DRIVing implementation and Evaluation of C2X communication technology in Europe

DRIVE C2X

Accelerate cooperative mobility

Deliverable D35.1

Report on implementation of DRIVE C2X data management to test sites

Version number	Version 1.1
Dissemination level	RE
Lead contractor	VTT
Due date	28.02.2013
Date of preparation	19.12.2013

SEVENTH FRAMEWORK PROGRAMME

Authors

Sami Koskinen (VTT),

Erik Svanberg (SAFER),

Ron Rotstein (Ygomi),

Joerg Rech (Facit),

Fabian de Ponte Müller (DLR),

Florian Häusler (Fraunhofer Gesellschaft),

Harry Wedemeijer (TNO)

Project co-ordinator

Matthias Schulze

Daimler AG

HPC 050 – G003

71059 Sindelfingen

Germany

Phone   +49 7031 4389 603

Mobile   +49 160 86 33308

Fax   +49 7031-4389 218

E-mail   matthias.m.schulze@daimler.com

Legal disclaimer

The information in this document is provided 'as is', and no guarantee or warranty is given that the information is fit for any particular purpose. The above referenced consortium members shall have no liability for damages of any kind including without limitation direct, special, indirect, or consequential damages that may result from the use of these materials subject to any liability which is mandatory due to applicable law.

© 2011 by DRIVE C2X Consortium

Executive summary

DRIVE C2X is an integrated EU project that carries out comprehensive assessment of cooperative systems through extensive European Field Operational Tests (FOTs).

The main purpose of this document is to describe the implementation of data management in DRIVE C2X project for participating test sites. This deliverable also provides a description of the data management practises applied in WP35 "Test data management" to WP43 "Data and data quality", which is where the data processing will continue. This deliverable covers log data sources, tools for monitoring and verifying data and description of data storages. It also discusses privacy. Finally it gives an overview of data enrichment and indicator processing, though post-processing is not a main test site activity.

This deliverable also provides the test sites a set of templates to describe the tests and collected data for collaborative evaluation purposes. From the perspective of an analyst it is essential to get a good top level view of the data i.e. what kind of systems were tested, with what kind of study design and what are the data elements that can be compared to evaluate a specific behaviour. This documentation describing data and tests is mostly referred to as "metadata" within the project, as it is information of the data itself.

Data management is an essential part of the Field Operational Tests (FOTs). Data management not only provides for storing the test results on specified files and databases, keeping track what's where. The scope of data management begins mainly by verifying, that data collection with logging tools matches project requirements. Another type of a starting point is overseeing formulation of agreements to be made with test users, ensuring that privacy, use of collected data and conditions for sharing the data between research partners are described in detail.

As the FOTs collect long periods of in-vehicle logger data – a FOT could e.g. cover the driving of a hundred drivers for one year – the sheer size of logs requires data management processes to be put up. Also, it's not only vehicle speed and other basic sensor data that's collected, but also filled questionnaires, travel diaries, internet service logs e.g. about local traffic status, and manual annotations of video material.

Data management sets up processes for collecting log files manually or automatically, sets up servers, decides backup practises and enables sharing the data. It monitors data collection and frequently validates stored data, both manually and with the help of dedicated software tools. Even basic software scripts can be very effective for pointing out errors in data collection or system failures. The main motivation for frequent monitoring of data collection is about not having to repeat expensive tests due to broken-down systems or e.g. configuration errors.

DRIVE C2X follows the previous TeleFOT project in setting up a central storage for collected data. A central data depository makes it easier to access the data and enables collaborative analysis of several tests. It can also help in harmonizing data formats and running common post-processing for all data. These common processing tools include data enrichment such as map matching and retrieval of weather data, and processing a common set of indicators such as average speed during a test drive.

When a test site moves their data to a central depository, they document the ownership of data and conditions for sharing it. They also document the test itself from an analyst's

perspective: for example what was the study design, vehicle properties and anonymous information of the test users such as their age and driving experience. Further, the format of collected log data is described in detail, including information of the sources that generated the data. These documents are test and data description documents, metadata.

All this documentation is necessary for an analyst who may have not been part of the data collection, to start analysing its data. Most aspects of the documentation are also necessary for the test leaders, to document details, dates and formats when they're still remembered, and to answer analysts' main questions only once.

This document outlines the key structure, processes and tools made available for DRIVE C2X test sites and analysts by Work Package 35. Key support for this deliverable has been provided by WPs 25 (regarding logging), 43 (regarding data processing and quality) and 45 (requirements of impact assessment).