



Needs, wants and behaviour of “Drivers” and automated vehicles users today and into the future

Contract No: 815001

D9.7: Revision of D9.4 Data Management Plan

Version 1.0

Work package	WP9: Project Management
Activity	A9.3
Deliverable 9.7	Revision of D9.4 Data Management Plan
Authors	Evangelia Gaitanidou, M. Britsas, M. Loukea, Evangelos Bekiaris (CERTH)
Status	Final
Version	Final
Dissemination Level	ORDP
Document date	27/10/20
Delivery due date	30/10/20
Actual delivery date	30/10/20
Reviewers	DEUSTO/ NTUA



This project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement no 815001.

D9.7: Revision of D9.4 Data Management Plan

Version History

Document history			
Version	Date	Modified by	Comments
0.1	22/09/20	M. Britsas	Table of Contents was created.
0.2	23/10/20	K. Touliou	DPIA templates results consolidated in one table.
0.3	12/10/20	M. Britsas	Chapters 1, 2 were added.
0.4	14/10/20	M. Loukea	Chapter 3 was added
0.5	22/10/20	E. Gaitanidou	Chapters 4 and 5 were added.
1.0	27/10/20	K. Touliou	Submitted for internal peer review.
Final	30/10/20	E. Gaitanidou	Incorporated feedback from reviewers and submitted to EC

Legal Disclaimer

This document reflects only the views of the author(s). Neither the Innovation and Networks Executive Agency (INEA) nor the European Commission is in any way responsible for any use that may be made of the information it contains.

Table of Contents

1 Introduction	6
1.1 Purpose of the Document	6
1.2 Methodology	6
1.3 Intended audience	6
1.4 Interrelations	6
1.5 Use of Google drive cloud for project storage	7
2. Drive2theFuture data	7
3. Data controllers and processors.....	15
4. Data Privacy Impact Assessment (DPIA).....	18
4.1. Do I have to do a DPIA?	19
4.2. Step 1: Identify the need for a DPIA.....	19
4.3. Step 2: Describe the processing	21
4.4. Step 3: Consultation process.....	25
4.5. Step 4: Assess necessity and proportionality	25
4.6. Step 5: Identify and assess risks.....	27
4.7. Step 6: Identify measures to reduce risk.....	27
4.8. Step 7: Sign off and record outcomes	28
5. Conclusion and next steps	29
References	30

List of Tables

Table 1. Data clusters template and instructions.....	8
Table 2. The characteristics of data clusters (WP2-WP5)	10
Table 3. The characteristics of data clusters (WP6-WP8)	13

List of Figures

Figure 1. Data flow (WP2).....	24
--------------------------------	----

D9.7: Revision of D9.4 Data Management Plan

Abbreviations List

Abbreviation	Definition
Art.	Article
AV	Autonomous Vehicle
C-ITS	Cooperative Intelligent Transport Systems
COVID	Coronavirus Disease
D	Deliverable
DMP	Data Management Plan
DPA	Data Protection Authority
DPIA	Data Privacy Impact Assessment
DPO	Data Protection Officer
GB	Giga Byte
GDPR	General Data Privacy Regulation
HMI	Human Machine Interface
HW	Hardware
ITS	Intelligent Transport System
M	Month
MAMCA	Multi-Actor Multi-Criteria Analysis
MB	Mega Byte
N/A	Not Applicable
ORDP	Open Research Data Pilot
SATI	SHAPE Automation Trust Index
SUS	System Usability Scale
SW	Software
TBD	To Be Decided
TLC	Time-to-Line-Crossing
TMC	Traffic Management Control
TTC	Time-to-Collision
UEQ	User Experience Questionnaire
UPQ	User Profile Questionnaire
VDL	Technology Acceptance Questionnaire (Van der Laan)
WP	Work Package

D9.7: Revision of D9.4 Data Management Plan

Executive Summary

This report is the update and revision of the information first presented in D9.4 'Data Management Plan' (DMP) for the Drive2theFuture project. The purpose of this update is to set out the main dimensions and characteristics of the project's data management policy for the datasets generated by the project following the regulations of the Pilot action on Open Access to Research Data of Horizon 2020. It is aligned with both D10.2 'H-Requirement no. 4', which is the Ethics policy and D9.4, which comprises the Data Privacy Policy of the project.

In addition, this updated Data Management Plan contains the complete structure of the database and descriptions of the metadata files to enable self-explainable (re)use of datasets by external parties. Long-term re-usability of these data is of substantial importance, especially in the field of automated driving experience. Embargos (if any) for parts/segments of data, models, evaluation dimensions (e.g. acceptance, trust), the surrogate and horizontal impact and metadata indicators/estimators will be set by the Partners who own them in collaboration with the Data Manager of the project. Additionally, the final location and format of open Drive2theFuture datasets is defined (i.e. Zenodo). Finally, the latest DPIA report is included along with the list of data controllers and processors so far.

In this version the following elements of the Data Management Plan framework are set:

- Purpose of the document, intended audience and interrelations (**Chapter 1**)
- Drive2theFuture data (**Chapter 2**)
- Data Privacy Impact Assessment (DPIA results) (**Chapter 3**)
- Data controllers and processors reports (**Chapter 4**)
- Conclusion and next steps (**Chapter 5**)

Post-processed datasets, free from any private/personal and identifiable information, will reside in the Drive2theFuture partners safe repositories. As the COVID situation has affected the pilots' timeline, not all analytic descriptions of the complete and (non) shareable datasets that will be created during the pilots (WP5), the analyses to follow (WP5 and WP6) and other WPs that require post-processing (e.g. for the preparation of the Use Cases in WP1 and the simulation modelling conducted in WP2) are included in the data clusters table. Thus, no decision on shareable datasets can be made if all pilots have not been conducted.

In conclusion, this Deliverable is a living document, and it will be regularly updated during the lifetime of the project. If official updates are necessary, then they will be included in the final version of the Project Management Plan (D9.8; M35) that will present the final data clusters' table, the selected datasets to be uploaded in Zenodo as well as the open-access publications. If another DPIA is required, it will be performed and reported in the final project report.

D9.7: Revision of D9.4 Data Management Plan

1 Introduction

1.1 Purpose of the Document

This Deliverable is a revision of D9.4 ‘Data Management Plan’, based on the current progress of the project in relation to data categories collected by the pilot sites so far and the second Data Privacy Impact Assessment (DPIA) performed across the project. Hence, this update includes a refined and elaborate account of data gathered so far and their “integration metadata” that might result from the integration of different systems to one platform as well as final indicators and labels fed to behavioural modelling (wherever these are available). Furthermore, this report includes the characteristics of the data and surrogate variables, their storage properties and the parts that can be communicated to public and shared with other research communities (see Table 2 in Chapter 2).

1.2 Methodology

The DPIA template (Chapter 4), the data clusters’ spreadsheet (Table 2) and the data controllers and processors template (Chapter 3) were initially communicated to each WP leader of the project to identify the needs and further the Activities and partners who might need to complete them. Then the authors established communication with the partners to ensure they have correctly completed the templates and, in the process, provide assistance and/ or clarification, in case it was required. DPIA was not performed by all partners. They first completed the initial exercise (first box in the DPIA template) to find out if they must perform a DPIA. If they answered ‘No’ to all Box 1 questions, it meant that they did not have at this stage to perform a DPIA and complete the respective template. At this phase of the project, this was true for WP1, WP4 and some activities in WP3. Some partners in WP5 also did not have to complete the DPIA. However, as some of the pilots are delayed due to the current COVID situation, some of the DPIAs will potentially be performed and reported at the final version of the project report. Therefore, the DPIA is performed only for partners who believe they might collect some data that could potentially create a risk, but still the risks are low and the controllers and processors have been identified for those data that are available. These are the data that are presented in Chapter 2 and they are not exhaustive. Data to be collected and subsequently decisions to be drawn on them will be made later before their collection. Hence, these decisions will be reported in the final DPIA report. Only agreed upon and anonymised datasets will be shared via Zenodo as well as the complete list of open source publications. Data sharing is part of the project’s research extraversion potentially and contribution to interested research communities that are considerably expanding in the area of autonomous transportation, ITS and smart integrated urban mobility.

1.3 Intended audience

This Deliverable serves to inform the Consortium about the updates in the existing data management plan and processes related to any data handling (i.e. collecting, controlling, storing, sharing, cleaning, pre-processing, analysing, and reporting) during the lifetime of the project. Primarily, partners involved in any of these data handling processes have contributed to this update and its content is important for evaluating the potentiality of data protection risks and respective mitigation strategies and reach decisions about the data collected/ will be collected, report on how they are/ will be managed, and what will happen to them after the end of the project. These decisions will be presented in the final update of this Deliverable when all data types have been specified and collected.

1.4 Interrelations

The interrelations have been highlighted in the previous version of this Deliverable and the information has not changed, so it will not be repeated here. However, we shall emphasise that the Data

D9.7: Revision of D9.4 Data Management Plan

Management Plan (DMP) complies with both the Ethics Manual (i.e. Ethics policy) and the data privacy policy (D9.4) and these constitute the primary interrelations.

1.5 Use of Google drive cloud for project storage

The project repository was moved from Dropbox to Google folders and this process is addressed here. This big cloud provider is GDPR compliant (<https://privacy.google.com/businesses/compliance/>). The use of Google documents is compliant to GDPR, but there are some rules partners must follow to ensure they remain compliant. The General Data Protection Regulation (GDPR) is mandated to protect private data. This means that partners have also to work in their internal policies to ensure their repositories, websites and data systems comply. The most common channel to communicate with end-users through the Google drive is the Google form. Therefore, we provide below the GDPR compliant way to communicate with end-users via Google forms.

End users are given specific rights by GDPR:

- **Right of access:** The end user should be informed somewhere (decision on where depends on the context of the form) in the form or website that their personal information is going to be collected and there will be no unauthorized sharing of the said data. The website or the email is the origin of communication or where the form is sent out from.
- **Right to be forgotten:** The end users of the forms should at any time request for the deletion of the collected data. At some point, entities will be required to delete any archived data. This protects the data from falling in the wrong hands.
- **Right to rectification:** The end users of the forms should have the right to rectify any erroneous information at any time.

The GDPR affects all signup forms where personal data of participants like name, address or email is collected. The rules do not apply to those random quizzes that one undertakes online. To make the forms GDPR compliant, a consent tick box will be added in the forms. This will define how the data is to be used and parameters for not sharing the data. It will also define the terms and private policy of the entity. Partners have been advised to avoid using the 'Share with anyone who has a link' option, when sharing Google documents.

2. Drive2theFuture data

It is imperative to consider that data collection adheres to the international, European, and national guidelines and legislations, as it has been discussed in D9.4 and the project's data policy, as presented again in D9.4. Any WP collecting data or offering technologies that handle data has been contacted and completed the three documents mentioned in the previous Chapter (i.e. the data clusters' spreadsheet, the DPIA template and the data controller and processor template).

In this template (Table 1), partners responsible for data handling defined these data types and subsequently, the data management technical, content, and quality specifications for storing and communicating the respective datasets. For data collected during the pilots, all security and ethical guidelines and standards do apply. In addition, data owners reach a decision upon data visibility and sharing limitations.

Post-processed datasets, free from any private/personal and identifiable information, will reside in the Drive2theFuture data repositories of each partner who controls them. Each one will include analytic descriptions (i.e. metadata for easily communicating them outside the Consortium) of the complete

D9.7: Revision of D9.4 Data Management Plan

and (non) shareable datasets that will be created during the pilots (WP5), the analyses to follow (WP5 and WP6) and other WPs requiring post-processing (e.g. for the preparation of the Use Cases in WP1 and the simulation modelling conducted in WP2). In addition, the final overview of these datasets will contain the complete structure of the database, descriptions of the metadata files to enable self-explainable (re)use of datasets by external parties. Long-term re-usability of these data is of substantial importance, especially in the field of automated driving experience. Embargos (if any) for parts/segments of data, models, evaluation dimensions (e.g. acceptance, trust), the surrogate and horizontal impact and metadata indicators/estimators will be set by the Partners who own them and the Consortium. Additionally, the final location and format of open Drive2theFuture datasets will be defined. For the already defined and collected data, these decisions have been made. For those datasets to be defined and collected at the pilot and user engagement activities to follow, we do not anticipate these decisions to change much. Nevertheless, they will be re-visited and reported when data collection is complete. Table 1 shows the data characteristics collected per data cluster by the partners. The partners received this template with instructions (right column) and with one row completed with dummy content (e.g. collection of pseudonymised questionnaires of 5MB) to ensure they understand what type of descriptions were expected and to avoid collecting erroneous or non-relevant information. It is evident from Table 2 that data categories are diverse, and partners have different storage processes and durations; however most of them are subjective data that are often easier to consolidate. Hence, when all data have been collected, data to be publicly shared (if any will be) need to be harmonised before sharing based on the following template (Table 1).

Table 1. Data clusters template and instructions

Partner	<i>Name of Partner</i>
WP/Activity	<i>Name of WP and Activity these data are collected</i>
Pilot site & Phase (if relevant)	<i>Name of Pilot site and Pilot phase (if relevant)</i>
Collected/ Created	<i>Collected/created</i>
Name	<i>Name of the data/ metadata/ exploitable result</i>
Description	<i>Description of the data/ metadata</i>
Category	<i>HW/SW/ Algorithm/Raw data/ Dissemination material/etc.</i>
Standards	<i>Important standards applied even for ensuring interoperability</i>
Type	<i>Document/video/images/Source code/etc.</i>
Format	<i>File extention/ prototype</i>
Size	<i>Size in MB/GB</i>
Owner	<i>Partner name/ Consortium/ external stakeholder</i>
Privacy level	<i>Public/ consortium/ partner/etc.</i>
Repository during the project (for private/public access)	<i>BAL.PM or other Open access repository/ partner storage(private cloud/private drop box), etc.</i>

D9.7: Revision of D9.4 Data Management Plan

Data sharing	Open (O): Open for public disposal	Embargo (E): It will become public when the embargo period applied by the publisher is over. In case it is categorized as embargo the end date of the embargo period must be written in DD/MM/YYYY format.	Restricted (R): Only for project internal use.	Each data set must have its distribution license.	Provide information about personal data and mention if the data is anonymized or not. Tell if the dataset entails personal data and how this issue is considered.
Back-up frequency	Daily/ Monthly/ Yearly/Once				
Destroyed at the end of the project?	No (1)/No (2)/No (3)/ Yes / Unnecessary				
Duration of preservation (in years)	Number of years				
Repository after the project	BAL.PM or other Open access repository/ partner storage (private cloud/private drop box), etc.				
Comments	Any other comments you may have and are important to categorise/ describe the data cluster.				

This template was distributed to all partners who are involved in data handling, either indirectly (through their technologies) or directly (through user engagement), to complete. Collected templates were consolidated and the current data clusters and characteristics are described in Table 2 (for WP2 to WP5) and in Table 3 for WP6 to WP8, respectively. These tables depict the current data handling status; however, they will be updated by the end of the project (i.e. in the final project report) with all project data descriptions. If needed, additional metadata files will be created to accompany the shared datasets on Zenodo. These metadata files will help interested external researchers to understand the project's data types and their characteristics in order to ensure their re-use is possible. The latter will happen only if the data owners agree to do so.

Other data (e.g. vehicle and traffic data like speed, acceleration, deceleration, Time to Collision (TTC), Time to Line Crossing (TLC), headway (sec), average number of passenger trips per day, average passenger-km travelled per day, etc.) are not included in these tables, as their descriptions are not available yet, but they will be included in their final versions.

D9.7: Revision of D9.4 Data Management Plan

Table 2. The characteristics of data clusters (WP2-WP5)

Partner	INF	FhG/IAO USTUTT	SWARCO	Tuco	VTI	VTI	TOI	PZM	PZM	DEUSTO	PZM
WP/Activity	WP2/ A2.5	WP3/ A3.2	WP5/ A5.1	WP 5 / A5.4	WP5/ A5.4	WP5/ A5.4	WP5/ A5.4	WP5/ WP5.4	WP5/ A5.5	WP5/ A5.8	WP5/A5. 8
Pilot site & Phase (if relevant)	N/A	All pilots involved in HMI testing and development, mainly Phase II	Italian Pilot site/Phase II	Danish Pilot/ Phase 1	RO8 Swedish pilot/ Phase II	RA1 Swedish pilot/ Phase II	Norwegian pilot Oslo/ Phase I	Polish pilot / Phase I and III	Polish pilot / Phase III	N/A	Polish pilot/ Phase I and III
Collected/ Created	Collect ed	Not yet received from partners	Created	Collected	Collected	Not collected yet	Collected	Collected and created	Collecte d	Collected	Collected
Name	Social media posts relevan t to the project' s objecti ves	HMI Questionnaire responses collected at pilot sites	Italian TMC's Operators questionnaire responses/Res ponse/Not defined if public or confidential	Questionnair e responses Anonymised	Questionnaire responses/ and interviews. Anonymised aggregated tables in an internal report on HMI.	Questionn aire responses/ and interviews. Anonymise d aggregated tables in an internal report on training.	Questionn aire responses/ Categorize d data from video recordings	Images and videos for disseminat ion purposes	Name, surnam e, email, phone number of pilot particip ants	Question naires' complete d during the pilots	Question naire response s (UPQ, UEQ, VDL, HMIQ, SUS, SATI) / Response restructu ring/ Anonymi sed aggregat ed tables in D5.8

D9.7: Revision of D9.4 Data Management Plan

Partner	INF	FhG/IAO USTUTT	SWARCO	Tuco	VTI	VTI	TOI	PZM	PZM	DEUSTO	PZM
Description	Posts in Twitter, Reddit and Youtube	Written answers in HMI questionnaire and if applicable, answer to interview questions, relevant to A3.2. Results will be included in anonymised and aggregated tables in D3.1.	Written answers to questions for the requirements of WP5, and results will be included in D5.1.	Written answers to questionnaire	Written answers to questions using the tools described in D5.3 for Phase II testing for RO8.	Written answers to questions using the tools described in D5.3 for Phase II testing for RA1.	Categorized & anonymised data & aggregated tables; to be included in D5.4.	Images and videos taken during the test-drives.	Registration details of phase III participants.	Analysis of data collected from each pilot of the project. Data is provided by pilots' leaders.	Written answers to printed questionnaires, results will be included in D5.8.
Category	Algorithm	Raw data and processed data from questionnaire	Raw data	Raw data	Raw data	Raw data	Partly raw data, partly categorized data	Dissemination material	Personal data	Raw data and processed data from questionnaires	Raw data
Standards	Related to algorithm/Python use	N/A	Not defined	N/A	N/A	N/A		N/A	N/A		N/A

D9.7: Revision of D9.4 Data Management Plan

Partner	INF	FhG/IAO USTUTT	SWARCO	Tuco	VTI	VTI	TOI	PZM	PZM	DEUSTO	PZM
Type	Source Code	Document with numbers, tables and charts, where applicable	Document	Document / Hard copy	User_profiledata.sav Afterdriving_questionnaires.sav	User_profiledata.sav Afterdriving_questionnaires.sav	Document	Video and images	Document	Document	Document
Format	Python files	doc/Excel	Doc/ Excel	Doc / Excel/ Paper	Netigate and SPSS	Netigate and SPSS	doc / Excel	MP4, jpg	Excel	doc / Excel	Excel
Size	3.1 GB	Not known yet	Max 5 MB	N/A	0.5 MB	0.5 MB	max 100 GB (in total)	5GB	unknown	N/known yet	Unknown
Owner	INF	Pilot partners (which exactly is TBD)	Swarco Mizar with contribution of IRU	TUCO	VTI	VTI	TOI	PZM	PZM	Pilots' leaders	PZM
Privacy level	Co	Co	Co	Co	Co	Co	Co	Public	Only for PZM	Co	Co
Repository during the project (for private/public access)	No repository	FhG-internal server	Cloud-based	Internal repository	Cloud based	Cloud based	Cloud based, Google doc	PZM storage	PZM storage	Cloud based, Google doc	PZM storage
Data sharing	R	R	R	R	R	R	restricted	O	R	R	R
Back-up frequency	Weekly	Daily	Once	Once	Daily	Daily	Monthly	Twice	Once	Daily	Twice
	3	Yes	N/A	N/A	3	3	3	No	Yes	No	Unnecessary
Duration of preservation	Until the end	1	Until the end of the project	Until the end of the project	3	3 years	1 year	Unlimited	1	1	Unlimited

D9.7: Revision of D9.4 Data Management Plan

Partner	INF	FhG/IAO USTUTT	SWARCO	Tuco	VTI	VTI	TOI	PZM	PZM	DEUSTO	PZM
on (in years)	of the project										
Repository after the project	TBD	TBD	Private cloud based and open access for Drive2theFuture partners	TBD	Partner storage	Partner storage	Cloud based	PZM storage	PZM storage	Cloud based	PZM storage
Comments	-	-	-	-	-	-	-	All visible persons signed an agreement /consent	Will be provided by participants themselves		No personal data

Table 3. The characteristics of data clusters (WP6-WP8)

Partner	VUB	Univ-EIFFEL	TOI	ACASA	IAM
WP/Activity	WP6/ A6.2	WP6/ A6.4	WP7 - all activities (A7.- A7.3)	WP8 / A8.1	WP8 / A8.3
Pilot site & Phase (if relevant)	-	-	-	Website subscribers	Survey to driving instructors
Collected/Created	Collected	Collected	Collected	Collected	Collected
Name	MAMCA weights and KPI prioritization. Aggregated data to be found in Figure 5 till Figure 34 and Table 50-51 in D6.1.	Questionnaire responses/ Response restructuring/ Anonymised aggregated tables.	A7.1-3 Expert interviews/Web survey responses from AV experts, anonymised in D7.1.	Subscribers to the newsletter	Driving Instructors (new drivers and specialized driving centers instructors)

D9.7: Revision of D9.4 Data Management Plan

Partner	VUB	Univ-EIFFEL	TOI	ACASA	IAM
Description	The data was collected through means of an online workshop for the requirements of WP6, A2 and the results are included in D6.1.	Written answers to questionnaires for the requirements of WP6.4, answer will be restructured, and aggregated results will be included in Zenodo.	Written answers to questions and results will be included in D7.1.	Volunteer subscription at the website to receive the project newsletters (name, email address, company).	Anonymous, but still collect email address.
Category	Raw data	Raw data	Raw data	Raw data	Raw data
Standards	N/A	N/A	N/A	N/A	N/A
Type	Images/ tables	Document	Document		
Format	Pictures (png type) / Excel	text document / spreadsheet	doc/excel/sav	wordpress Database	Survey Monkey
Size	15 MB	TBD	TBD	TBD	TBD
Owner	VUB	U of EIFFEL	TOI	ACASA	IAM
Privacy level	Co	Co	Co	Co	Co
Repository during the project (for private/public access)	Private cloud VUB sharepoint	Private cloud	Cloud based, QuenchTec, google doc	Wordpress Cloud based database	
Data sharing	R	R	R	R	R
Back-up frequency	Daily	Monthly	Monthly	Wordpress standards	TBD
	No	3	3	3	3
Duration of preservation (in years)	5	1	1	1	1
Repository after the project	Private cloud VUB sharepoint	Private cloud, open access (TBD)	Cloud based (QuenchTec)	TBD	TBD

3. Data controllers and processors

*NOTE: The information requested here is in line with the requirement to maintain data processing records under the GDPR and is **specific to personal data**. All data controllers and processors must also keep records of data set descriptions according to the latest Data Management Plan and DPIA. Where applicable, this information must be verified by the organizational Data Protection Officer. These documents should be treated by responsible persons as living, which are responsible to keep up to date and share them with the management team whenever they are requested. The current reporting of data controllers and processors in the project can be found below.*

I. Data controller's record of processing activities

1	Contact details of Data Controller
Email	<p>PZM: marzena.jougounoux@pzm.pl TUCO: Jonas@tuco.dk ACASA: Isabel.clos@racc.es / rssdrive2thefuture@gmail.com TUB: baris.cogan@tu-berlin.de VTI: anna.anund@vti.se</p>
Company address	<p>PZM: Polish Automobile and Motorcycle Federation (PZM), ul. Kazimierzowska 66, 02-518 Warszawa, Poland TUCO: Krogsbjergvej 2, Faaborg, 5600, Denmark ACASA: Diagonal, 687 08028 BARCELONA (SPAIN) TUB: Technische Universität Berlin, Straße des 17. Juni 135, 10623 Berlin VTI: Olaus Magnus väg 35, 581 95 Linköping Sweden</p>
Telephone	<p>PZM: +48 22 542 01 00 TUCO: +45 20143898 ACASA: +34 696996406 TUB: +4931422710 VTI: +46 709218287</p>
2	Purpose of processing
<p>PZM:</p> <ul style="list-style-type: none"> - AV awareness and acceptance analysis - Training needs analysis - HMI analysis - Dissemination - Logistics of the pilot (COVID-related and registration for test-drives) <p>TUCO: We do not handle any personal data. All surveys and questionnaires are anonymised. And nothing of personal value according to GDPR will be obtained and stored in any archive, physical data drive or a cloud system.</p> <p>ACASA: Allow the people interested in AV and in our project to subscribe to the project newsletter.</p> <p>TUB: Technische Universität Berlin (TUB) will be involved in pilot tests for optimizing a Human-Machine-Interface for a potential use in automated railway systems. This will be based on receiving user data collected on the test sites by the responsible pilot leader. The pilot test will identify needs and preferences of future users. The role of TUB is to analyse the data and draw conclusions to improve the HMI and to make suggestions for increasing user acceptance.</p> <p>VTI: The aim of the data collections was two-folded:</p>	

D9.7: Revision of D9.4 Data Management Plan

	<ul style="list-style-type: none"> ▪ Focus on bus-drivers and passengers' interaction with the aim to optimize the HMI for safety and trust during automated docking at bus stop. The applied method is Co-simulation using VR. Data is collected through surveys and interviews. ▪ Co-simulation between dispatcher and train drivers to develop a training tool version 2.0. Co-simulation using a train and a dispatcher simulator is used as method. Data is collected through surveys and interviews.
3	Description of categories of data subjects and of the categories of personal data
	<p>PZM:</p> <ul style="list-style-type: none"> - Anonymous data in questionnaires - Personal data in the form of image (photos and videos) - Personal data (name, surname, phone number, email) for registration <p>TUCO: We do not handle any personal data. All surveys and questionnaires are anonymised. And nothing of personal value according to GDPR will be obtained and stored in any archive, physical data drive or a cloud system.</p> <p>ACASA: Categories</p> <ol style="list-style-type: none"> 1. Whoever is interested on receiving the project information through the project newsletter. <u>Identifying data:</u> Name, Company, Email address. 2. Driving instructors or people who has answered the survey addressed to driving instructors. <u>Identifying data:</u> Email address (volunteer, only people that want to receive more information) <p>TUB: Subjective data will be collected on pilot sites via user surveys (online or pen-paper). Objective data (vehicle data on time, speed etc.) will be collected by pilot conductors or equipment used in pilot tests. The nature of data collected might include personal information such as age, gender, preferences, habits as well as emotional responses of individuals during the pilot tests.</p> <p>VTI: The subjects were normal persons without any special needs. The personal data was name, address and email in order to be able to administrate the payment of incentives.</p>
4	Categories of recipients to whom the personal data have been or will be disclosed including recipients in third countries or international organisations
	<p>PZM:</p> <ul style="list-style-type: none"> - Anonymous data in questionnaires will be transferred to DEUSTO, Spain for analysis - Images will be public online <p>TUCO: None.</p> <p>ACASA: Punctually could be asked by the project coordinator (CERTH).</p> <p>TUB: Data will not be shared with third parties that are not in the D2TF consortium.</p> <p>VTI: Individual data will not be disclosed.</p>
5	Where applicable, transfers of personal data to a third country or an international organisation, including the identification of that third country or international organisation
	<p>PZM: See above</p> <p>TUCO: None</p> <p>ACASA: Decision on Wordpress data will be made soon and will be reported as soon as it is available.</p> <p>VTI: N/A</p>
5	Where possible, the envisaged time limits for erasure of the different categories of data
	<p>PZM:</p> <ul style="list-style-type: none"> - Questionnaire data – unlimited - Images – unlimited - COVID-related – 2 weeks - Registration – 1 year <p>TUCO: No personal data are stored.</p> <p>ACASA: 1 year.</p>

D9.7: Revision of D9.4 Data Management Plan

<p>TUB: TUB only stores the data for as long as requires using in the research during the project, and for 3 months after the end of the project.</p> <p>VTI: Data, without keycode, will be stored at VTI for 3 years and destroyed after the project has ended.</p>	
6	Where possible, a general description of the technical and organisational security measures for
a	<p><i>the pseudonymisation and encryption of personal data;</i></p> <p>TUCO: We do not handle any personal data. All surveys and questionnaires are anonymised. And nothing of personal value according to GDPR will be obtained and stored in any archive, physical data drive or a cloud system.</p> <p>ACASA: N/A</p> <p>VTI: The participants receive a number when included. The key between name and number is only used for incentives. The key is stored by the controller on an area restricted with a password.</p>
b	<p><i>the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;</i></p> <p>TUCO: We do not handle any personal data.</p> <p>All surveys and questionnaires are anonymised. And nothing of personal value according to GDPR will be obtained and stored in any archive, physical data drive or a cloud system.</p> <p>ACASA: We use Wordpress as web design platform and collecting data.</p> <p>VTI: The data are limited. Only researcher with an invited access to the data can analyse it. However, no personal data will be available. The code stays locally with the controller.</p>
c	<p><i>the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;</i></p> <p>TUCO: We do not handle any personal data, as such we have nothing to restore.</p> <p>ACASA: We use Wordpress as web design platform and collecting data.</p> <p>VTI: The SharePoint solution and the controllers' computer is backup once each day.</p>
d	<p><i>a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing;</i></p> <p>TUCO: We do not currently have executed such processes internally at Tuco.</p> <p>ACASA: We use Wordpress as web design platform and collecting data.</p> <p>VTI: There is a special function at the Institute that handles this continuously.</p>

II. Data processor's record of processing activities

1	Contact details of Data Processor
Email	<p>TUB: baris.cogan@tu-berlin.de</p> <p>VTI: anna.anund@vti.se</p> <p>DEUSTO: unai.hernandez@deusto.es</p>
Company address	<p>TUB: Technische Universität Berlin, Straße des 17. Juni 135, 10623 Berlin</p> <p>VTI: Olaus Magnus väg 35, 581 95 Linköping Sweden.</p> <p>DEUSTO: Avenida de las Universidades 24, 48007 Bilbao, Spain</p>
Telephone	<p>TUB: +4931422710</p> <p>VTI: +46 709 218287</p> <p>DEUSTO: +34944139000</p>
2	Categories of processing carried out on behalf of the Controller

D9.7: Revision of D9.4 Data Management Plan

	<p>TUB: Data will be analysed based on the purpose of the project.</p> <p>VTI: Not needed.</p> <p>DEUSTO: Data to be processed will be provided by pilots' coordinators</p>
3	<p>Where applicable, transfers of personal data to a third country or an international organisation, including the identification of that third country or international organisation</p>
	<p>TUB: Data will not be shared with third parties that are not in the Drive2theFuture consortium.</p> <p>VTI: Not needed.</p> <p>DEUSTO: Data will not be shared with third parties that are not in the Drive2theFuture consortium.</p>
4	<p>Where possible, a general description of the technical and organisational security measures for</p>
a	<p><i>the pseudonymisation and encryption of personal data;</i></p> <p>TUB: Data will not be shared with third parties that are not in the Drive2theFuture consortium.</p> <p>VTI: See section I.</p> <p>DEUSTO: Data will not be shared with third parties that are not in the Drive2theFuture consortium.</p>
b	<p><i>the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;</i></p> <p>TUB: The collected data will always be stored safely and securely with an access only to appointed employees.</p> <p>VTI: See section I.</p> <p>DEUSTO: The collected data will always be stored safely and securely with an access only to appointed employees.</p>
c	<p><i>the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;</i></p> <p>TUB: The collected data will always be stored safely and securely with an access only to appointed employees.</p> <p>VTI: See section I.</p> <p>DEUSTO: The collected data will always be stored safely and securely with an access only to appointed employees.</p>
d	<p><i>a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing;</i></p> <p>TUB: Security of the storage and processing of data are ensured by the organization.</p> <p>VTI: See section I.</p> <p>DEUSTO: Security of the storage and processing of data are ensured by the organization.</p>

4. Data Privacy Impact Assessment (DPIA)

The Privacy Impact Assessment is required under Article 35 of the General Data Protection Regulation (EU) 2016/679. A DPIA is a process which helps assessing privacy risks to individuals in the collection, use and disclosure of information. DPIAs help identifying privacy risks, foresee problems and bringing forward solutions. DPIA is an evolving process in the project and a preliminary version was included in the D9.4. It is important to note that due to the close interrelation between WP5 and WP6, unless it is stated otherwise, what is reported as part of WP5 by partners applies also for WP6. The planning and conducting of the pilots are very closely related and intertwined.

D9.7: Revision of D9.4 Data Management Plan

4.1. Do I have to do a DPIA?

Determining if you need to do a PIA - screening questions

Answering yes to any of these questions indicates that a PIA is necessary.

Will the project involve the collection of new information about individuals?

WP2 (INF), WP5 (TUCO; SWARCO), WP6(Uni. Eiffel): No, WP8 (ACASA): No

Will the project compel individuals to provide information about themselves?

WP2 (INF), WP5 (TUCO SWARCO), WP6(Uni. Eiffel): No, WP8 (ACASA): No

Will information about individuals be disclosed to organisations or people who have not previously had routine access to the information?

WP2 (INF), WP5 (TUCO SWARCO), WP6(Uni. Eiffel): No, WP8 (ACASA): No

Are you using information about individuals for a purpose it is not currently used for, or in a way it is not currently used?

WP5 (TUCO; SWARCO), WP6(Uni. Eiffel): No, WP8 (ACASA): No, WP2 (INF): Maybe

Does the project involve you using new technology which might be perceived as being privacy intrusive? For example, the use of biometrics or facial recognition.

WP2 (INF): WP5 (TUCO; SWARCO), WP6(Uni. Eiffel): No, WP8 (ACASA): No

Will the project result in you making decisions or acting against individuals in ways which can have a significant impact on them?

WP2 (INF): WP5 (TUCO; SWARCO), WP6(Uni. Eiffel): No, WP8 (ACASA): No

Is the information about individuals of a kind particularly likely to raise privacy concerns or expectations? For example, health records, criminal records or other information that people would consider to be particularly private.

WP2 (INF): WP5 (TUCO; SWARCO), WP6(Uni. Eiffel): No, WP8 (ACASA): No

Will the project require you to contact individuals in ways which they may find intrusive?

WP2 (INF): WP5 (TUCO; SWARCO), WP6(Uni. Eiffel): No, WP8 (ACASA): No

However, some partners completed the whole DPIA, even if they answered 'No' in all of the above questions to ensure no risks can be created by combinations of data or metadata.

4.2. Step 1: Identify the need for a DPIA

Explain broadly what aims to achieve and what type of processing it involves. You may find it helpful to refer or link to other documents, such as relevant deliverables and other supportive documents that reside in SharePoint. Summarize why you identified the need for a DPIA.

WP2 (NTUA): *The project aims to create a simulation platform for simulating the behaviour of automated driving, estimating and predicting user acceptance upon autonomous vehicle functions, as well as conducting an impact assessment of autonomous driving on traffic, safety and energy. For simulating the behaviour of AVs, a microsimulation model will be developed based on the principles of the deep reinforcement learning and will be trained and validated using data from the RO2 pilot, led by FZI. Additionally, the results of the use cases applied in this pilot and the survey conducted will be used for formulating an AV profile accepted by the pilot participants (either in the virtual simulator or the real test track) and insert it in the simulation platform.*

D9.7: Revision of D9.4 Data Management Plan

WP2 (INF): The project involves the collection of social media posts relevant to autonomous mobility, and the analysis of such posts for determining the general public's sentiments and fears towards automated mobility. The system is designed by INF end-to-end. Such analysis aims to offer insight to service providers about the user acceptance of their tools.

WP5 (VTI): The data collections aim to achieve the following:

- An improved HMI to be able to have safe and trustable automated docking at bus stop from both a passenger and a bus driver perspective.
- An understanding of the roles when interacting between train drivers and dispatchers. Taking both perspectives into consideration.

WP5 (PZM): Pilot site RO-4 involves PZM collecting anonymous data from drivers and passengers (through questionnaires developed by DEUSTO in Google Forms, translated into Polish and printed) taking part in the tests in order to investigate the awareness and acceptance of automated cars (level 2 & 3) as well as associated with training needs in a country with a very high average age of a car. PZM will then translate the answers into English and pass on the data to DEUSTO in an Excel file for analysis. The only personal data processed in Phase I is image on pictures and videos acquired after receiving written agreement of participants and will be used mainly for dissemination purposes. Additionally, we were forced to gather COVID-statements from the participants, including their phone numbers. However, those are required for contact tracking by the local authorities due to the current pandemic and are to be destroyed after 2 weeks. In phase III, the above will remain the same, however, more personal data will be collected as PZM plans to organize a recruitment for the tests by email – to ask people, eager to participate in the test, to send an email to register. Potential participants will share their data with PZM but PZM will not transfer it any further. The purpose is to build a schedule of test-drives.

WP5 (TUB): TUB will be involved in pilot tests for optimizing a Human-Machine-Interface for a potential use in automated railway systems. This will be based on receiving user data collected on the test sites by the responsible pilot leader. The pilot test will identify needs and preferences of future users. The role of TUB is to analyse the data and draw conclusions to improve the HMI and to make suggestions for increasing user acceptance. The need for a DPIA arises from the collection of information on people's preferences, habits, and behaviours. Only the anonymized information will be used in the analysis.

WP5 (DEUSTO): University of Deusto (DEUSTO) will be involved in the Pilots results consolidation task. The objective of this Activity is to analyse the results of the Drive2theFuture pilots, find synergies with other projects and draw conclusions from the rest WP5 Activities. The need for a DPIA arises from the collection of information generated in all the pilots for its analysis. Only the anonymized information will be used in the analysis.

WP5 (SWARCO): The aim of the RO 7 Pilot is to assess the acceptance and operation capacity of the Traffic Management Centre (TMC) operators towards autonomous vehicles and mixed flows. This will be based on TMC operators' training schemes on Cooperative Intelligent Transport Systems (C-ITS) and automated functions that will be evaluated through questionnaires. The data to be collected will be related to the experiences, opinions, and relevant conclusions of the participants in terms of C-ITS systems and automated functions considering their roles in each one of the TMC's.

WP7 (TOI): WP7 will investigate sociocultural, ethical, safety, security and legal issues by literature analysis and expert interviews. The interview data will supplement the literature on these issues with the most updated observations and development as seen by main experts in the field. The data are collected through web surveys from (minimum 15) experts in each of the three activities (A7.1-A7.3). The questionnaires ask for active consent from the interviewees.

4.3. Step 2: Describe the processing

Describe the nature of the processing

WP2 (NTUA): NTUA will use and analyse data collected from surveys during the project real experiments as well as kinematic data (such as acceleration, speed, headway, etc) from the RO2 pilot, led by FZI. The project partners responsible for the pilots' realization and organization will collect the data taking into consideration that this data will be used as input from NTUA team to conduct its analysis. The data received will be completely anonymized as any personal information will be deleted. Therefore, no risk involved in any type of data processing is anticipated. The anonymized data may also be shared among the other activity participants - CERTH/HIT, IFSTTAR, TUM, FZI

WP2 (INF): The tool gathers social media posts by querying APIs from popular social media websites, namely Twitter, Reddit and Facebook. The data are analysed exclusively in house. No personal data are collected, only the social media posts body (actual textual content of the post) as well as the date and source.

WP5 (VTI): The data will be collected, used, stored and deleted in the following way.

- Data collected are user surveys focusing on user profile, trust, acceptance, and HMI (see D5.3). In addition, interview with bus drivers and passengers will take place. The data will be stored in Netigate and as an SPSS file (format .sav). The data will be backup once each day and will be stored at VTI for 3 years. There is no plan to share this data with other partners than DEUSTO that is responsible for the consolidated report. After the project ends the data will be deleted.
- Data collected are user surveys with focus on user profile, trust, acceptance, training effectiveness and HMI (see D5.3). In addition, interview with bus drivers and passengers will take place. The data will be stored in Netigate and as an SPSS file (format .sav). There is no plan to share this data with other partners than DEUSTO that is responsible for the consolidated report. After the project ends the data will be deleted.

WP5 (PZM): Non-personal data was collected in printed questionnaires. One of PZM workers will fill in the data from printed questionnaires into Google Forms, then we will extract the data into Excel, translate it into English and transfer to DEUSTO by email for analysis. Image of the participants was recorded during the test-drives by a third-party company which also prepared the video about the pilot. The videos are stored on the PZM computer. COVID-statements with phone numbers were collected on paper and destroyed after 2 weeks. Personal data for registration to phase III will be collected by email and stored in an Excel file on the PZM computer. Non-personal data was collected in printed questionnaires filled in by hand by the participants of the pilot. Image of the participants was recorded during the test-drives by a third-party company with several video and photo cameras Personal data for registration to phase III will be sent to PZM directly by the participants. DEUSTO will analyse non-personal data. Image of the participants shared with public after users' agreement. Personal data for registration to phase III will not be shared with anyone.

WP5 (TUB): The forms and questionnaires will be developed in cooperation with the responsible consortium partners. Subjective data will be collected on pilot sites via user surveys (online or pen-paper). Objective data (vehicle data on time, speed etc.) will be collected by pilot conductors or equipment used in pilot tests. The collected data will always be stored safely and securely with an access only to appointed employees. Information will be received and stored in a form that no inferences on individual participants are possible. TUB will not share the data with anyone outside the consortium unless this was agreed on.

WP5 (DEUSTO): Data consolidation will be carried out with the outputs of the project's pilots. This information will be sent to DEUSTO in Excel format. The collected data will always be stored safely and securely with an access only to appointed employees. Information will be received and stored in a form

D9.7: Revision of D9.4 Data Management Plan

that no inferences on individual participants are possible. DEUSTO will not share the data with anyone outside the consortium unless this was agreed on.

WP5 (SWARCO): *Data will be collected by TMC operators who will participate in the project's training scheme through questionnaires. The collected information (e.g., surveys, interviews, questionnaires) will be disclosed only with the project consortium members. No risks are anticipated in relation to data handling.*

WP7 (TOI): *The expert interviews are organised and sent out as a web survey by using the web survey programme QuenchTec. The data will be saved as excel- (.xls) and SPSS-files (.sav). The data will only be shared by researchers in WP7. The expert interviews will be reported anonymously. 2 years after the project is ended, the personal data will be deleted, as specified in agreement with the participants.*

Describe the scope of the processing

WP2 (NTUA): *The data processed belong in two categories. Firstly, kinematic and driving characteristics will be collected from the RO2 pilot such as speed, reaction times, lateral position, etc either from real test tracks or through the virtual simulator. The data will be provided to NTUA for the training and validation of the microscopic model that it will develop. The second category includes data collected from the questionnaires (either from RO2 pilot or other pilots) for analysing accepted behaviours of automated vehicles as well as the interaction between the driver and the vehicle. Both types of data will be provided to NTUA completely anonymized, i.e. NTUA will not have any personal information and thus it could not match any data with any person.*

WP2 (INF): *The data includes textual information from social media posts. The posts are openly available to the public and up to date more than 40000 individual posts have been analysed. The analysis includes determining the fears and opinions of the post using deep learning sentiment analysis algorithms. The data are not stored. Only their sentiment values and statistics are kept for demonstrating the tools performance and offering insight to service providers of autonomous mobility solutions. No geographical restrictions are put upon the data.*

WP5 (VTI): *The data is collected only once and at the time of the experiment. It consists of data from answering surveys using a webtool called Netigate. In addition, interviews with the bus drivers, the passengers, the train drivers, and the dispatchers will take place. Written notes will be saved from those. The data will be kept at VTI for 3 years and deleted at the end of the project. For study I in Phase II, in total 5 bus drivers and 15 passengers were included. For study II in Phase II, in total 10 train drivers and 10 dispatchers will be included. They are all from Sweden and the area around Linköping.*

WP5 (PZM): *No special category of data is collected from approximately 60 persons from Poland (first time in phase I and second time in phase III). Data will be stored until the work within Drive2theFuture is completed.*

WP5 (TUB): *The nature of data collected might include personal information such as age, gender, preferences, habits as well as emotional responses of individuals during the pilot tests. No special category or criminal offence data will be included. The data will be collected in two experiments with approximately 30 participants in total. TUB stored data only for the purposes of the project and keeps them for 3 months after the end of the project. Any data we store during this period will be treated in accordance with our privacy policy. The data collection will take place in Germany.*

WP5 (DEUSTO): *Data will be collected at the end of each pilot using the format defined in coordination with all the pilots' coordinators and supervised by Activity and WP leaders. The analysis of all data is*

D9.7: Revision of D9.4 Data Management Plan

anonymous and will not be shared with third parties out from the project consortium. Data will be processed internally in coordination with all the partners involved in the activity.

WP5 (SWARCO): *Each participant will answer 20-30 questions once at the end of the training scheme. Data will be kept until the end of the project. As the activity is under development, the number of participants has not been determined yet; however, participants will come from the geographical area of Rome, Italy.*

WP7 (TOI): *45-50 expert interviews will be conducted with international experts in the various topics in the three activities. The questionnaire is short, only around 10 questions for open-answer responses. The interviews will only be conducted once (M18).*

Describe the context of the processing

WP2 (NTUA): *NTUA is not responsible for collecting the data and therefore it does not have any relationship with the individuals participating in the pilots. Since NTUA will not be present in the data collection, it cannot ensure that the participants will be aware of the purpose the collected data will be used for. Additionally, the sample will also include Vulnerable Road Users (at least pedestrians will participate in the RO2 pilot).*

WP2 (INF): *The analysis is anonymous. The users voluntarily post in the social media services and their analysis adheres to the platform's as well as national and EU data protection regulations.*

WP5 (VTI): *There is no relationship with the participants, they were informed that they have the right to stop without giving a reason. They were informed about the reason of the study beforehand and signed an informed consent. Both studies are novel, both in terms of co-simulation using 2 VR systems and the real simulations for co-simulation between train drivers and dispatchers. There are no public issues to be concerned and no children or persons with special needs were included.*

WP5 (PZM): *The individuals are participants of the test-drives and they sign an agreement about the publication of their visual (for the photo and video). The process is clear to them and no children are involved. There are no public issues or special needs to be considered. In PZM we work with an external company on GDPR. Our internal lawyer always reviews agreements to be signed by participants of PZM events.*

WP5 (TUB): *The participation to the pilot test will be voluntary basis. The participants have the right to withdraw from the study at any point. All data will be deleted at their request. The participants will be informed about the study and the data collection and processing procedure. Children or other vulnerable groups will not be included. The data processing method used is common in the similar research areas. There are no known concerns over the processing or security flaws. TUB respects and follows the Federal Data Protection Act to protect the collection and use of personal data in research in Germany. Federal Data Protection Act (Bundesdatenschutzgesetz), 30.06.2017.*

English Version: https://www.gesetze-im-internet.de/englisch_bdsch/englisch_bdsch.pdf

WP5 (DEUSTO): *Data from all pilots will be processed to obtain information about the impact of autonomous driving in order to improve the user acceptance, find synergies with other projects and draw conclusions from the rest WP5 Activities. Then, main goal of this process is to collect, structure and analyze the pilot findings towards enhancing user awareness and acceptance for AVs, while investigating the commonalities and diversities per mode, user cluster and location.*

D9.7: Revision of D9.4 Data Management Plan

WP5 (SWARCO): Users will have no relation with the entity, and they will choose the questions they want to answer. There are no security, novelty or any other concerns related to users and the data to be collected.

WP7 (TOI): The experts are selected by the three activity leaders' familiarity and knowledge on who are the important experts in the fields. Participation is voluntary, and responses are based on their willingness to share their opinion.

Describe the purposes of the processing

WP2 (NTUA): The purpose of the processing is to train and validate the microscopic simulation model developed by NTUA that will mimic the behaviour of the automated vehicles and will be then integrated in the simulation platform for assessing the impacts of AVs on critical areas. Additionally, various scenarios will be simulated for investigating the impacts of AVs under different conditions in terms of road networks, traffic density, demand, traffic control, etc. Furthermore, the data from the use cases and the surveys implemented and conducted in the pilots will contribute to formulating an accepted behaviour as well as increasing the acceptance of the participants and users towards automated functions. To conclude, the aim of this processing is to raise public acceptance towards AVs.

WP2 (INF): This service aims at offering insight of the public's perception of autonomous vehicles. The benefit is mainly towards assessing current fears. This can help service providers in both their service design and marketing strategy in increasing the acceptance of their solution.

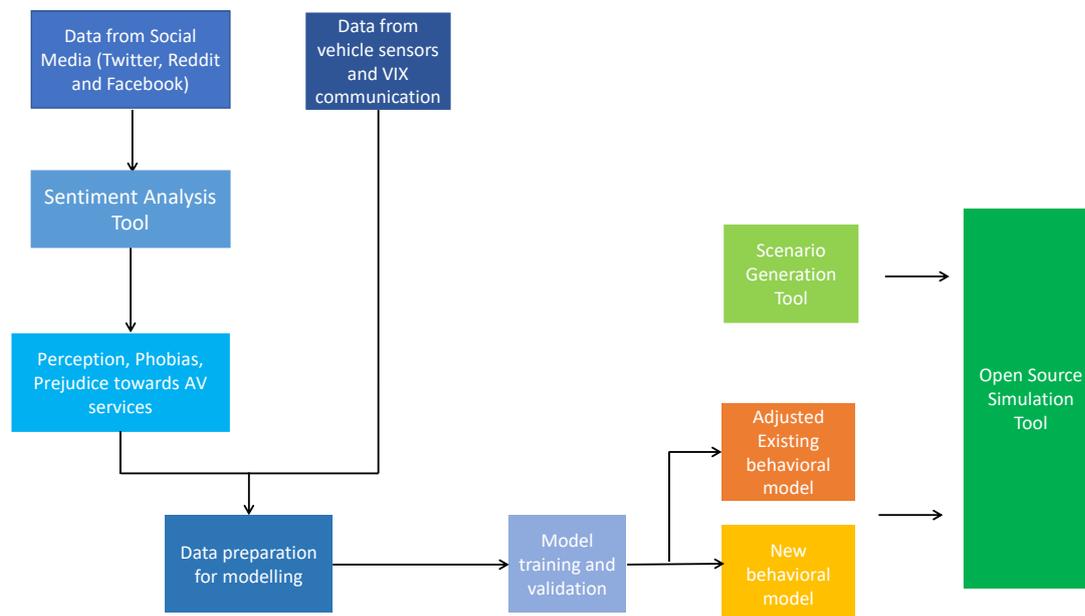


Figure 1. Data flow (WP2)

WP5 (VTI): For study 1 the purpose was to improve the HMI when docking a bus stop automatically. VR was used to minimize the risk, but also to visualize future scenarios for automation. For study 2 the purpose was to increase train drivers and dispatchers understanding of their roles and the task behind. No effect on the individual was expected. The participants received 30 Euro as a compensation for their participation.

WP5 (PZM): We want to investigate the awareness and acceptance of automated cars (level 2 & 3) as well as associated it with training needs in a country with a very high average vehicle age. Individuals, participating in the surveys are supposed to reflect on their experience with safety systems and autonomous functions of the new cars. They can share their opinion and thus further influence the

D9.7: Revision of D9.4 Data Management Plan

development of HMI and training for autonomous vehicles drivers/passengers and operators. Thanks to the processing we will know if there is any influence of the profile of the participant on his opinions related to the autonomous vehicles. The processed video will help us to present the project to a broader public.

WP5 (TUB): *The data collected from the participants in the pilot tests are essential to identify best practices and potential problems of the HMIs and to improve user acceptance. As proposed in the project framework, the main intend is to increase user awareness and acceptance of automated vehicles. Findings from the study will also promote future developments in designing human-centric HMI.*

WP5 (SWARCO): *Assess the acceptance and operation capacity of the Traffic Management Centre (TMC) operators towards autonomous vehicles and mixed flows. The conclusions to be drawn rely solely on this data collection and there is intended effect on users.*

WP7 (TOI): *The aim of the expert interviews in WP7 is to come up with the most recent and important issues and items related to the societal (sociocultural, ethical, safety, security and legal) impacts of the future adoption of AVs.*

4.4. Step 3: Consultation process

Consider how to consult with relevant stakeholders

WP2 (NTUA): *Same as stated in the first DPIA.*

WP2 (INF): *The mining and analysis is adhering to current EU regulations and the services the platforms terms and conditions. We do not foresee any issue with consulting other stakeholders.*

WP5 (VTI): *No consultation process was needed.*

WP5 (PZM): *At PZM I consulted our internal lawyer to have the right agreement for the use of visual on photos and videos – to be signed by the test-drive participants. Our internal PZM lawyer stays in contact with an external GDPR-consulting company (AUDYTEL S.A.).*

WP5 (TUB): *There is no need for a consultation from other stakeholders.*

WP5 (DEUSTO): *No consultation process was needed.*

WP5 (SWARCO): *Considering the activities planned during training, involvement of more people or experts is not required.*

WP7 (TOI): *The web surveys are sent out in M18. The questionnaire is developed in close cooperation between the activity leaders and the WP7 leader. No other researchers or experts are needed for selecting or analysing the data.*

4.5. Step 4: Assess necessity and proportionality

Describe compliance and proportionality measures, in particular

WP2 (NTUA): *The processing technique that will be followed is appropriate for training and validating data driven models for simulating driver behaviour as well as for formulating an accepted behavioural profile for AVs mimicking the human driving behaviour. The data quality necessary for a reliable result will be ensured by the pilot leaders. NTUA will only receive the data collected from them and it will be anonymized, i.e. any personal information will have been previously removed. NTUA will only process and analyse the*

D9.7: Revision of D9.4 Data Management Plan

data which will not be transferred or sent to any external body. The dataset will be kept within the project community.

WP2 (INF): *The General Data Protection Regulation (GDPR) foresees two provisions for “publicly accessible data”. Article 14 for Information obligations on Data Controllers for indirectly collected personal data and Article 9 for exceptions on the prohibition of the processing of special categories of personal data (herein known as “Sensitive Data”¹. For “manifestly made public data”, the Article 29 Working Party explained in WP2 for the Law Enforcement Directive (EU2016/680) that for Sensitive Data, publicly available sources must be interpreted as the data subject “was aware that the respective data will be publicly available which means to everyone including authorities”². The Data Protection Authorities go on to establish a difference between information found on an internet biography, press or public website and social networks. For the latter, the WP2 states “In such cases most of the users probably do not actively take notice [national police authorities have access to personal data] and are in fact not aware that their data are available to police authorities”³³.*

WP5 (VTI): *The data collected is stored by the supervisor and not shared. Using VR and simulations for visualizations are a safe and effective way to get different stakeholders’ perspective without putting them into risk. Those tools are also useful to visualize future possible technical solutions to make sure that future functions take the users’ perspective and need into consideration. The individuals were well informed before the decision to join. Information was sent home before hand and a verbal instruction was given during the arrival. Due to COVID 19 all instructions on how to use VR and simulation were video recorded to make sure the participants were informed and to minimize the physical contact between participant and supervisor. The participants signed informed consent after being informed they have the right to stop whenever they want without giving an explanation. The responsible processors in both studies have a PhD degree and they are experienced with data collections. All data collected were analysed and taken into consideration in further developing the HMI in study 1 and for the version 2.0 of the training for train drivers and dispatchers. To the best of our knowledge no international safeguard is needed.*

WP5 (PZM): *Participants consent, and all the other questions are answered in D9.4, WP6.5, WP5.3.*

WP5 (TUB): *Informed user consent will be collected to obtain freely given consent for data collection and processing. Analysing collected data is the basis for user-oriented research. This information cannot be obtained using other measures. Data will only be used for the declared purpose. The participation to the pilot test will be voluntary basis and the participants have the right to withdraw from the study at any point. All data will be deleted at their request. The participants will be informed about the study and the data collection and processing procedure.*

WP5 (DEUSTO): *The data collected is stored by the supervisor and not shared with third parties. All data used during activity A5.8 will be collected by all pilots’ coordinators following the rules and regulations about GDPR.*

WP5 (SWARCO): *There is no sensitive information required during training. However, if necessary, participants consent could be a possible lawful basis for processing. Considering the purpose of the training, to ensure data privacy, only the necessary, relevant, and adequate information will be acquired through the questionnaires. The information that will be presented during the training is in preparation.*

¹ Special categories of personal data defined as: “personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person’s sex life or sexual orientation”.

² The EU organ regrouping the European National Data Protection Authorities or “WP29”.

³ Article 29 Working Party, WP258, Opinion on some key issues of the Law Enforcement Directive (EU 2016/680), pg 10.

D9.7: Revision of D9.4 Data Management Plan

The collected information acquired through the questionnaires will be used only for project research purposes. Secure file exchange platforms will be used, if necessary, to share the information collected.

WP7 (TOI): In the questionnaire the interviewees are asked for active consent. We could have chosen a different data collection methodology, for instance a face-to-face or a phone interview. The COVID situation has hindered the first option. In choosing between a phone or a web survey interview, the latter was selected to make it easier to explain the aim and topics of the interview and to allow the participants to respond in their own time using their own words. This will ensure that the experts can provide the information they deem necessary ensuring high quality and detailed answers. Results will be reported anonymously.

4.6. Step 5: Identify and assess risks

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary.	Likelihood of harm	Severity of harm	Overall risk
WP2 (NTUA): Not enough data for training and validating the microscopic behavioural model.	Possible (due to COVID 19 situation)	Medium	Medium
WP5 (VTI): Study 1. In general, no major risk is foreseen since it is a VR study. <ul style="list-style-type: none"> The risk that might happen is motion sickness. The risk of COVID exposure. 	Remote	Minimal	Low
WP5 (PZM): All the other questions are answered in D9.4, WP6.7, WP5.3.	Remote	Minimal	Low
WP5 (TUB): Leakage of user data during data transfer to consortium partners. Impact: Anonymized user data will be leaked. Data transfer and storage will be conducted using secure file exchange platforms.	Remote	High	Low
WP5 (SWARCO): No risks identified as medium or high risk.	Remote	Low	Low
WP7 (TOI): The recognition of the individual expert and his/her opinion/statement.	Remote	High	Low

4.7. Step 6: Identify measures to reduce risk

For each of the identified risks, the following measures were selected.

Identify additional measures you could take to reduce or eliminate risks identified as medium or high risk in step 5				
Risk	Options to reduce or eliminate risk (risks not related to data privacy and assessment is not thus relevant)	Effect on risk [eliminated; reduced; accepted]	Residual risk [low; medium; high]	Measure approved [Yes/No]

D9.7: Revision of D9.4 Data Management Plan

1, 2, 3, 4, 5	<p>WP2 (NTUA): Train and validate the model with the amount of data provided.</p> <p>WP5 (VTI): They could easily remove the VR system and extra efforts was put on how to put the system on/off by using video instructions.</p> <p>WP5 (VTI): The risk of getting COVID was handled by no contact with participants, gloves when using the hand controls and extra cleaning. In addition, all instructions were made in a big room, only 2 participants could join at the same time, and all equipment were cleaned carefully after each session.</p> <p>WP5 (TUB and SWARCO): No risks identified as medium or high risk.</p> <p>WP7 (TOI): Anonymise the statements properly</p>	<p>Reduced</p> <p>Reduced</p> <p>Reduced</p> <p>Reduced</p> <p>Reduced</p>	<p>Low</p> <p>Low</p> <p>Low</p> <p>Low</p> <p>Low</p>	<p>No need</p> <p>No need</p> <p>No need</p> <p>No need</p> <p>No need</p>
---------------------	---	--	--	--

4.8. Step 7: Sign off and record outcomes

This table was completed only for partners who seek the advice of their DPOs and this is reported below.

Item	Name/date	Notes
Measures approved by:	<p>WP5 (VTI): Anna Anund</p> <p>WP7 (TOI): Trine Dale</p>	Integrate actions back into project plan, with date and responsibility for completion
Residual risks approved by:	WP5 (VTI): Not relevant – no risks are identified	If accepting any residual high risk, consult the ICO before going ahead
DPO advice provided:	WP5 (VTI): Yes	DPO should advise on compliance, step 6 measures and whether processing can proceed
Summary of DPO advice: WP5 (VTI): Ms Louise Dahlgren Be aware of the risk to store information in SharePoint. It might be unlawful due to Schrems (C-311/18, Data Protection Commissioner v. Facebook Ireland Limited, Maximilian Schrems). At this stage, I do not see a need to make further comments.		
This DPIA will kept under review by:	<p>WP5 (VTI): Anna Anund</p> <p>WP7 (TOI): Trine Dale</p>	The DPO should also review ongoing compliance with DPIA

Overall: CERTH is responsible for integrating back into the plan and any future data privacy related communication. For implementing the solutions, depends on who has developed the corresponding part.

Action to be taken	Date for completion of actions	Responsibility for action
No action is required at this point	26/10/20	CERTH as Data Management Plan responsible

5. Conclusion and next steps

This version of the Data Management Plan addresses the data clusters, second version of the data privacy impact assessment, and data management notes in relation to change from Dropbox to Google drive as project storage.

Post-processed datasets, free from any private/personal and identifiable information, will reside in the Drive2theFuture partners safe repositories. As the COVID situation has affected the pilots' timeline, not all analytic descriptions of the complete and (non) shareable datasets that will be created during the pilots (WP5), the analyses to follow (WP5 and WP6) and other WPs that require post-processing (e.g. for the preparation of the Use Cases in WP1 and the simulation modelling conducted in WP2) are included in the data clusters table and thus no decision on shareable datasets can be made if all pilots have not been conducted.

In conclusion, this Deliverable is a living document, and it will be regularly updated during the lifetime of the project. Updates will be necessary and will be included in the final version of the Project Management Plan (D9.8; M35) that will present the final data clusters' table, the selected datasets to be uploaded in Zenodo as well as the open-access publications. If another DPIA is required, it will be performed and reported in the final project report.

References

1. The EU General Data Protection Regulation (GDPR) website (<https://eugdpr.org/>). Last accessed: 17/10/19.
2. Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020(http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf)
3. Panou, M., Gaitanidou, E. (2019). D10.2 H- Ethics Requirement no. 4. Needs, wants and behaviour of “Drivers” and automated vehicles users today and into the future (Drive2theFuture), Grant Agreement no. 815001.
4. Panou, M., Gaitanidou, E. (2019). D9.4 Data Management Plan. Needs, wants and behaviour of “Drivers” and automated vehicles users today and into the future (Drive2theFuture), Grant Agreement no. 815001.