Funded by the Horizon 2020 research and innovation programme
of the European Union (GA 769638)

Project Acronym:	**INTEND**
Project Title:	INtentify future Transport rEsearch NeeDs
Project Number:	769638
Topic:	**MG-8-7-2017**
Type of Action:	**Coordination and support action**

D5.2 Data Management Plan

(Version 1.0, 12/12/2017)

Deliverable:	D5.2 Data Management Plan
Work Package:	WP5: Dissemination, Communication and Exploitation
Due Date:	M3
Submission Date:	15/12/2017
Start Date of Project:	01/10/2017
Duration of Project:	12 Months
Organisation Responsible of Deliverable:	Coventry University Enterprises Ltd
Version:	0.1
Status:	Final
Author name(s):	Eleni Anoyrkati, Alba Avarello
Reviewer(s):	All partners
Nature:	☒ R – Report ☐ P – Prototype ☐ D – Demonstrator ☐ O - Other
Dissemination level:	☒ PU - Public ☐ CO - Confidential, only for members of the consortium (including the Commission) ☐ RE - Restricted to a group specified by the consortium (including the Commission Services)

Document history			
Version	**Date**	**Modified by**	**Comments**
0.1	10/11/2017	Eleni Anoyrkati, Alba Avarello	Draft
0.2	05/12/2017	Merja Hoppe, Thomas Trachsel	
0.3	06/12/2017	Norman Doege, Massimo Moraglio, Alkiviadis Tromaras	
0.4	08/12/2017	Slobodan Mitrovic, Vladislav Maras, Mirjana Bugarinovic, Norman Doege	
1.0	12/12/2017	Eleni Anoyrkati, Alba Avarello	Final version

Contents

Abbreviations

CERTH	Ethniko Kentro Erevnas Kai Technologikis Anaptyxis
CUE	Coventry University Enterprises Ltd
EC	European Commission
FTTE	Univerzitet u Beogradu – Saobracajni fakultet
TUB	Technische Universitaet Berlin
WP	Work Package
ZHAW	Zurcher Hochschule fur Angewandte Wissenschaften

Executive summary

This document establishes a Data Management Plan for the INTEND project. It has been implemented by Coventry University Enterprises Project Coordination Team with the review from INTEND Work Package Leaders, and describes how data will be collected, processed and generated during and after the project lifetime.

This document presents the basic descriptions of users and groups logical structure with roles. It also elaborates whether and how the data will be protected (due to intellectual property, personal or security reasons) or made open.

Introduction

The INTEND project participates to the Open Research Data Pilot (ORD pilot). The ORD pilot aims to improve and maximise access to and re-use of research data generated by Horizon 2020 projects. A Data Management Plan is required for all projects participating in the extended ORD pilot.

A Data Management Plan is a key element of good data management. It describes the management life cycle of the data collected and produced within the project. As part of making research data findable, accessible, interoperable and re-usable (FAIR), the Data Management Plan will need to include information on:

- the handling of research data during and after the end of the project
- the type of data collected, processed and generated
- the methodology applied
- whether the data will be shared or made open access
- the way data will be curated and preserved, also after the end of the project.

Therefore, the INTEND Data Management Plan will sensitize project partners on data management, give some common data management rules, make it easier to find project data when needed.

1 Purpose of data collection

The overall objective of the INTEND project is to deliver an elaborated study of the research needs and priorities in the transport sector utilising a systematic data collection method. Therefore, the project collects and generates data for internal use and further processing by the project partners to produce analysis, reports and plans, as well as data that will be made accessible for external users such as the project deliverables, project summaries stored in the transport research database and communication and dissemination material.

The following sections elaborate on the purpose of data collection, the origin and types of data, the format and storage methods and concentrate on Work Packages 2 to 5. This is because Work Packages 1 and 6 only comprise confidential data.

1.1 Data collection in WP2

The data collected in Work Package 2 supports an annotated literature review to identify future transport technologies and mobility concepts with a time horizon of 2020-2035 and the main political imperatives and visions regarding transport. This will result in the production of three deliverables:

- D2.1 Transport projects & future technologies synopses handbook
- D2.2 Report on key transport concepts of the future
- D2.3 Report on political imperatives

The data will come mainly from literature at European and international level, produced through sponsored research projects, scientific publications, forward looking exercises, industry studies and strategic research agendas, with emphasis on transport. In particular, the results coming from D2.1 report will also be publicly accessible via the transport research database on the website (for more details, please see section 3).

1.2 Data collection in WP3

The data collected in Work Package 3 provides the definition and evaluation of the most important trends (Megatrends) and technological advances in passenger and freight transportation, which impact the realization of the transport concepts of the future and the political imperatives identified in Work Package 2. This will result in the production of two deliverables:

- D3.1 Report on the main Megatrends
- D3.2 Report on Megatrends validation and impact assessment

The data will come from:

- a thorough literature review looking at relevant European, global and national projects, academic literature, reports from business sector, consultancy firms and worldwide research organizations elaborating megatrends in forward looking transport projects, etc.

- ad hoc queries conducted through online questionnaire to 90 people through the LimeSurvey software
- optional webinar offered to the people who want to participate to the online questionnaire

1.3 Data collection in WP4

The data collected in Work Package 4 provides guidelines for a forward-looking transport sector based on an understanding of the nature of the systemic change in the sector and the research needs arising from it. This will result in the production of three deliverables:

- D4.1 Sketch of future transport system
- D4.2 Gap Analysis
- D4.3 Transport research agenda: blueprint on transport research needs, priorities and opportunities

The data will come from:

- desk research that will put together the results coming from WP2 and WP3, along with additional research about technology, infrastructure and policy using studies about the future of mobility, international research agendas, international transport policy strategies, projects and literature
- qualitative interviews to 15-20 experts using the license-based software MAXQDA
- online survey to 100-150 people using the online survey tool Unipark

1.4 Data collection in WP5

The data collected in Work Package 5 delivers the formal structure and processes to enable an effective communication and dissemination of the knowledge gathered during the project as well as the outputs produced during its lifetime. This will result in the production of three deliverables:

- D5.1 Dissemination and exploitation strategy plan
- D5.2 Data Management Plan
- D5.3 Web tools

Due to the nature of the activities in the Work Package, it will produce a wide range of data:

- printed communication material like flyers and brochures
- online communication material like the project website (including news in short and long version), Facebook, LinkedIn and Twitter contributions, 3 newsletters
- publications and papers related to the work done and the results achieved in the project
- communications with TRIMIS, green car congress, UITP, ELTIS and with non-academic technology newsites to further disseminate the results of the project
- database of people registered on the website to receive the INTEND newsletters
- project documents repository containing project material for internal use only

- transport projects database, accessible via the website and containing mainly the transport projects reviewed in D2.1 report
- public deliverables and other materials published on the project website
- data included in the Transport Synopsis Tool
- presentations in power point or pdf format related to project consortium meetings and events attended.

In addition to this, we also have all dissemination activities that are documented in the Dissemination and exploitation strategy plan.

1.5 Data formats and storage methods

The generated data sets will be stored as MS Excel files. The analysis of the metadata (normally summary graphs) will be available within the relevant deliverables as MS Word and Adobe Reader pdf files. All related data sets will be backed up and stored for at least five years by CUE.

Other raw data that will be generated will be related to the megatrends Analytic Network Process matrices and the analysis with the relevant software as described in WP3. These will all be stored in file formats according to the software and type of data.

Further raw data will be generated in WP4 through conducting qualitative expert interviews (using the license-based evaluation software MAXQDA) and an online survey (using the online survey tool Unipark). All raw data resulting from this will be stored in common file formats such as XLS, CSV or SPSS for their further use in MS office applications.

Data files of the transport projects database will be stored in adequate data formats on the web server together with all other files of the project documents repository and files related to the project website. Users of the transport research database will have read-only permission. The access to the project documents repository will be password-protected, only project partners will have access to it. All the data stored on the web server will exist for five years: this will be established in a contract with the company hosting the website. After five years the data will be deleted.

As regards the collected data sets for the Transport Synopsis Tool, this will have to be named and stored separately as raw data. The Transport Synopsis Tool that will be hosted on a separate website shows a graphical representation of the data developed in WP3. This graphical representation is accessible for everyone. TUB will have password-protected access to this data stored on the corresponding webserver.

2 FAIR data

The data collected and generated in the project should be made findable, accessible, interoperable and re-usable (FAIR).

INTEND data will not be **findable** with metadata, but the majority of the material produced, including all public deliverables, will be findable on the project website. It will be named according to the following convention and will provide clear version numbers: Project name + item name + version number. For example for the deliverables it will be: INTEND_Dx.x_shorttitle_vx.x.docx (or .pptx/.xlsx/.pdf…).

The INTEND data will be **accessible** mainly through the public deliverables that will be published on the project website. The confidential deliverables and other data (such as meeting presentations, meeting minutes, deliverable drafts, project tasks guidelines, etc.) will be also available in the documents repository accessible through the website just by the project partners.

The data will not contain private information about transport sector stakeholders or participants to the project activities. This refers in particular to:

- individual results of the online survey of WP3
- the contents shared during the webinar in WP3
- the reporting of the interviews to the experts in WP4
- individual results of the survey in WP4

This data is considered confidential. In general, data sharing will comply with privacy and ethics guidelines of the project.

The INTEND data will be **interoperable**. It will adhere to standards for formats compliant with available software applications and will be using standard vocabularies. This will allow data exchange and re-use between researchers, institutions, organisations and countries.

The data will be **re-usable** by third party, also after the end of the project, with appropriate reference to the INTEND project. The public deliverables will be published on the project website once they are approved by the European Commission.

3 Procedures for data collection, storage, protection, retention and destruction

All the data sets will be managed in line with the Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020.

A central database / platform (repository) named as “**Green Archives**” (Self-archiving) will be developed on the project’s website. All the aforementioned literature for D2.1 report will be stored centrally in order to maximise standardization of the collected files which will have to be titled and categorised according to mode or theme. This database will help project partners to share and easily access literature, but will also allow external persons to gather data about the several transport projects. It will be accessible via the project home page (www.intend-project.eu) and function like a Wiki, allowing its user to browse through the different thematic fields of D2.1 report in order to access the project summaries. A tab for each transport mode will be available where the user can click on each mode and get the full list of projects that are relevant. The projects will be classified by thematic areas (competitiveness, environment, energy, infrastructure, systems based on D2.1 report) and technology cluster, while the user will be able to see the technologies that each project has researched. In addition, the user will be able to filter the results by thematic areas and sector (passenger/freight). This will offer a brief thesaurus of projects and technologies that have been covered in D2.1 report. The Project Coordinator will be responsible for the management of this “green” database / platform.

Together with the *transport projects database* accessible by everybody, there will be a **project documents repository** that will contain all project-related data like confidential and public deliverables, other confidential documents, working documents related to project tasks, project meetings presentations. It will be accessible via the project homepage just by the INTEND partners; they will have a personal password-protected account to access the repository. The main purpose of this repository is to have an actual database that will help optimizing data exchange and collaborative work processes in the project, as well as store confidential documents. As already mentioned above, data determined to be accessible for an external audience (like public deliverables, electronic leaflets, transport projects, future technologies synopses handbook) will be made freely accessible to the public or interested parties separately via the homepage of the project website.

A **database of people** will also be developed within WP5. This will contain a list of people registered on the website to receive the INTEND newsletters. The subscribers only needs to enter their first and last name and their email address. This information will be stored in form of a list on the INTEND webpage server in the documents repository. As it contains some sensitive information, it will be managed and treated in full compliance of data protection national and EU legislation. In specific, the subscribers have to confirm the subscription via an automated mail and they also receive a mail confirmation with the information on how to cancel the subscription. They can easily unsubscribe via the link presented on the newsletter or by sending an email to the contact person(s) so that they can manually be remove from the list of subscribers. Only CUE, TUB and CERTH and the external subcontractor will have access to this list.

As regards the **online questionnaire** expected to take place within WP3, all the input data will be obtained using the Survey service (LimeSurvey software). Survey service will be hosted on a physically secured server, operating in the Secure Socket Layer (SSL) mode and also all data will be stored in the same server storage pools, protected with Access Control List (ACL). None of the questions related to participants’ personal data (except Country of Business) will be offered in the Survey. The Main Survey module contains questions from the Main Survey and the Additional Survey. Survey raw data will contain the IP address of the participant only during single Survey session. If the specific question (that belongs to the Main Survey) does not fulfil the required answer scheme, an additional question that belongs to the Additional Survey will be activated (by predefined logical schema, stored in the Main Survey Module). In this way, both surveys could be completed within the same Survey session. Just upon the Survey session end, the participant’s IP address and other session data will be erased from the participant record, in order to avoid any personal metadata collection. The IP address will be erased automatically from the raw record by the customized script.

The **webinar** service in WP3 will also be hosted on the physically secured server, operating in the Secure Socket Layer (SSL) mode. All data will be stored in the same server storage pools, protected with Access Control List (ACL), as well. During webinar realization, the so called “recording mode” will be disabled at both the server and the client side. Also, none of the webinar session data will be permanently stored on the hosting server. All personal documents, personal session data and the server log data will be erased just upon the end of the webinar by activating the log data erase script: this searches for any log and session data acquired during the webinar realization and erases it in all related log files.

Both webinar and survey processes will be conducted without the usage of the 3rd party software and services: all data traffic will be based on direct client-server connection without intermediate rerouting to 3rd party servers (e.g. outer cloud services, etc.).

As regards User Access Control Hierarchy related to the specific user research and operational role, the considered access control include several operational levels:

- Visual access to research data (read-only user access type), on the server by using the provided secured computer with authentication, authorization and auditing capabilities (AAA). Usage of devices for data recording by optical devices (i.e. cameras), physical media (USB flash drive, memory card, optical medium, hard disk, etc.), as well as the possibility of sending data over the intranet or external (internet) connection is strictly prohibited (in technical and organizational way);
- Visual access to research data and possibility to change (read-write user access type), on the server by using the provided secured computer with authentication, authorization and auditing capabilities (AAA). Usage of devices for data recording by optical devices (i.e. cameras), physical media (USB flash drive, memory card, optical medium, hard disk, etc.), as well as the possibility of sending data over the intranet or external (internet) connection is strictly prohibited (in technical and organizational way);
- Visual access to research data (read-only user access type), on the server by using the provided secured computer with authentication, authorization and auditing capabilities (AAA). Usage of devices for data recording by optical devices (i.e. cameras), physical media (USB flash drive, memory card, optical medium, hard disk, etc.), is allowed, but the possibility of sending data over the intranet or external (internet) connection is forbidden. The process is allowed and monitored by the Project data security officer whose task is to register all devices and storage media used for data copy in order to apply erase/destroy-data procedures upon data manipulation;
- Visual access to research data and possibility to change (read-write user access type), on the server by using the provided secured computer with authentication, authorization and auditing capabilities (AAA). Usage of devices for data recording by optical devices (i.e. cameras), physical media (USB flash drive, memory card, optical medium, hard disk, etc.), is allowed, but the possibility of sending data over the intranet or external (internet) connection is forbidden. The process is allowed and monitored by the Project data security officer whose task is to register all devices and storage media used for data copy in order to apply erase/destroy-data procedures upon data manipulation;
- Data manipulation on computers / other devices used as destination in copy procedures mentioned above. Data manipulation is allowed and monitored by the Project data security officer whose task is to apply erase/destroy-data procedures on destination device upon data manipulation;
- Data transmission between research/operational personnel using the secured/encrypted intranet and Internet connections hosted by the devices/servers with authentication, authorization and auditing capabilities (AAA).

Survey and webinar servers are physically secured in rooms where access is allowed only to authorized personnel. All processes related to survey and webinar are exclusively hosted by these servers strictly excluding the use of any external (3rd party) services and their servers. Complete communication is SSL encrypted, using digital certificates issued by relevant CA (Certified Authority) institutions.

As regards WP4 **qualitative expert interviews**, all the input data will be generated either digitally (audiofiles) or written as flow texts. Using the license-based qualitative evaluation software MAXQDA, raw data will be transcribed and afterwards stored on internal server. To guarantee protection of personal and sensitive data throughout the whole evaluation process, the software will be installed on a limited number of computers only, which are personalized through password-protection. For further scientific evaluations and with regard to the dissemination material, all information that could possibly allow inference on the interviewee will be disguised or entirely removed from the transcripts.

All the input data produced within WP4 **online survey** will be generated by using the online survey software Unipark. With this software tool, the customer is solely responsible for who will participate in the survey, how it is made available to participants and what data finally will be collected. Any information stored on the provider's server is treated confidentially according to GDPR (General Data Protection Regulation of the EU) and access to data is only possible for authorized personnel. For customers, located in the European Economic Area (EEA), all processing of Personal Data is – according to the provider – performed in accordance with privacy rights and regulations following the EU Directive 95/46/EC of the European Parliament and of the Council of 24th October 1995 (the Directive), and the implementations of the Directive in local legislation.

All data generated in WP4 will be held securely on password-protected computers and server systems, which are periodically maintained by the IT-division of ZHAW. The computers as well as the server systems are kept in office rooms, where access is allowed to authorized personnel only. Whenever the computers are moved to another location or used externally, the owner of the computer has to follow the strict guidelines and safety regulations of ZHAW concerning the handling of internal devices with sensitive data abroad.

4 Open access to scientific publications

The Project Partners who want to publish peer reviewed scientific papers that contain data and results from the project, or promote the project, will have to ensure that any effort will be made for the papers to be made available as "open access" if possible.

Most of the papers will be published in conferences which is a much quick process with a great outreach. The proceedings of the conferences will be freely available.

All articles will have to contain the project's acronym, reference to the words EU and Horizon 2020 ensuring the promotion of the funding scheme and the identification and accessibility of the work in the future.

5 Intellectual Property Rights

As the INTEND project is a Coordination and Support Action, new knowledge created will form a common ground to accumulate new knowledge. The management of intellectual property will be on the focus of the dissemination manager and the coordinator.

The dissemination manager, CERTH, possesses significant experience in IPR and will make sure that the following rules will be applied:

- Pre-existing know-how will remain the property of the partner having brought it into the project.
- Pre-existing know how will be made available, by their owners, to the project participants on the need to know basis. Usage outside the project will be decided among the owners and the potential users on a case-by-case basis.
- Knowledge will remain the property of those the partner involved in its generation / production.
- Knowledge jointly generated (without possibility to identify the individual share of work) shall be the joint property of the partners concerned.

6 Allocation of resources, data security and ethical aspects

The INTEND data will be stored in the documents repository, in the transport projects database, on the project website and in the Transport Synopsis Tool database. The costs will be covered by the budget.

Each project partner is responsible for a reliable data management related to their work within the project. However, the Project Coordinator is responsible for the overall data management of the project.

Each project partner is also responsible for the security and preservation of their data. The project partners' servers are regularly and continuously backed-up, as well as the project website. Moreover, access to the documents repository and the Synopsis Tool (both through the project website) will be controlled by use of user name and password. In the event of an incident, the data will be recovered according to the necessary procedures of the data owner.

As regards ethical aspects, Work Package 6 is entirely dedicated to specify all relevant ethical aspects related to the involvement of people to the project activities. Two deliverables will be submitted: D6.1 "H – Requirement No. 1" and D6.2 "H – Requirement No. 2". Specifically, D6.1 includes information on the procedures used to identify and recruit research participants to be involved in the diverse project activities (e.g. webinar, surveys and interviews) and on the consent procedures for the participation of humans, comprising a template of the informed consent form to be filled in by the participants before taking part in any project activity. D6.2 provides details about the procedures that will be implemented for data collection, storage, protection, retention, destruction and confirmation in order to protect individuals' privacy.