SHared automation Operating models for Worldwide adoption

SHOW

Grant Agreement Number: 875530

D14.2 Data Management Plan (DMP) – 1st version

This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation programme under Grant Agreement number 875530

Legal Disclaimer

This report is subject to a disclaimer and copyright. This report has been carried out under a contract awarded by the European Commission, contract number: 875530. The content of this publication is the sole responsibility of the SHOW project.

Executive Summary

The Data Management Plan (DMP) has been developed as required by the implementation of a Horizon 2020 limited pilot action on open access to research data.

This deliverable is the first version of the DMP elaborated within the SHOW project framework, referring to what data the project will generate, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved.

The SHOW Data Management Plan refers to the latest EC DMP guidelines. Data that will be produced within the project lifecycle shall be Findable, Accessible, Interoperable and Reusable (FAIR). It describes what kind of data will be collected, processed or generated and following what methodology and standards, whether and how this data will be shared and/or made open, and how it will be curated and preserved

The SHOW Data Management Plan is a living document and as such will follow and reflect the evolution in the form of dataset updates and/or respective changes in Consortium policies. It should be noted that this first version of the DMP, submitted by M6, will be continuously updated and it will be followed by the second and final version of the DMP (D14.3) by Month 24 of the project.

Document Control Sheet

Start date of project:	01 January 2020
Duration:	48 months
SHOW Del. ID & Title:	D14.2 Data Management Plan (DMP) – 1st version
Dissemination level:	PU (for Public)
Relevant Activities:	A14.6: Data Management - ERTICO
Work package:	WP14
Lead authors:	Nikolaos Tsampieris (ERTICO)
Other authors involved:	Maria Gemou, Matina Loukea (CERTH/HIT)
Internal Reviewers:	ICCS, TECNALIA.
External Reviewers:	N/A
Actual submission date:	30/06/2020
Status:	SUBMITTED
File Name:	SHOW_D14.2 _DMP first version_SUBMITTED

Document Revision History

Version	Date	Reason	Editor
0.1	30/04/2020	Table of Contents for feedback	Nikolaos Tsampieris
1.0	15/06/2020	Version sent for internal peer review	Nikolaos Tsampieris
2.0	30/06/2020	Final version sent for submission	Nikolaos Tsampieris

Table of Contents

List of Tables

List of Figures

Abbreviation List

Abbreviation	Definition
AV	Autonomous Vehicle
C-ITS	Collaborative ITS
CCAV	Cooperative Connected Automated Vehicle
D	Deliverable
D3WC	International Boundary & Water Commission
DCAT-AP	Data Catalogue Application Profile
DMP	Data Management Plan
DOI	Digital Object Identifier
DPIA	Data Privacy Impact Assessment
DPO	Data Protection Officer
EC	European Commission
EN	European Standard, telecommunications series
ETA	Estimated Time of Arrival
ETSI	European Telecommunications Standards Institute
EU	European Union
FAIR	Findable, Accessible, Interoperable, and Reusable
GDPR	General Data Protection Regulation
GPS	Global Positioning System
IoT	Internet of Things
IPR	Intellectual Property Rights
ISA	Interoperability solutions for public administrations, businesses and citizens

Abbreviation	Definition
AV	Autonomous Vehicle
MaaS	Mobility-as-a-Service
OA	Open Access
OEM	Original Equipment Manufacturer
ORDP	Open Research Data Pilot
POPD	Protection Of Personal Data
RT	Reaction Time
SME	Small and Medium Size enterprise
VEC	Vulnerable to Exclusion
W3C	World Wide Web Consortium
WOT	Web of Things
WP	Work Package

1 Introduction

1.1 Purpose of the document

The purpose of the deliverable -*D14.2-Data Management Plan*- 1st version (DMP) is to provide an analysis of the main elements of the data management policy that will be used by the consortium, concerning all the datasets that will be generated and/or collected during the project.

SHOW is a H2020 funded project, and within this framework, the DMP must at least cover the specific aspects about the project's datasets. Particularly, in accordance with the EC DMP guidelines [1], it describes the data management life cycle for all datasets to be collected, processed or generated by a research project. It must cover:

- the treatment of research data during the project lifecycle and beyond;
- which data to be collected, processed and/or generated;
- which methodology and standards will be applied;
- if will data be shared or it will be open access;
- how data will be curated and preserved.

The DMP plays a crucial role in the project's success because it ensures the availability and the quality of the datasets, which will be generated/used within the Project framework, serving the achievement of the project's objectives.

The SHOW Data Management Plan refers to the latest EC DMP guidelines.
This version has explicit recommendations for full life cycle management through the implementation of the FAIR principles, which state that the data produced shall be Findable, Accessible, Interoperable, and Reusable (FAIR).

It should be noted that the DMP – 1st version, is a Deliverable on Month 6; however, it will be a live document and it will evolve during the project lifespan according to the progress of project activities, capable to capture and reflect the evolution in the form of dataset updates and/or changes in Consortium policies to be reflected in *Deliverable D14.3 DMP – final version* on Month 24.

1.2 Intended Audience

The SHOW project addresses highly innovative concepts. As such, foreseen intended audience, due to the strong expected impact of the project on their respective domains, is the project authorities (Cities, Municipalities, Ministries) and policymakers, OEM's and transport operators, Tier 1 suppliers, telecom operators and technology providers, SME's operating in the area, road operators, passengers as well as all road users encompassing VEC (Vulnerable to Exclusion) citizens (through their associations).

Moreover, a further intended audience is the scientific community (research and academia) with an emphasis in the areas of intelligent transport systems, road safety, and automotive engineering, and standardisation organisations.

1.3 Interrelations

The DMP is part of WP14 and is closely linked to data collection and processing activities, such as WP5 Big Data collection, processing, and analytics, the pilots, where the primary data sources will come for the pilot related activities (WP9-WP14), taking into consideration the data clusters and specifications that will be defined in WP9, the data relating processes for treating data (collecting, storing, sharing, analysing, and reporting data) in WP11, 12, 13, and 14, but also when only data sharing takes place to perform other activities (e.g. simulations in WP10). Some

partners will be only processors (e.g. those involved in WP10 – simulations), so their role is dependent on the data collection taking place in other activities (e.g. WP1, WP11, WP12).

2 Data Summary in SHOW

2.1 Overview

This section aims to:

- provide a first categorisation of the data;
- identify a list of the data types that will be generated;
- provide a list of metadata that will be used to describe the generated data and enable their reuse;
- provide recommendations on data collection and sharing processes during the project and beyond.

The project involves data collection (in the context of the piloting and validation phase) and a set of SHOW demonstrations to assess the technologies and effectiveness of the proposed solutions in real-life conditions. It should be noted here, that due to the fact that the project will collect (partially) personal-related data, the Consortium will comply with any related European and national legislation and Directives relevant to the country where the data collections are taking place; this will be dealt in detail in Deliverables D3.2, D3.4 and D3.5 Ethics Section 4-5.

The SHOW project will collect a large amount of raw data including – among others – real-time GPS-data, vehicle sensor data, traffic plans, routes, occupancy levels, RT delays, ETAs, road infrastructure data etc.

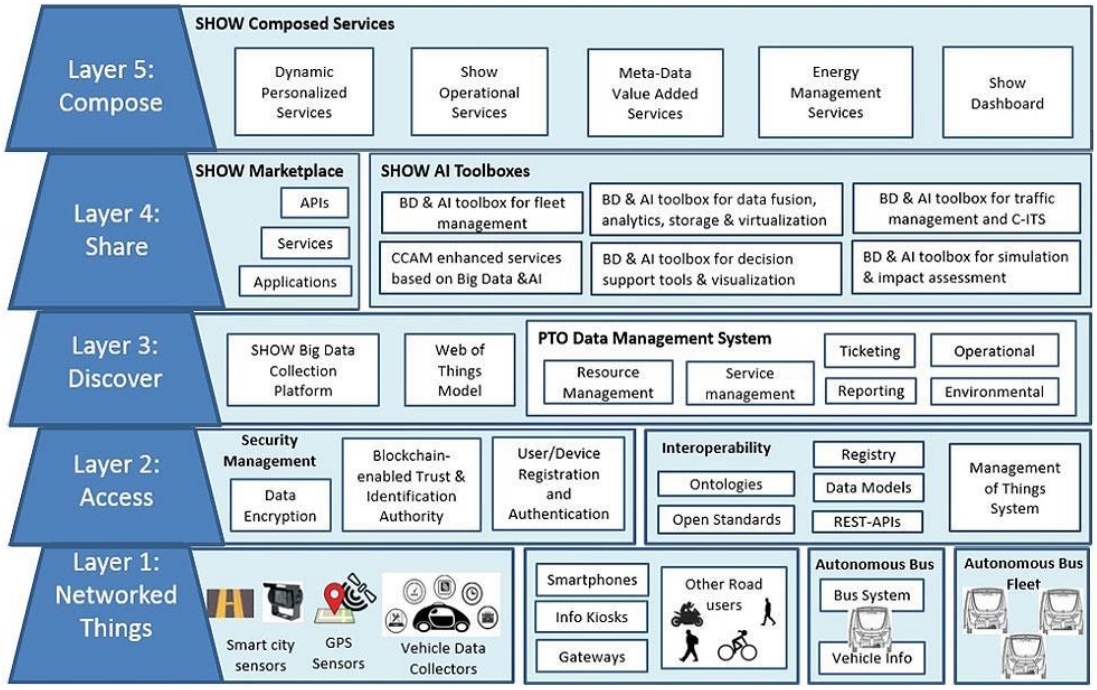

Figure 1: SHOW Architecture

SHOW will devise a reference architecture for AVs by providing an interoperable environment to boost innovative connected vehicle-based business and services to allow communication and cooperation between vehicles, infrastructure and with other road users and to enable automated, smart mobility services, innovative fleet management concepts and higher performance of automated vehicle functions. Thus, at the architecture level, SHOW will improve the current conceptualisation of the W3C

new incoming global standard Web of Things layer architecture utilizing the best practices of both worlds (Web and IoT) and providing several functionalities needed for enabling the domain of the connected vehicles by adopting the current W3C WOT layer architecture that includes the following 4 layers: **Access** (provide access to a thing), **Find** (discover things), **Share** (enables things exchange) and **Compose** (enables Things creation from existing ones).

Focusing on the lower Layer 1 Networked Things (Compose), there is a wide collection of heterogeneous sensors and edge IoT devices which include but are not limited to GPS and smart city sensors, data collectors from conventional vehicles, user's mobile devices (and wearables), info kiosks, gateways, autonomous buses Geo-spatial and other fleet/passenger/traffic situation/service real-time information)

Moreover, subjective data will be collected through all types of qualitative surveys, questionnaires, personal and group interviews (WP1) that take place in the project. The data may be collected from users/customers, data providers and operators as well as other types of stakeholders. In all cases, the data is/will be anonymised to ensure privacy and protection of the participants' identity (as in D3.2, D3.4). In the case of the users/customers, subjective data will mostly deal with travel preferences and needs. Additional data sources will be used and are required for the development of tools and new services, exploiting Big Data and AI, for the data management platform (Table 1).

Table 1: Data sources and services used in SHOW

Data sources	Services
Event data (event type, start time, location and other attributes), mobility data (taxi, buses, autonomous cars location, availability, delays, traffic congestion, etc.)	Demand prediction accounting for the impact of special events for travellers, traffic management centres, public transport service providers, cities authority, event managers, etc.
Bus operation data (location, delays, routes, demand, etc.)	Bus arrival-time / travel-time prediction for travellers, public transport operation centres, traffic management centres/operators.
Fleet location data, origin-destination demand data (e.g. passenger request through apps or MaaS platforms.), traffic congestion data (e.g. floating car data)	Predictive routing for a fleet of autonomous vehicles for OEMs, drivers (SAE 3), travellers, navigation service providers, fleet operators, traffic management operators
Telematics, floating car data and external (historical) sources (for example meteorological data)	Fleet management for professional fleet operators, freight forwarders and relevant technology providers
Data from conventional equipment such as loops, cameras and radars along with innovative ones such as probe data and data from connected vehicles. CCAV can be an important source of data, since they can assist in getting a better view of the entire urban environment	Traffic management and C-ITS for traffic managers, traffic control centres and relevant technology providers
Floating car data, Probe data, Social media data, Bluetooth detections	Mobility patterns identification and prediction for mobility service providers, public authorities and public transport operators

Data sources	Services
Traffic flows, vehicle counts, speed measurements, time/speed profiles, mobility patterns, traffic status etc.	Decision support tools for public administration officials, public transport operators, mobility service providers

The above-mentioned data will be collected and managed through the Big Data Platform and Data Management Portal, where all relevant internal and external data assets will be continually collected and registered, with their appropriate metadata according to relevant EC standards, such as DCAT-AP (D3WC initiative) and any extensions that may be done to it in the context of the ISA2 programme.

2.2 Datatypes, Datasets and Dataset Categories

This section presents a description of the data that will be collected, generated and managed in SHOW. More specifically, data is clustered into three different sections: data types, datasets and datasets categories. Data types are related to the source of the data, i.e. vehicle data, traffic data, etc., datasets refer to the file extension of the data and, finally, dataset categories refer to the level of process that the data has undergone.

The types of data generated, collected and managed within SHOW fall into the following categories:

- Subjective data (user profile and user request related data)
- Vehicle related data
- Urban environment related data (traffic and city related data)
- Infrastructure related data

As far as datasets are concerned, the following will be handled by SHOW:

- Reports (in the form of word or pdf documents)
- Excel files with raw data, as received by sensors, surveys etc.
- Video signals
- Database

The following table is a template will be used to describe the different datasets, including their reference, file format, standards, methodologies and metadata and repository to be used.

Table 2: Dataset Description Template

Dataset Reference	SHOW_WPX_TX.X_XX Each dataset will have a reference that will be generated by the combination of the name of the project, the Work Package and Task in which it is generated
Dataset Name	Name of the dataset
Dataset Description	Each dataset will have a full data description explaining the data provenance, origin and usefulness. Reference may be made to existing data that could be reused.
Standards and metadata	The metadata attributes list The used methodologies
File format	All the format that defines data
Data Sharing	Explanation of the sharing policies related to the dataset between the next options: **Open**: Open for public disposal **Embargo**: 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**: Only for project internal use. Each dataset 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 taken into account.
Archiving and Preservation	The preservation guarantee and the data storage during and after the project (for example : databases, institutional repositories, public repositories …)

The SHOW project will produce different categories of datasets:

- Context data: data that describe the context of an experiment.
- Acquired and derived data: data that contain all the collected information related to an experiment.
- Subjective data: questionnaires, surveys, personal and group interviews.
- Raw/unprocessed data: data collected directly from the source (either objective or subjective).
- Metadata: descriptions of data that will facilitate the data analysis and data pre-processing.
- Aggregated data: data summary obtained by reduction of acquired data and generally used for data analysis.
- Consolidated data: data collected across sites and per data type.

As the nature and extent of these datasets can evolve during the project, more detailed descriptions will be provided in the M24 version of the DMP. It is our understanding that the DMP will be enriched and updated throughout the project lifecycle.

A data collection template, with required information per data cluster (e.g. survey, questionnaires, vehicle, automated, purely formative) has been circulated in a first attempt to identify the higher-level clusters and the information needed for them.

In detail, the template collects information about the following:

- The name of the data
- Whether the data was collected or created
- Data description
- Data category
- Data type
- Data format
- Data size
- Data Ownership
- Privacy level
- Data repository during the project (for private/public access)
- Data sharing
- Back-up frequency
- Status of data at the end of the project (destroyed or not)
- The duration of the data preservation (in years)
- Data repository after the project is complete

Table 3 shows a sample example of the aforementioned data collection template.

Table 3: Data collection template

Data clustering template														
Collected/ Created	**Name**	**Description**	**Category**	**Type**	**Format**	**Size**	**Owner**	**Privacy level**	**Repository during the project (for private/public access)**	**Data sharing**	**Back-up frequency**	**Destroyed at the end of the project?**	**Duration of preservation (in years)**	**Repository after the project**
Collected/ created	Name of the data/ metadata/ exploitable result	Description of the data/ metadata - Please provide the description of the information you will collect/create (example: is it information about mobility habits from pilot users? Is it information related to vehicles that are used in the pilots? What is the source of the information?	what category does the information match? FW/SW/ Algorithm/Raw data/ Dissemination material/etc.	What is the type of information? Document/video/images/Source code/etc.	File extention/ prototype (docx, .txt., etc)	size in MB/GB	Partner name/ Consortium/ external stakeholder	Public/ consortium/ partner/ etc.	BAL.PM or other Open access repository/ partner storage(private cloud/private drop box)/etc.	**Open:** Open for public disposal **Embargo:** 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:** 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 taken into account.	How often will you make backups? daily/ monthly / yearly/once	NO (1)/No (2)/NO (3)/ Yes / Unnecessary	number of years	Open access repository/ partner storage (private cloud/private drop box), etc.

Data clustering template														
Collected/ Created	**Name**	**Description**	**Category**	**Type**	**Format**	**Size**	**Owner**	**Privacy level**	**Repository during the project (for private/public access)**	**Data sharing**	**Back-up frequency**	**Destroyed at the end of the project?**	**Duration of preservation (in years)**	**Repository after the project**
Collected		Vehicle-kilometres driven	Raw Data		File extension		Navya	partner	partner storage	Restricted	Daily	No	Undefined	Undefined
Collected		vehicle-kilometres automatic driven	Raw Data		File extension		Navya	partner	partner storage	Restricted	Daily	No	Undefined	Undefined
Collected		locational losses signal numbers	Raw Data		File extension		Navya	partner	partner storage	Restricted	Daily	No	Undefined	Undefined
Collected		trajectory's mistakes	Raw Data		File extension		Navya	partner	partner storage	Restricted	Daily	No	Undefined	Undefined

The information enables us to identify the types of data and determine how to manage it within the Data Management Plan and as instructed by the General Data Protection Regulation (GDPR) principles.

Personal-related data will be centrally stored in an anonymised and secure standards-abiding way and in accordance with the GDPR [Regulation (EU) 2016/679 of the European Parliament]. Also, data will be scrambled where possible and abstracted in a way that will not affect the final project outcome. Hence, it becomes apparent that the majority of research data generated by the project will be made open & will be offered to the Open Research Data Pilot, in which SHOW has declared its intention to participate (Section 3). Also, the allocation of specific roles in the Consortium in compliance with the GDPR requirements, briefly touched upon for completeness in Section 5, will be treated in detail in Deliverables D3.2 and D3.5 – Ethics. The type and format of data upon the FAIR principles (Section 3) their way/means of collection (against enhanced GDPR compliant templates) and the open access layer in each category will be also be addressed in the Deliverable D3.2 where the Data Protection Policy is included, and in D3.4 and D3.5 where additionally, the Data Privacy Policy will be included.

3 FAIR Data Principles

The data generated within the project framework should be Findable, Accessible, Interoperable and Reusable (FAIR). These qualities neither affect implementation choices nor imposing specific technologies, standards, or implementations. FAIR data principles, therefore, should not be viewed as a standard but rather as a framework to follow, when designing a Data Management Plan, since they ensure that the most important components for lifecycle data management are covered.

3.1 Making Data Findable

Data to be made openly available in SHOW, according to the principle of making data findable, special care should be applied to prevent storage of data for open access, unless appropriate metadata will be provided.

Internal and external data will be continually collected and registered, with their appropriate metadata according to relevant EC standards, such as DCAT-AP (D3WC initiative) and any extensions that may be done to it in the context of the ISA2 programme, thus making data visible in a searchable context based on the semantics of data.

3.2 Making data openly Accessible

Data in SHOW, will be made available with open access. This will not apply to the datasets that will be deemed by the Consortium as IPR protected such as proprietary data owned by consortium beneficiaries or third legal parties, as outlined in the DoA, and also in Deliverables D3.2 and D3.5 - Ethics manual & Data Protection Policy and Data Privacy Impact Assessment (DPIA) and the project Consortium Agreement. Especially for generated data or for data declared as background assets, access will be granted only to authorised users (Appendix I – ref Article 29.3).

The SHOW's open data portal, where all relevant internal and external data assets will be continually collected and registered, will be the common means for exchanging data either among the project partners or between the consortium and third parties.

All datasets will be available for sharing and re-use via the Data Management Portal of the project.

The specifications of how and which the portal will be used in the context of data sharing and reuse will be defined in the next months (up to M9) by the consortium and will be included in the final version of the DMP. These specifications have to be referred to any new features that have to be added to the current portal infrastructure for dealing with the different data manipulation needs.

All project datasets will be stored in a private cloud-based repository, e.g. ZENODO (https://zenodo.org/) a free service developed by CERN under the EU FP7 project OpenAIREplus (grant agreement no.283595). The repository shall also include information regarding the software, tools and instruments that were used by the dataset creator(s) so that secondary data users can access and then validate the results. The SHOW data collection will be accessed in ZENODO repository in a similar address as the following link: https://zenodo.org/collection/<<show>>

The datasets in the cloud repository will be linked to the management/exploitation portals of the project. They will be assigned to DOIs in order for third parties to be able to access them. Through the use of the above repository (or similar, e.g. own-Cloud), we will ensure that the most up to date security features will be applied out of the box, e.g., firewall, password protection, encryption, etc. If any IPR issue exist on sharing,

they will be handed accordingly. Moreover, all necessary material that supplements each dataset (software for parsing the datasets, standards documents, etc.) will be provided by the consortium via the data management portal.

3.3 Making data Interoperable

Data assets collected in the Data management portal, are accompanied with their appropriate metadata according to relevant EC standards, such as DCAT-AP (D3WC initiative) and any extensions that may be done to it in the context of the ISA2 programme. This application profile is a specification for metadata records to meet the specific application needs of data portals in Europe while providing semantic interoperability with other applications on the basis of reuse of established controlled vocabularies (e.g. EuroVoc) and mappings to existing metadata vocabularies (e.g. Dublin Core, etc.). The SHOW data management portal will be based on existing open source technology, so as to maximize interoperability with open data federators and to facilitate deployment and will become a relevant asset inside the project as well as to those outsider institutions that are willing to reuse the data assets gathered and produced by the project.

Data interoperability is foreseen in the project through conformance to standards Specifically, SHOW data will conform to a number of standards (see also Appendix II), including the following:

- ETSI EN 302 637-2 [2]
- ETSI EN 302 637-3 [3]
- DATEX II, TPEG2 [4,5]

3.4 Increase data Re-use

SHOW participates in the Pilot on Open Research Data launched by the European Commission along with the Horizon2020 Programme. As such, all data produced by the project are to be published with open access. Particularly, access will be given through Creative Commons CC0 license for all datasets at this project stage, unless it is defined otherwise.

All datasets will be maintained for the entire duration of the project as well as for 2 additional years after its conclusion. After the project ends, all datasets will be stored in a centralized facility to minimize maintenance costs. Datasets with acknowledged long-term value may be kept for a longer period. The long-term value of a dataset will be decided according to the exploitation plan as well as its relation to a scientific publication.

For the period in which the data will be available for open access, no restrictions will be imposed on their access. In case IPR protected or proprietary data will be generated or made available as background assets in the context of the project by one or more beneficiaries, access to this data will be treated in accordance with the rules and regulations foreseen in the project Consortium Agreement. For all this data, access will be granted only to authorised users.

All data to be available as open data will undergo a quality assurance process.

At a first level, each partner in charge follows specific procedures to assure the quality of the data and conformance to standards as referred to each dataset's description in Section 2. Those procedures regard calibration of the measurements, as well as, internal post-measurement review. On a higher level, the quality will be also monitored periodically on the different versions of the dataset as conservation experts of the consortium will review each update. Finally, additional peer-reviews will take place in case of publication in Journals.

4 SHOW Open Strategy & Participation in the Open Research Data Pilot

The SHOW project has agreed to participate in the Pilot on Open Research Data in Horizon 2020 and uses the specific Horizon 2020 guidelines associated with 'open' access to ensure that the results of the project results provide the greatest impact possible. SHOW will ensure open access to all peer-reviewed scientific publications relating to its results and will provide access to the research data needed to validate the results presented in deposited scientific publications.

The following lists the minimum fields of metadata that should come with a SHOW project-generated scientific publication in the Information HUB section of the project website https://show-project.eu/

- The terms: "European Union (EU)", "Horizon 2020"
- Name of the action (Research and Innovation Action)
- Acronym and grant number (SHOW, 875530)
- Publication date
- Publication type
- Publication place
- Publication name
- Publication authors
- Length of embargo period if applicable
- Persistent identifier

Apart from the scientific publications index, SHOW will publish all its public Deliverables and dissemination material in its web site (under "Library"). When referencing Open access data, SHOW will include at a minimum the following statement demonstrating EU support (with relevant in-formation included into the repository metadata):

"This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 875530".

Finally, SHOW will target **Open Access journals**, when applicable, whereas it will also target Gold OA publications and Green OA, wherever Gold OA is not possible. The target is to maximize the impact on scientific excellence through result publication in open access yet highly appreciated journals, without releasing any confidential information that could potentially violate the security nature of the project.

The SHOW consortium will strive to make many of the collected datasets open access. When this is not the case, the data sharing section for that particular dataset will describe why access has been restricted (see also section 3.2).

In regards to the specific repositories available to the SHOW consortium, numerous project partners maintain institutional repositories that will be listed in the next DMP version (Deliverable D14.3), where project scientific publications and in some instances, research data will be deposited. The use of a specific repository will depend primarily on the primary creator of the publication and on the data in question.

Some other project partners will not operate publically accessible institutional repositories. When depositing scientific publications they shall use either a domain specific repository or use the EU recommended service OpenAIRE (http://www.openaire.eu) as an initial step to finding resources to determine relevant repositories.

Project research data shall be deposited on an online data repository (see also Section 3.2) about ZENODO. In summary, as a baseline, SHOW partners shall deposit:

- Scientific publications – on their respective institute repositories in addition (when relevant) to the SHOW online data repository
- Research data – to the SHOW online data collection (when possible)
- Other project output files – to the SHOW online data collection (when relevant)

This version of the DMP does not include the actual metadata about the Research Data being produced in SHOW project. Details about technical means and services for building repositories and accessing this metadata will be provided in the next version of the DMP (deliverable D14.3).

5 Allocation of Resources

SHOW partners have to observe the policies set out in this DMP and datasets have to be created, managed and stored appropriately, to respond to the data management challenges in an efficient manner. The Data controller/processor responsibilities, also in reference to the Data Protection Officer (DPO), have been described in Deliverable D3.2 -Ethics manual & Data Protection Policy and Data Privacy Impact Assessment (DPIA).

A preliminary (DPIA) will be conducted at each pilot site in collaboration with the data collectors and processors (respective templates have been annexed in Appendix III), the project DPO and the Data Manager. Any ethical treatments and data protection mechanisms regarding the main data clusters on pilot and project level will also be included in the aforementioned Deliverables as required in POPD – Requirement POPD – H – Requirement No. 1 (D18.1) and (POPD – Requirement No. 3 D18.2 and the Data Management Plan (D14.2; ERTICO).

It should be noted that regarding the ORDP, the data controller must ensure that data are shared and easily available. Each data producer and WP leader is responsible for the integrity and compatibility of its data during the project lifetime. The data producer is responsible for sharing its datasets through open access repositories. He is in charge of providing the latest version. As the SHOW open data will be hosted by an open free of charge platform (e.g. ZENODO), no additional costs will be required for hosting the data.

The Data Manager, Dr. Nikolaos Tsampieris (ERTICO), will coordinate the actions related to data management; being responsible for the actual implementation of the DMP successive versions and the compliance to Open Research Data Pilot guidelines.

6 Conclusions

The first version of the Data Management Plan aims at presenting the general principles of DMP and a first data management analysis in anticipation of the respective procedures and infrastructures to be put in place in SHOW, to efficiently manage the generated and/or collected data.

In Section 2 the document provides an overview of the data that SHOW project will produce and outlines, in summary, identified datasets (categories and brief descriptions) which will be incrementally enriched along the project lifetime.

In Section 3, the datasets treatment according to FAIR principles is referenced while in Section 4 the SHOW Open Strategy & Participation in the Open Research Data Pilot is outlined.

The next version of the data management plan (D14.3), which is scheduled for M24, will include detailed dataset descriptions as the data types from the project data sources and the end-user applications and services will have been comprehensively defined in detail; hence, an enhanced and revised version of the SHOW DMP will be provided.

The Data Management Plan is a living document. In order to prepare for the next version, the contribution guidelines will be presented in a dedicated workshop (Face-to-face meetings or webinar tools will be used).The workshop will be organized by ERTICO (Data Manager) not later than M12, during which the scope of the data management plan will be explained and recommendations to be taken into account during the data collection/generation of their datasets will be suggested.

References

[1] EUROPEAN COMMISSION Directorate-General for Research & Innovation H2020 Programme Guidelines on FAIR Data Management in Horizon 2020- Version 3.0 26 July 2016 http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] ETSI EN 302 637-2 V1.3.1 2014-09 - Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part2: Specification of Cooperative Awareness Basic Service

[3] ETSI EN 302 637-3 V1.2.1 2014-09 - Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part3: Specification of Decentralized Environmental Notification Basic Service

[4] DATEX II extension: Safety Related Messages ITSTF13004v3. 2013-12-13)

[5] Safety-related message sets – Selection of DATEX II Codes, TPEG2-TEC-Causes

Appendix I - Data Management Plan Context

Reference to the Grant Agreement 875530 SHOW

a. Article 29.3

29.3 Open access to research data

*Regarding the digital research data generated in the action ('**data**'), the beneficiaries must:*

(a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:

(i) the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;

*(ii) other data, including associated metadata, as specified and within the deadlines laid down in the **'data management plan'** (see Annex 1);*

(b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

This does not change the obligation to protect results in Article 27, the confidentiality obligations in Article 36, the security obligations in Article 37 or the obligations to protect personal data in Article 39, all of which still apply.

As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data if the achievement of the action's main objective, as described in Annex 1, would be jeopardised by making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access.

Appendix II - SHOW relevant standards

Standard to abide	For which work item and how we comply	Anticipated contribution (Yes/No)	How (if Yes)
ETSI EN 302 637-2 Intelligent Transport Systems (ITS); Veh+B43icular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service	A4.2	Yes	Inputs for new messages and service standards
ETSI EN 302 637-3 Intelligent Transport Systems (ITS);Vehicular Communications; Basic Set of Applications; Part 3: Specifications of Decentralized Environmental Notification Basic Service	A4.2	Yes	Inputs for new messages and service standards
ETSI TS 103 301 Intelligent Transport Systems (ITS);Vehicular Communications; Basic Set of Applications; Facilities layer protocols and communication requirements for infrastructure services	A4.2	Yes	Inputs for new messages and service standards
ETSI TS 102 731 Security Architecture	A4.4 (Cyber Security view)	Tbd in the project Tbd in the project	
ETSI TS 102 940 ITS communication security architecture and security management	A4.4 (Cyber Security view		
ETSI TS 102 941 Trust and Privacy Management	A4.4 (Cyber Security view		
ETSI TS 102 942 Access Control	A4.4 (Cyber Security view		
ETSI TS 102 943 Confidentiality Services	A4.4 (Cyber Security view		
ETSI TS 103 097 Security header and certificate format	A4.4 (Cyber Security view		
SAE J2735 Dedicated Short Range Communications (DSRC) Message Set Dictionary	A4.4 (Cyber Security view		
IEEE 1609.2 Security für C-ITS	A4.4 (Cyber Security view		
CEN TS13149 : onboard data communication	A4.2	Yes	Proposal of new services
CEN EN 12896 : Transmodel - Reference Data Model for Public Transport	A4.2	Yes	Proposal of data model update
CEN EN/TS 15531 : SIRI - Service Interface for Real time Information	A4.2	No	
CEN TS 16614 : NeTEx - Network Timetable Exchange	A4.2	No	
ETSI EN 302 895 - Local Dynamic Map (LDM)	A4.2	No	

Appendix III - Data Controller & Data Processor Record of processing activities

*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.*

I. Data controller's record of processing activities

1	**Contact details of Data Controller**
Email	
Company address	
Telephone	
2	**Purpose of processing**
3	**Description of categories of data subjects and of the categories of personal data**
4	**Categories of recipients to whom the personal data have been or will be disclosed including recipients in third countries or international organisations**
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**

5	**Where possible, the envisaged time limits for erasure of the different categories of data**
6	**Where possible, a general description of the technical and organisational security measures for**
a	the pseudonymisation and encryption of personal data;
b	the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
c	the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident
d	a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing;

II. Data processor's record of processing activities

1	**Contact details of Data Processor**
Email	
Company address	
Telephone	
2	**Categories of processing carried out on behalf of the Controller**
3	**Where applicable, transfers of personal data to a third country or an international organisation, including the identification of that third country or international organisation**

4	**Where possible, a general description of the technical and organisational security measures for**
a	the pseudonymisation and encryption of personal data;
b	the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
c	the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident
d	a process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing;