twinsafe

Twin-Safe: Advancing Road Safety Through Twinning

Deliverable 3.1

Open database of collected data

https://twin-safe.com

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them

Project identification

Grant Agreement No	101159114
Acronym	TWIN-SAFE
Project Title	TWIN-SAFE: Advancing Road Safety Through Twinning
Start Date	1 May 2024
End-Date	30 April 2027
Project URL	https://twin-safe.com

Document summary

Deliverable No	3.1
Deliverable Title	Open database of collected data
Work Package	3 - Research of road infrastructure impacts on driver behaviour
Contractual due date	M18
Actual submission date	M18
Nature	Data
Dissemination level	Public
Lead Beneficiary	University of Zagreb, Faculty of Transport and Traffic Sciences
Responsible Author	Tom Brijs
Contributions from	University of Zagreb, Faculty of Transport and Traffic Sciences: Dario Babić; Hasselt University: Tom Brijs

Please refer to the document as:

Brijs, T & Babic, D. et al. (2025). *Open database of collected data.* Deliverable 3.1 of the EC Horizon Europe project TWIN-SAFE: Advancing Road Safety Through Twinning.

Revision history (including peer review & quality control)

Version	Issue date	% Complete	Changes	Contributor(s)
v1.0	October 16, 2025	100%	Finalised deliverable	Dario Babić, Tom Brijs

Disclaimer

Copyright

Table of contents

List of Figures

Glossary and abbreviations

Word / Abbreviation	Description
WP	Work package
TWIN-SAFE	The project on the topic of Advancing Road Safety Through Twinning
FTTS	University of Zagreb, Faculty of Transport and Traffic Sciences
HU	Hasselt University

1 Introduction

Work Package (WP) three of TWIN-SAFE project represents a research component of the overall project that aims to enhance road safety on rural roads by examining the complex relationship between driver characteristics, infrastructure, and driving behaviour. The study will identify high-risk road segments where human factors and infrastructure elements interact to increase hazards. To achieve this, the project team recruited a representative sample of Croatian car drivers and motorcycle riders, for a naturalistic driving study. The project team used TOECAN mobile app to capture critical driving parameters such as speed, acceleration, and mobile phone use. In parallel, comprehensive infrastructure data is being gathered through cooperation with Croatian roads Ltd. and GIS-based assessments, including the application of the iRAP methodology to evaluate road safety. The collected data will undergo detailed statistical analysis to identify correlations between driver behaviour and infrastructure features, leading to the identification of risky road segments. Ultimately, WP3 aims to develop and test tailored safety solutions for high-risk areas, advancing the principles of the Vision Zero and Safe System approaches.

One of the activities in WP3 is related to open data sharing, i.e. providing the collected data related to naturalistic driving study to the general research community in accordance with open access principles. Therefore, this deliverable provides insight into the overall data collection process, description of the available dataset and the procedure of obtaining aforementioned data.

Overall, the deliverable serves as a source of information on naturalistic driving study conducted under WP3 of TWIN-SAFE project, as well as simplified handbook on how to conduct naturalistic driving study and publish openly collected data. Therefore, the deliverable is primarily aimed at the FTTS researchers as it provides framework for conducting naturalistic driving studies and publishing open access data. Furthermore, as this deliverable provides protocol on how to obtain the data collected in TWIN-SAFE project, it is also of high interest to research community in the road safety field.

2 TOECAN app

The TOECAN mobile application was developed under Horizon 2020 i-DREAMS project (Grant Agreement No 814761) coordinated by Hasselt University. The goal of the mobile application is to collect individual data on the driving habits of car drivers and motorcyclists, aiming to understand how different individual characteristics influence driving patterns and behaviour across various segments of rural roads under different conditions.

The system integrates the O7SDK, which collects data from smartphone sensors (e.g., GPS, accelerometer, gyroscope) during a trip. This data is used for both real-time and post-trip interventions, and is subsequently sent to the OSeven backend for processing after the trip concludes. The system has several features such as:

- Detection of handheld mobile phone use during driving
- Display of trip information and assignment of ratings based on various safety parameters
- Comparison of the user's driving results with other drivers in the project
- Provision of relevant information on safe driving practices
- Encouragement for users to set goals for improving safe driving
- Answers to frequently asked questions (FAQ)

Within TWIN-SAFE project, the application is used solely for monitoring and collecting information such as speeding, acceleration, deceleration, and mobile phone use during driving. Some of the other features normally available in the app were visible within the application but their access was restricted.

When installed, each trip has been automatically recorded and chronologically organized from the most recent to the oldest in the trip overview function. In addition, the app calculates the overall score for each based on the results of various parameters. By selecting a specific trip, app provides more details about the selected journey. The displayed score indicates how well the driver performed regarding the specific parameter for that trip. Swiping left or right allows users to view results for other trip parameters. Tapping on "visualize on map" displays the trip and all recorded events during the journey on an interactive map. Tapping on a specific event on the map reveals additional information about the event (e.g., event time, event status = moderate/high, event duration).

The "Ratings" function allows users to view scores for a selected time period or for different parameter types. In addition to individual ratings for parameters such as acceleration and deceleration, a rating for vehicle control is also displayed. The chosen time period for ratings can be adjusted using the calendar function.

Ratings are calculated on a per-trip and per-parameter basis. The rating for each parameter is based on the number and intensity (and in some cases, duration) of events measured during the trip for that parameter. A higher number of events with greater intensity (status = high) over a longer duration results in a poorer rating for that parameter for the trip. Each driving parameter has a specific formula developed for scoring. The parameter scores are accumulated into overall trip, daily, weekly, monthly, and total ratings.

The aggregation of scores considers the trip length (or duration of movement for some parameters). Therefore, a poor score on a short trip has less impact on the daily rating compared to a poor score on a long trip.

1. Speeding

   The speed rating considers how many speeding violations were recorded during the trip, along with their duration and the percentage of the trip spent above the speed limit. The measured vehicle speed (taking into account a tolerance of 5%) is compared with the permitted speed obtained from the digital speed limit map. The speed limit map is frequently updated to include the latest information on speed restrictions for road segments. Although highly accurate, it can never be 100% precise.

   An overall score below 100 does not necessarily indicate that the driver is not respecting the speed limits. Possible inaccuracies may exist in the measured vehicle speed or in the speed limit information on the map. It is more important to compare speeding scores between drivers and over time to assess improvements or deteriorations.

2. Sudden acceleration/deceleration

   Sudden acceleration, deceleration, or steering are signs of unsafe driving behaviour and can lead to increased fuel consumption, brake wear, and higher risk of accidents. TOECAN uses reference threshold values (G-force) based on results from the i-DREAMS project, which collected over 300,000 hours of driving data and more than 3.5 million kilometres of travel data.

   Trip events are indicated on the dashboard with colour codes that reflect the intensity of the event, ranging from light yellow to dark purple.

3. Distraction (mobile phone use during driving)

   Regardless of trip length, mobile phone use during driving results in a zero score (score = 0), as such behaviour is strictly penalized.

Figure 1 TOECAN app

3 Participant recruitment and data collection process

To recruit participants, the project team utilized a range of platforms to effectively reach the target audience. These included social media channels and the Faculty of Transport and Traffic Sciences website. Additionally, the team engaged with local communities, leveraged personal contacts, and reached out to individuals who had previously participated in studies conducted at FTTS. Several motorcycle clubs, including Moto Gymkhana Croatia, the Motorcycle Association "Karavana za život," and the BMW Motorcycle Club Zagreb, were also contacted to broaden the recruitment efforts.

During the recruitment process, project coordinator explained to participants the purpose of the study, data collection and privacy, participant rights and responsibilities. Also, instructions for use for the TEOCAN app were provided to all participants, and project team replied to all the participants questions regarding the study itself and their role in it.

The entire recruitment process complied with FTTS and University of Zagreb ethical guidelines, and the research team obtained all necessary approvals and consents from the ethics committee. Special attention was paid to data privacy in accordance with the Data Management Plan.

From the beginning of the study until October 27th 2024, a total of 96 participants, and agreed to take part in the study. Of these 96 participants, 24 withdrew from the study. The final number of participants is 72. The sample consists of 23 women (31.94%) and 49 men (68.06%). Of these, 19 (26.39%) are motorcyclists, while 53 (73.61%) are car drivers. Among the 19 motorcyclists, 3 are women (15,79%) and 16 are men (84.21%), while among car drivers, 20 are women (37.74%) and 33 are men (62.26%). Overall, the number of recruited participants significantly overpasses the milestone - at least 40 participants (25 car drivers and 15 motorcyclists).

The project team ensured that the sample is balanced to represent the structure of Croatian drivers. Overall, in Croatia, there are 2,318,679 registered drivers, with 995,289 (42.92%) being female and 1,323,390 (57.08%) male. Out of 757,412 motorcycle drivers, 38,710 (5.11%) are female, while 718,702 (94.89%) are male.

In June 2025 data collection phase based on the naturalistic driving study using TOECAN app was finalised. In total, the collected sample consisted of 28750 trips for car drivers and 6674 trips for motorcyclists. Overall, 8338 hours of driving for car drivers and 2106 for motorcyclists was collected, reaching in total 417399 km for car drivers and 107659 for motorcyclists. This means achievement of second KPI related to O1 - “collected minimum of 2000 kilometres of recorded data in Croatia”.

4 Data description and availability

As stated, the collected dataset contains naturalistic driving data of various trips of car drivers and motorcycle riders recruited in the EU Horizon Europe TWIN-SAFE project.

The data consists of several main data parts[1] and is structured in .JSON format:

1) Overall trip characteristics: these include trip ID, start date and time, end date and time, trip distance and trip duration
2) Trip scores: for each trip, an aggregated score (min score = 0, max score = 100) is calculated for several driving behavioural parameters including: acceleration, deceleration, speeding, handheld mobile phone use.
3) Trip event data: this section of the data contains for each trip a list of trip events with specific information for each event.
    - Acceleration: location (lat, lon), status (low, medium, high), duration (seconds)
    - Deceleration: location (lat, lon), status (low, medium, high), duration (seconds)
    - Speeding: location (lat, lon), status (low, medium, high), duration (seconds), distance (km), maximum speed (kph), speed limit (kph)
    - Hand-held mobile phone use: location (lat, lon), duration (seconds), distance (km)

No driver information, neither starting or end location of the trip is included in the dataset to safeguard the privacy of the participants in the study. Individual trips can therefore not be linked to an individual.

The data were collected by means of a smartphone telematics application with automatic trip recording (start and stop of a trip). No manual intervention from the participants was needed with the app running in the background.

A sample data set is available via the Zenodo platform. It can be downloaded from the following link. The data access procedure is as follows:

1) Download the sample data from the Zenodo platform (see below for the link) and develop an understanding of the data structure and metadata;
2) Download the data access application form, fill and submit it via e-mail provided in the application form;
3) The TWIN-SAFE project co-ordinator will read the details provided by the data user. Upon approval of the data request, a member of the TWIN-SAFE team will work on a delivery schedule with the data user.
4) The data user will comply with all the terms and conditions of the data use and appropriately cite the TWIN-SAFE project data in all research products (i.e. research project deliverables, research articles).

The data access application form is available at this link: https://zenodo.org/records/17279578.

[1] The dataset also contains 'null' information for a few other parameters (e.g. steering, collision avoidance, headway, etc). Since these were not collected as part of the TWIN-SAFE project, these parameters contain no data.

Annex 1: Data access application form

twinsafe

DATA ACCESS APPLICATION FORM

Contact Details of Data User

Name: _____

Position: _____

Organisation: _____

Address: _____

Telephone: _____

E-mail: _____

Web: _____

Project Description and Data Use

Q1 Who is the ultimate client for this request for use of the TWIN-SAFE Data? (This information is for our records only and will not be used to contact the client directly. In case it is same as above just write "same as above".)

Agency/Company Name:

Address:

Contact Name/s:

Telephone:

E-mail:

Web:

Q2 Please provide a project title and a description of the scope and objectives of the project, (if the data is requested for a specific task of a large project, please also discussed the scope and objective of that task).

Q3 Please provide a description of how you plan to make use of the TWIN-SAFE data (e.g. the modelling methodology, fusing other data sources.)

twinsafe

DATA ACCESS APPLICATION FORM

Terms & conditions:

1) **Please sign below to confirm that you understand and are aware that the TWIN-SAFE Project team accepts no liability for any errors or omissions in the data.**
2) **The user understands that the requested data is available once a non-disclosure agreement is signed between the two parties and the user of the data is bound to respect the terms and conditions stated in it.**
3) **The user understands that the data originates from a 3rd party source and use of the data is at the users own risk. Following completion, please return to project co-ordinator (E-mail: dbabic2@fpz.unizg.hr).**

Once your project has been accepted, a member of the TWIN-SAFE project team will work on sending a non-disclosure agreement and also data delivery schedule with you.

I have read and understood the Terms of Use. I agree to follow the Protocol for use of the data	Signature: Print name:
Date of application:	

For internal use only	
Date of Receipt:	
Approved by:	
Approval Signature:	
Approval Date:	
Project number and region	

Annex 2: Sample data


```
{
  "trip_id": "3BF6C352-A72B-427F-875F-B1D0A0BE44EB",
  "trip_start": "2024-11-06T06:25:54.080000+00:00",
  "trip_stop": "2024-11-06T06:33:38.942000+00:00",
  "distance": 4.0,
  "duration": 461,
  "scores": {
   "health": {
    "score": 100,
    "fatigue": {
     "score": null
    },
    "distraction": {
     "score": 100
    }
   },
   "vehicle_control": {
    "score": 0.0,
    "acceleration": {
     "score": 0
    },
    "deceleration": {
     "score": 0
    },
    "steering": {
     "score": null
    }
   },
   "road_sharing": {
    "score": null,
    "tailgating": {
     "score": null
    },
    "lane_discipline": {
     "score": null
    },
    "overtaking": {
     "score": null
```

```
    },
    "forward_collision_avoidance": {
      "score": null
    },
    "vulnerable_road_user_collision_avoidance": {
      "score": null
    }
  },
  "speed_management": {
    "score": 0.0,
    "speeding": {
      "score": 0.0
    }
  },
  "safety_devices": {
    "score": null,
    "mobileye_use": {
      "score": null
    },
    "idreams_app_use": {
      "score": null
    }
  }
}
```