown
Null value	Unknown

Data field ID	R8
Data field description	**Was advance signing present?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If yes, enable R9. If no or unknown, disable R9
Options	Yes; no; unknown
Null value	Unknown

Data field ID	R9
Data field description	**Advance signing type**
User interface format	Drop-down list
Internal database format	Integer

Precedents	Enabled by R8 = yes
Dependents	None
Options	Hard signs only (static); hard signs only (mobile); electronic signs only (static); electronic signs only (mobile); combination of both hard and electronic (static); combination of both hard and electronic (mobile); unknown
Null value	Unknown. If disabled, 'not relevant'.

Data field ID	R10
Data field description	**Traffic control measures**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	None; automatic traffic signals; manual traffic control; unknown
Null value	Unknown

Data field ID	R11
Data field description	**Approx length of road works (metres)**
User interface format	Integer input or unknown
Internal database format	Integer
Precedents	None
Dependents	None
Options	N/A
Null value	Unknown

This should be the distance from the first physical traffic management intervention to the end of the works end zone.

Data field ID	R12
Data field description	**Did the road works involve a contraflow?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If yes, enable R13. If no or unknown, disable R13.
Options	Yes;no;unknown
Null value	Unknown

Data field ID	R13
Data field description	**Contraflow type**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by R12 = yes
Dependents	None
Options	Full; tidal; unknown
Null value	Unknown. If disabled, 'not relevant'

Data field ID	R14
Data field description	**Was the incident on the carriageway or off the**

	carriageway?
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If off, enable R15 and disable R16 and R17. If on, enable R16 and disable R15. If unknown, disable R15 and R16
Options	On carriageway; off carriageway; unknown
Null value	Unknown

Data field ID	R15
Data field description	**Was the incident on the verge or in the central reserve?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by R14 = off carriageway
Dependents	None
Options	Verge; central reserve; unknown
Null value	Unknown. If disabled, 'not relevant'

Data field ID	R16
Data field description	**Was the incident mainly on the hard shoulder, on a live carriageway or within the closure?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by R14 = on carriageway or both.
Dependents	If 'within closure', enable R17. If any other option, disable R17.
Options	On hard shoulder; on live carriageway (i.e. on lane open to traffic); within closure; unknown
Null value	Unknown. If disabled, 'not relevant'

If the incident occurred within a hard shoulder closure, the user should select 'within closure'.

Data field ID	R17
Data field description	**What delineation was present?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by R16 = within closure
Dependents	None
Options	Incident occurred behind cones; incident occurred behind barrier; incident occurred behind panels; unknown
Null value	Unknown. If disabled, 'not relevant'.

Data field ID	R18
Data field description	**Incident location zone**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	Advanced warning zone; transition area (entrance); transition area (exit); works zone; unknown
Null value	Unknown

3.3.4 Environmental data

Data field ID	E1
Data field description	**Time of day**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If daylight, disable E2. If darkness or dawn / dusk, enable E2.
Options	Daylight; darkness; dawn / dusk
Null value	Unknown

Data field ID	E2
Data field description	**Lighting in use (works lighting or street lighting)**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by E1 = darkness or dawn / dusk.
Dependents	If lighting in use, enable E3. If no lighting in use or unknown, disable E3.
Options	Lighting in use; no lighting in use; unknown
Null value	Unknown. If disabled, ‘not relevant’.

Data field ID	E3
Data field description	**Type of lighting**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by E2 = lighting in use
Dependents	None
Options	Works lights; street lights only; unknown
Null value	Unknown. If disabled, ‘not relevant’.

Data field ID	E4
Data field description	**Weather conditions**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	Fine; rain / snow; fog / mist; unknown
Null value	Unknown

Data field ID	E5
Data field description	**Visibility conditions**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If poor, enable E6. If good or unknown, disable E6.
Options	Good; poor; unknown

Null value	Unknown

The description (or associated help message) should include examples of visibility contraints to assist the user in understanding conditions that may be described as poor.

Data field ID	E6
Data field description	**Visibility constraint**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by E5 = poor
Dependents	None
Options	Weather; spray from vehicles; low sun; road layout (e.g. bend, hill crest); object / vegetation blocking view; other; unknown.
Null value	Unknown. If disabled, 'not relevant'

If more than one constraint is relevant, the user should select the option felt to have been the most significant factor.

Data field ID	E7
Data field description	**Road surface condition**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	Dry; wet/damp; snow; frost/ice; flood (surface water over 3cm deep); unknown
Null value	Unknown

3.3.5 Vehicles data

Vehicles records are created from B8, one record for each vehicle. V1 is created and pre-filled automatically for each.

Data field ID	V1
Data field description	**Vehicle ID**
User interface format	Automated
Internal database format	Text string
Precedents	Created from B1 in the format 'Incident ID_n'
Dependents	None
Options	N/A
Null value	N/A

If no further details are known, the record for the vehicle is still created.

Data field ID	V2
Data field description	**Was this a road user vehicle or a road works-related vehicle?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If road worker vehicle, enable V3 and disable V4. If road user vehicle, enable V4 and disable V3. If unknown, disable V3 and V4.

Options	Road works related vehicle; road user vehicle; unknown
Null value	Unknown

Data field ID	V3
Data field description	**Was the vehicle equipped with a crash cushion?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by V2 = road works related vehicle
Dependents	None
Options	Yes; no; unknown
Null value	Unknown. If disabled, 'not relevant'

Data field ID	V4
Data field description	**Road user vehicle type**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by V2 = road user vehicle
Dependents	None
Options	PTW; bus/coach; car (inc minibus up to 9 people); light goods (<=3.5 tonnes); medium goods (>3.5 tonnes but <7.5 tonnes or a minibus 9 or more people); heavy goods (7.5 tonnes or more); other motorised (e.g. agricultural); non-motorised; unknown
Null value	Unknown. If disabled, 'not relevant'.

Data field ID	V5
Data field description	**Intended manoeuvre**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	None
Options	Reversing; parked; slowing / stopping; moving off; going ahead; turning; changing lane / overtaking; unknown
Null value	Unknown

Data field ID	V6
Data field description	**Was this vehicle involved in a collision or a near miss?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If collision, enable V7 and V8. If near miss, disable all remaining 'V' fields. Also feeds into O2.
Options	Collision; near miss
Null value	Unknown

Data field ID	V7
Data field description	**How many collisions was this vehicle involved with?**
User interface format	Drop-down list

Internal database format	Integer
Precedents	Enabled by V6 = yes
Dependents	For each collision enable V8, V9 and V10.
Options	Integers (excluding zero)
Null value	Unknown. If disabled, 'not relevant'

For each collision that this individual vehicle was involved in (recorded in V7), the following data fields should be enabled, thus creating subsets of collision data for each vehicle. This is so that the data for a vehicle which is involved in several collisions within the same incident can be recorded (e.g. a vehicle collides with another and is pushed into a barrier).

Data field ID	V8
Data field description	**What type of collision?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by V6=yes.
Dependents	If vehicle/equipment, enable V9 and V10 and disable V11. If vehicle/vehicle, enable V9 and V11 and disable V10. If vehicle/pedestrian, enable V9 and disable V10 and V11. If unknown, disable V9, V10, V11.
Options	Vehicle / vehicle; vehicle / pedestrian; vehicle / equipment; unknown
Null value	Unknown. If disabled, 'not relevant'

Data field ID	V9
Data field description	**First point of impact on vehicle**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by V8 <> 'unknown'.
Dependents	None
Options	Front; back; offside; nearside; unknown
Null value	Unknown. If disabled, 'not relevant'

This field is independent of whether the vehicle hit something or was hit by something.

Data field ID	V10
Data field description	**Equipment / object hit**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by V8 = vehicle / equipment
Dependents	None
Options	Permanent road feature; temporary sign; temporary barrier (i.e. temporary vehicle restraint); delineator (e.g. cones); works lighting; works tools or equipment; construction materials; unknown
Null value	Unknown. If disabled, 'not relevant'

Data field ID	V11
Data field description	**Other vehicle ID**
User interface format	Drop-down list
Internal database format	Text string
Precedents	Enabled by V8 = vehicle / vehicle

Dependents	None
Options	List automatically populated by B8 and associated V1 for each vehicle.
Null value	Unknown. If disabled, 'not relevant'

3.3.6 People data

Person records are created from B9 and B10, one record for each road user and injured road worker. P1 is created and pre-filled automatically for each and P2 is pre-filled for each. P6 and P8 can be pre-filled for road workers.

Data field ID	P1
Data field description	**Person ID**
User interface format	Automated
Internal database format	Text string
Precedents	Created from B1 in the format 'Incident ID_n'
Dependents	None
Options	N/A
Null value	N/A

If no further details are known, the record for the person is still created.

Data field ID	P2
Data field description	**Person type**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Pre-filled depending on B9 or B10.
Dependents	If road worker, enable P5. If road user, disable P5.
Options	Road worker; road user; unknown
Null value	Unknown

Data field ID	P3
Data field description	**Person location**
User interface format	Drop-down list
Internal database format	Integer
Precedents	None
Dependents	If driver or passenger, enable P4. If on foot, other or unknown, disable P4. If P3 = on foot AND P6 = yes, enable P7.
Options	Driver; passenger; on foot; other; unknown
Null value	Unknown.

Data field ID	P4
Data field description	**Associated vehicle**
User interface format	Drop-down list
Internal database format	Text string
Precedents	Enabled by P3 = driver or passenger
Dependents	None
Options	List automatically populated by B8 and associated V1 for each vehicle.
Null value	Unknown. If disabled, 'not relevant'

Data field ID	P5
Data field description	**Road worker activity**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by P2 = road worker
Dependents	None
Options	Installing/removing TTM; conducting works/inspection; vehicle recovery; other; unknown
Null value	Unknown. If disabled, 'not relevant'

Data field ID	P6
Data field description	**Was this person involved in a collision?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Can be pre-filled for road workers.
Dependents	If yes and if P3 = on foot, enable P7. Also feeds into O2.
Options	Yes; no; unknown
Null value	Unknown.

Data field ID	P7
Data field description	**Associated vehicle (pedestrian collision)**
User interface format	Drop-down list
Internal database format	Text string
Precedents	Enabled by P3 = on foot AND P6 = yes
Dependents	None
Options	List automatically populated by B8 and associated V1 for each vehicle.
Null value	Unknown. If disabled, 'not relevant'

Data field ID	P8
Data field description	**Was this person injured?**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by P6 = yes. Can be pre-filled for road workers.
Dependents	If yes, enable P9. If no or unknown, disable P9. Also feeds into O2.
Options	Yes; no; unknown
Null value	Unknown. If disabled, 'not relevant'

Data field ID	P9
Data field description	**Injury level**
User interface format	Drop-down list
Internal database format	Integer
Precedents	Enabled by P8 = yes
Dependents	Feeds into O2.
Options	Killed; major injury; minor injury; unknown
Null value	Unknown. If disabled, 'not relevant'

For this field, 'killed' means the injured person died on site or within 30 days of the incident, 'major injury' means that the injured person was hospitalised or had seven or more consecutive days off work, 'minor injury' means that the injured person was treated at the scene and had less than seven consecutive days off work.

3.3.7 Other data

Data field ID	O1
Data field description	**Any other consequences**
User interface format	Free text
Internal database format	Text
Precedents	None
Dependents	None
Options	N/A
Null value	Unknown

As for the incident description, for different countries this will be in different languages and is included only for reference.

Data field ID	O2
Data field description	**Overall incident level**
User interface format	N/A
Internal database format	Automated
Precedents	Created from V6, P6, P8, P9 (see below)
Dependents	None
Options	Near miss, non-injury collision, minor injury collision, major injury collision, fatal collision; unknown
Null value	Unknown.

The overall incident level is a derived field created from the data input into fields:

- V6 (Was this vehicle involved in a collision?)
- P6 (Was this person involved in a collision?)
- P8 (Was this person injured?)
- P9 (Injury level)

The following logic applies:

- If V6 = No and P6 = No for ALL vehicles and people involved in the incident then O2 = near miss.
- If V6 = Yes or P6 = Yes for any vehicle or person involved in the incident, then if P8 = No for ALL people involved in the incident, then O2 = 'non-injury collision'.
- If V6 = Yes, P6 = Yes, P8 = Yes for any vehicle or person involved in the incident, then the 'most serious' injury level recorded in P9 (for any person) defines the injury level for the whole incident, e.g. O2 = 'minor injury collision', 'major injury collision' or 'fatal collision'.
- If V6 and P6 are unknown for all vehicles and people involved in the incident, then O2= unknown.

4 Implementation

4.1 Database implementation

The concept of the EuRoWCas data specification is that, provided a database complies with the EuRoWCas common standard, the internal architecture is not relevant. This provides for flexibility in implementation and is not based on a specific data architecture.

The selection of database arrangement will be down to the individual organisation, region or nation recording EuRoWCas-compatible data. Generally, database designs can be either 'relational' or 'flat'. Within a relational database, data is held in multiple tables that are related to each other using reference fields. These tables are organised in tiers, each of which consists of one (or more) tables that relate to the tier above.

This arrangement allows flexibility in the number of values recorded in the lower tiers that relate to a single record at the highest level (e.g. multiple vehicles associated with one incident). This provides for a compact database which has significant advantages when considering import and export of data. It also enables some analyses of the data to be carried out that cannot be carried out using a 'flat' database. It is, however, more difficult to manage than a 'flat' database.

'Flat' databases consist of a single table of data. A flat database can also have one entry for each incident, but difficulties are encountered if the number of associated fields is variable (as for the number of vehicles, people and collisions associated with one incident). In the case of variable numbers of fields, this requires that the database is pre-defined using the maximum number of vehicles, people and collisions which are ever likely to be recorded. This is difficult to define and restricts the flexibility of the system. This issue could be overcome through the use of multiple entries for one incident with a common incident reference ID, however this can restrict the analytical functionality as there is not a unique entry for each incident.

Traditionally, the approach used for accident databases is the relational structure, and it is likely that this approach will be used for any demonstration version of EuRoWCas. However there is no reason that EuRoWCas has to follow this approach if another is preferred, provided it complies with the common standard. For example, if a country was to record only limited base data, this could be kept in a single 'flat' table or even recorded in a spreadsheet provided the field specification structure was followed and the import/export function operated according to the specification.

National-level databases (or databases covering individual regions or contractors) which feed into the EuRoWCas can be created using the same or different approaches. This flexibility is enabled through the harmonisation of the data format at all levels.

Similarly the software product used for the database is not relevant, and the same product may not necessarily be the best solution for all end users. As above, for a simple table a spreadsheet-style application may be best, whereas for a multi-user system this would not be suitable. Specifying the use of one system may subsequently limit the options in terms of software and operating system for end users; as such, the requirement to use a single specific product in order to participate in the data collection process for EuRoWCas may represent an obstacle to delivering a pan-European road worker accident database.

From this it can be seen that any data handling software running under any operating system can be used to host a EuRoWCas-compliant database provided it complies with the common data standard. As an example, possible database software options under different operating systems could include:

- Filemaker (MacOSx)
- Access (Windows)
- Oracle (Linux)
- FireBird (open source multi-platform)
- MySQL server (multi-platform)
- Specific customised software (e.g. the Highways Agency's original AIRS system).
- Web-based system (e.g. the Highways Agency AIRSWeb system)

The common standard allows for any of these options. It also allows for all these options to communicate via a common interface, i.e. the specified data import / export format. As described previously, simple databases could be held in a spreadsheet application provided the data import and export functions were implemented correctly.

4.2 Impact of different levels of data collection

As previously discussed, the highest quality data will be obtained when all fields in the EuRoWCas common data standard are completed for every incident. However, in practice and particularly during initial implementation of the system, this is unlikely to be possible in all countries. As a result, it is useful to examine the impact of different levels of data collection on the usability of EuRoWCas data.

There are four critical fields – country, date, time and location – which must be completed in order to create an incident record. The benefit of having the set of critical fields is that it allows incidents to be counted, but little more. However, even this will be of benefit for road authorities who currently do not have any clear information regarding road worker injury accidents.

Completion of all fields in the database is strongly recommended, but providing the whole database and making completion of some fields mandatory and others optional would allow road authorities to carry out a staged implementation if desired. (Note that regardless of the fields collected, the full dataset must still be stored using the null values for any missing fields. It may be beneficial for subsequent data analysis to record the reason why null values have been entered in this situation, i.e. whether the data field has not been completed or whether the data field has ben disabled.)

As more data are collected, more detailed analysis becomes possible. In terms of a relational database, the lowest level tier of data collected provides the maximum detail at which it is possible to count. For example if only incident circumstances data are collected (i.e. no vehicle or people information) all analytical queries must be of the form *'how many accidents were there where…'*. The criteria forming the filter part of this query (i.e. the '*where…')* depends on the availability of data fields collected within the incident circumstances dataset; collection of more fields enables more filtering to be carried out and so allows for more detailed analysis.

If 'vehicle level' data are collected (or similarly 'people level' data), the questions can become *'how many vehicles / people were….*' or indeed *'how many accidents were there where one or more vehicles / people…'*. As before, the level at which the vehicles or people data can be filtered depends on the level of data collected in each of these areas.

Therefore the data collection levels define what analysis can be carried out on the data and so defines the impact of data collection. If only high-level data are collected, this will provide a baseline but little else, which will limit the potential impact of the collected data. Such data

would be useful for comparison of national road works safety performance and for demonstrating the success of safety interventions but would not be capable of providing detailed information to identify accident mechanisms from first principles or develop operational safety controls (risk mitigation measures).

The impact of providing more detailed data is that the level of support that the data can provide to road worker safety intervention, monitoring and hazard identification also increases. More detail can guide validation of external theories relating incident risk to road works or site characteristics or other factors. Extensive detail will provide the opportunity to identify risks and issues proactively from the data.Unless all fields are mandatory, it may also be beneficial for subsequent analysis to calculate performance indicators showing the level of completion of the (enabled) data fields; this could form an additional derived field for each entry in the database.

4.3 Data input, import and export

The import / export format of the data must be defined and be non-proprietary (that is to say not based on any specific database or software import/export format). The data should be delimited, using plain ASCII character set, which allows for import into text editors. The import / export file should be structured by incident record; the data within the record will define the number of vehicle (field B8), people (B9) and collision (V7) records that will be associated with the incident. The numbers in fields B8, B9 and V7 will define how many data fields for vehicles, people and collisions the receiving database should expect and so import.

Irrespective of whether a fixed record system is used for the database architecture, if any of the fields B8, B9 and V7 are zero value the data exported should not contain any vehicle, people or collision data respectively.

Since the format of the data is standardised, the user input interface (and indeed the data visualisation for data output) can similarly vary and be customised as required. The most practical application of this flexibility is the ability for different languages to be used in different implementations of the system. Additionally, the input interface can be different for different countries, database levels (national, regional, local) or users (contractors, sub-contractors, auditors etc.). A demonstration option for the interface will be created as part of later workpackages.

The EuRoWCas approach is sufficiently flexible to allow for a paper-based form to be used for data capture, allowing offline data capture and for details to be entered later. This is possible, but not recommended as it is best for data to be entered directly onto the database by the person reporting the incident, as using paper forms and ‘punching’ the data can result in loss of information or misinterpretation of the meaning of the data. In addition, given the precedents and dependencies between the data fields, an electronic version is recommended as it provides the opportunity for data validation on entry.

4.4 Data output / visualisation

4.4.1 General

A benefit of the database specification approach outlined in this deliverable is that it enables data output and visualisation applications to be developed and tailored as desired by any end user. There are a range of potential end users, for example:

- A data analyst may require selected fields from the data to be provided as an extract compatible with a spreadsheet package (e.g. Microsoft Excel)
- A statistician may require the “raw data” for import into a specialist data software (e.g. IBM SPSS)
- A non-specialist user may want simple tabulations and basic graphs for management reporting
- A safety manager may want the data reported in specific categories that align to the different reporting criteria particular to their national safety laws

Given the specified design of the data stored, any piece of software that is able to import these data can then be used by any user to process, display, analyse and extract information from the data and then show this in tables, lists or graphically for display purposes.

The data output and visualisation is dependent on the structure of the data that can be extracted from the database. If this is a basic structure, for example comma separated values (CSV), this will import into most spreadsheet and analysis software without issue. However, one of the challenges within this approach is accounting for the differences in number of records if the database has a relational structure (this can be addressed via use of null records where a variable can have multiple entries within the data)

For other visualisation software where basic data structures are not appropriate, the output format of the data will need to be set in different implementations of the system to output data in a native format for the data visualisation package selected. The level of data available in the database will define which data visualisation packages are most suitable. For example, using GPS tagged data would allow overlay of the data onto a GIS application, which would require the data to be in the proprietary format to import or upload into external third party software. If suitable mapping references are not included then the data can only be counted at the lowest practicable level.

Standardised reports would be possible at different levels, depending on the scope of the database. This could include national benchmarking for performance monitoring or improvement or international benchmarking for comparison and cooperation.

Since different implementations of the database will collect different levels of data, it is not simple nor desirable to specify a ‘one size fits all’ data visualisation package. This would either need to allow for visualisation of all data, resulting in redundant functionality for many users, or would potentially have shortcomings for users who were collecting all data. There is also the issue of personal preference when selecting analysis tools and visualisation software, which guides selection away from any single application.

This section considers some of the potential functionality that should be considered in any potential application(s) for data output and visualisation, and provides some options for implementation.

4.4.2 Functionality

BRoWSER deliverable D1.1 - Benefits Case (July 2013) reported on the results of a consultation exercise in which National Road Authorities, Local / Regional Road Authorities and other other relelvant organisations were interviewed. In total, twelve organisations were interviewed across nine European countries, namely:

- Highways Agency, England
- Transport Scotland
- Welsh Government
- Department for Regional Development, Northern Ireland
- Land Nordrhein-Westfalen, Germany
- Hessen, Germany
- Flemish Road Authority, Belgium
- Belgian Federation of Road Contractors
- Walloon Road Directorate, Belgium
- National Roads Authority, Ireland
- Rijkswaterstaat, Netherlands
- Družbe za avtoceste v Republiki Sloveniji, Slovenia

One aspect that was investigated through these interviews was the intended or desired use of these data; interviewees were asked both 'how would you use such data?' and 'what would [the ability to compare road worker accident data with other countries] allow you to do and why would this be useful?'. Some of the responses to these questions are reproduced and summarised below:

- Benchmarking / comparing performance with other countries
- Demonstrating safety record
- Metrics and trend analysis
- Monitoring performance
- Highlighting need to improve performance
- Pinpointing the issues and emerging trends
- Understanding the impact of different policies
- Case building / justification of policy including sccess to more data on which to base policy decisions
- Work scheme improvements
- Operational practice improvements
- Harmonisation of road work practices
- Raising awareness of road contractors / workers
- Understanding accident circumstances to improve operational practice and procedures
- Risk assessment and analysis of operational practices
- Calculation of risk exposition by road work types
- Sharing of best practice

This emphasises the point made in the previous section – that the data output and visualisation should be customised according to the user's needs and requirements, depending on the intended use of the system, rather than based around a presumption of user requirements.

However, there are a number of functionality aspects that can be considered as highly likely requirements for most users, and which are possible with the lowest level of data (i.e. the critical fields). These include:

- Viewing all incidents on a map (at various levels)
- Viewing selected incidents on a map (e.g. by country or date)
- Viewing incidents in 'list view'
- Filtering incidents in 'list view' by any available and relevant parameter
- Selecting an incident from map or list view to display all associated data

As discussed in Section 4.2, as more data fields are collected, more detailed analysis becomes possible. With the four critical fields – country, date, time and location - which must be completed in order to create an incident record, the data visualisation package will be able to do little more than incident mapping (assuming sufficient GIS tagging or equivalent), count the number of incidents, and do basic analysis such as displaying the number of incidents versus time of day. The functionality required would therefore be limited. If more data fields were collected then the list of functionality requirements above could be extended to include more graphical reporting, summary statistics, detailed analytics, trend analysis, performance monitoring and even proactive identification of improvements in road works schemes and operational practice.

4.4.3 Options and examples

As previously discussed, the data visualisation package should not be standardised across all implementations, indicating that there should not be a "EuRoWCas visualisation" product. When identifying a solution there are three potential options that can be considered by an end user. These are an existing product, a custom solution or adaptation of an existing product to meet end user need.

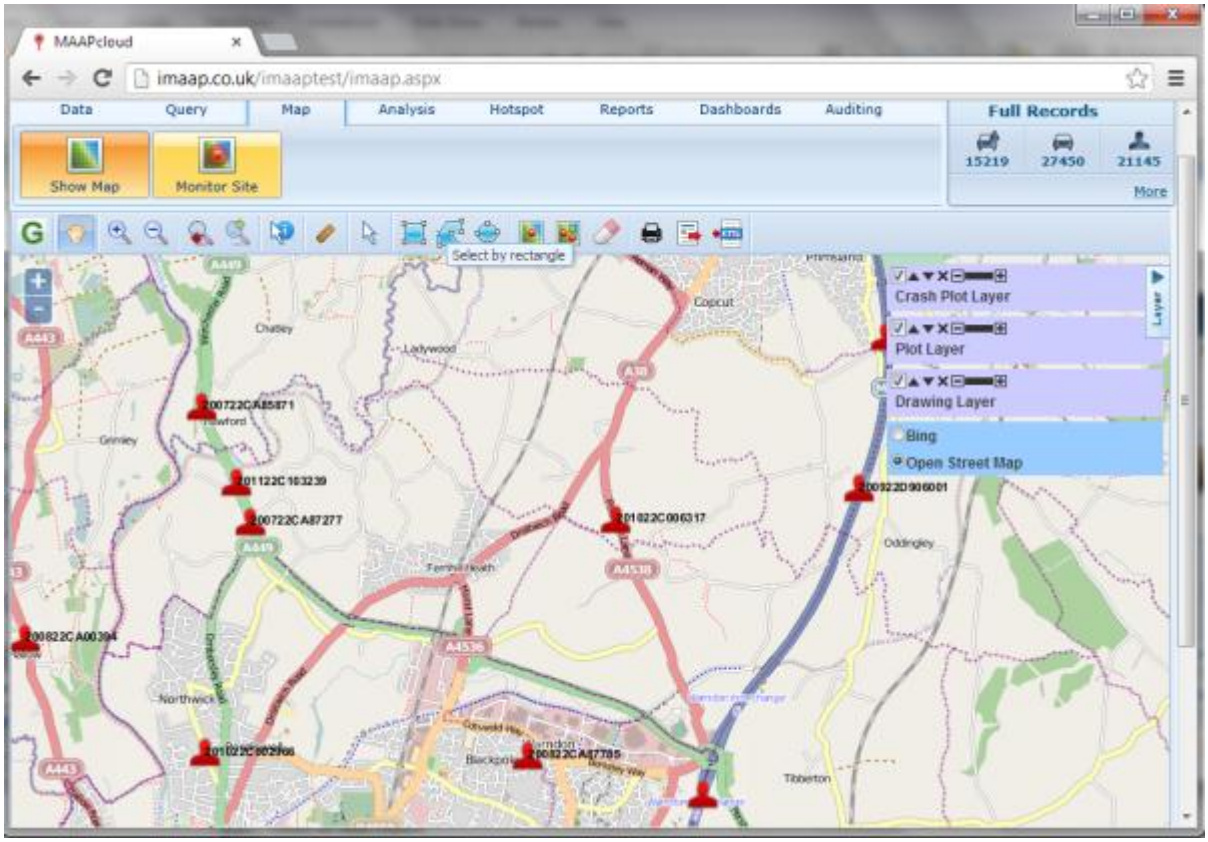
The first option is for an existing commercial off-the-shelf product can be used, requiring no adaptation or customisation. Whilst products exist that are close to this, any existing product is likely to require compromises in the output, the visualisation, the available analysis tools or potentially all three. This is because EuRoWCas is a new database application and so no existing product has the specific functionality to handle all the potential data held in a EuRoWCas database. Therefore, it is believed that, currently, use of a commercial off-the-shelf product is not a possibility.

Secondly, there is the option for a new software product to be built specifically for this purpose, using the database and export specification. This could either be designed to meet an individual user's needs or be a universal tool making use of all data held in EuRoWCas. The latter would require central development; the former is likely to be the most expensive option open to authorities, but would, by definition, fulfil all the user's requirements.

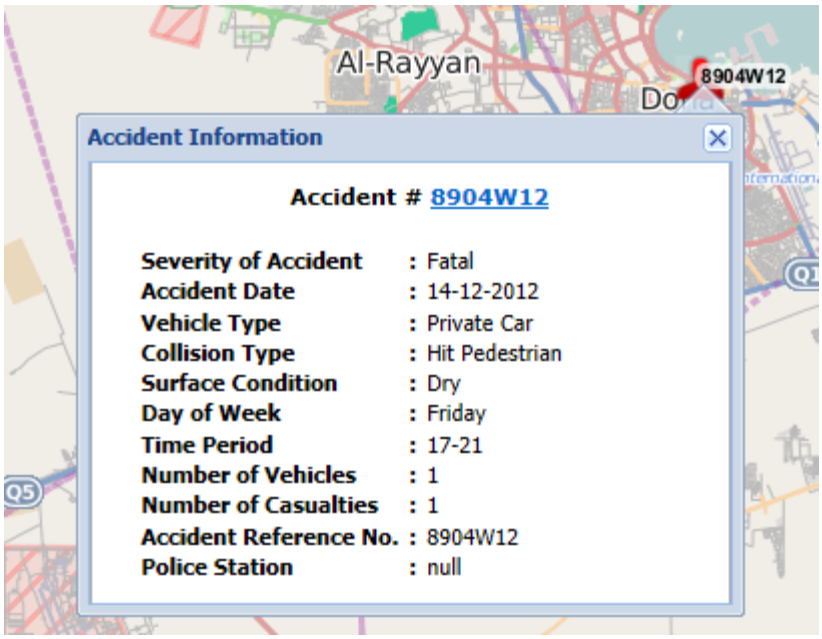
The third option is the adaptation and customisation of an existing product to meet the needs of EuRoWCas as discussed in the previous section. This option is likely to be the most practicable and hence a brief consideration of two existing examples is considered in the remainder of this section.

Customisation of software to handle data from different sources is a well-established technique for developing end-user specific products. One such approach is the iMAAP crash data system - http://www.trl.co.uk/solutions/safety-security/road-safety/imaap/, This is a software product that stores crash data, provides in-depth analyses of the data, monitors trends, identifies and monitors hotspots and produces customised and tailored reporting. Whilst it would not be suitable for off-the-shelf use with EuRoWCas, previous experience of customising iMAAP suggests it could provides a platform based around an existing accident reporting product that could be used for these purposes.

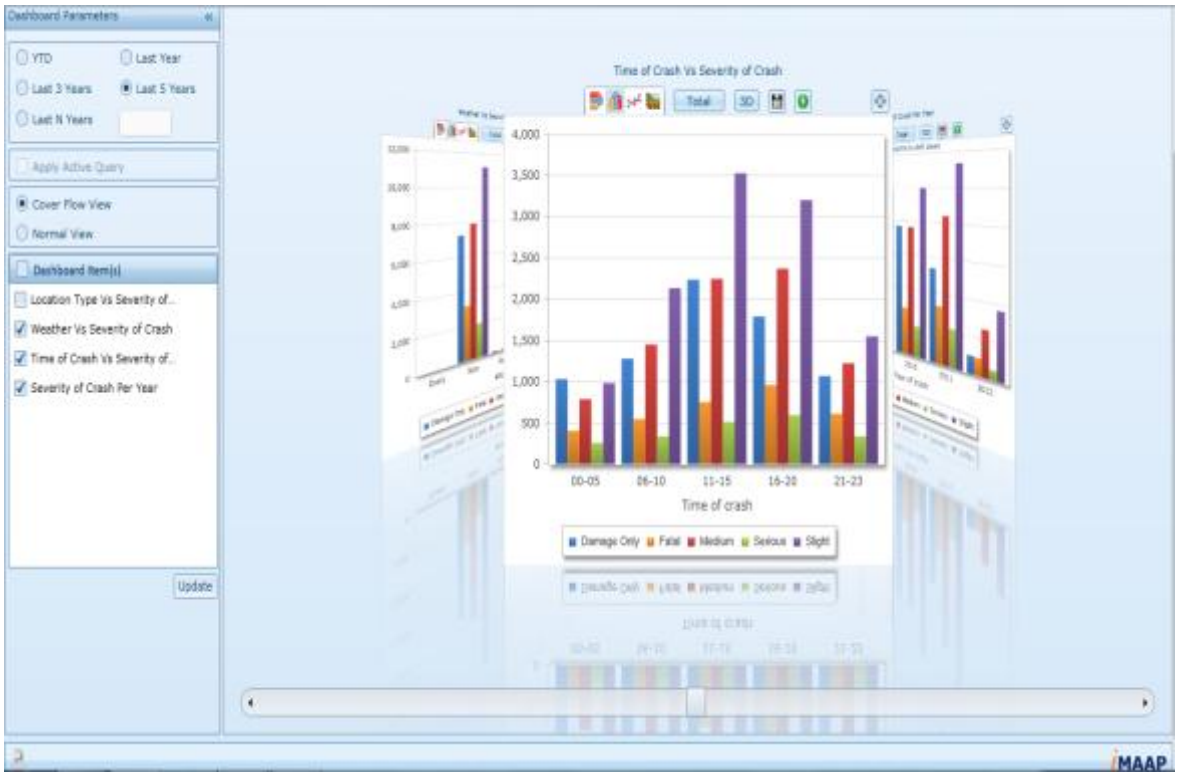
Crash location mapping



Crash details



Analysis and reporting



Another example with similar potential for adaptation and customisation is Ubipix - . http://www.ubipix.com/. Ubipix is a platform that enables GPS tagged images and video to be recorded, uploaded, published, analysed and shared. Current applications include network maintenance (asset inspection, infrastructure condition, sign and markings), road safety survey, traffic management, noise modelling and environment surveys.

Site mapping

Mapping with associated video information

Mapping with graphical analysis

4.5 Next steps

A demonstration of a possible option for this data input and data visualisation will be constructed as part of later work packages in this project. This will include:

- Demonstration (relational) database
- Demonstration data input interface or 'front end'
- Potential data visualisation via external third party software

Note that the data import and export functionality will not be demonstrated as such as part of this work as this implementation will be standalone and import / export connections will not be available to send or receive data.