Mobilit**y** as a Service in a multimodal European cross-border **Corridor** (**MyCorridor**)

Deliverable 2.1

Data management plan

Aimilia Bantouna, WINGS ICT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723384

MyCorridor	*D2.1*
Dissemination level:	*PU*
Work package:	*WP2*
Lead beneficiary:	*Aimilia Bantouna (WINGS ICT)*
Other beneficiaries involved:	*Marie-Claire Smith (OC), Vassilis Mizaras (SWARCO Hellas), Vassilis Foteinos (WINGS ICT), Nelly Giannopoulou(WINGS ICT), Andreas Georgakopoulos (WINGS ICT), Panagiotis Demestichas (WINGS ICT), Jeremy Godley (OC)*
Date due to EC:	30/11/2017 (M6)
Date of Delivery to EC:	07/12/2017
Status (F: final; D: draft; RD: revised draft):	F
File Name:	MyCorridor_D2.1 Data Management Plan_Final
Version:	Final

Document history

Version No.	Date	Details
0.1	04/10/2017	1st draft version – Table of Contents
1.0	20/11/2017	Full Draft – ready for peer-review
1.1	29/11/2017	Peer-reviewed version
Final	07/12/2017	Final version submitted to EC

Reviewers List

Name	Company
Gino Franco	External Expert
Laura Coconea (MyCorridor Quality Assurance Manager)	SWARCO MIZAR
Mr Dionysis Kehagias	CERTH-ITI
Ms Marie-Claire Smith	OC

This project is co-funded by the European Union under the Horizon 2020 Research and Innovation Programme. The content of this document reflects solely the views of its authors. The European Commission is not liable for any use that may be made of the information contained therein.

The members of the MyCorridor project Consortium shall have no liability for damages of any kind including, without limitation, direct, special, indirect, or consequential damages that may result from the use of these materials.

This deliverable is a draft document subject to revision until formal approval by the European Commission.

The MyCorridor project consortium consists of:

No.	Name	Short name	Country
1	NEWCASTLE UNIVERSITY	UNEW	UK
2	ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS	CERTH	EL
3	OSBORNE CLARKE LLP	OC LLP	UK
4	WINGS ICT SOLUTIONS INFORMATION & COMMUNICATION TECHNOLOGIES EPE	WINGS ICT	EL
5	SWARCO MIZAR SRL	SWARCO MIZAR	IT
6	EFARMOGES EXYPNOU LOGISMIKOU KYKLOFORIAS & METAFORON AE	SWARCO Hellas	EL
7	CHAPS SPOL SRO	CHAPS	CZ
8	HACON INGENIEURGESELLSCHAFT MBH	HACON	DE
9	MAP TRAFFIC MANAGEMENT BV	MAPtm	NL
10	VIVA WALLET HOLDINGS - SOFTWARE DEVELOPMENT SA	VivaWallet	EL
11	AMCO OLOKLIROMENA SYSTIMATA YPSILIS TECHNOLOGIAS ANONYMI VIOMICHANIKI KAI EMPORIKI ETAIRIA	AMCO	EL
12	TOMTOM DEVELOPMENT GERMANY GMBH	TOMTOM	DE
13	ROMA SERVIZI PER LA MOBILITA SRL	RSM	IT
14	TTS Italia	TTS	IT
15	PANEPISTIMIO PATRON	UPAT	EL
16	IRU PROJECTS ASBL	IRU	BE
17	SALZBURG RESEARCH FORSCHUNGSGESELLSCHAFT M.B.H.	SFRG	AT

Table of Contents

List of tables

Abbreviation List

Abbreviation	Definition
C-ITS	Cooperative Intelligent Transport Systems
DMP	Data Management Plan
DPIA	Data Privacy Impact Assessment
EMP	Ethics Management Panel
FAIR	Findable, Accessible, Interoperable and Re-usable
ITS	Intelligent Transportation Systems
MaaS	Mobility as a Service
ORDP	Open Research Data Pilot
PT	Public Transportation
UCD	User-Centred Design
UI	User Interface
VAS	Value-Added Services

Executive Summary

The current document constitutes MyCorridor first version of Data Management Plan, namely D2.1: “Data Management plan”, and the 1st version of MyCorridor’s guide on how MyCorridor consortium will manage the data to be used throughout the whole project, consortium decisions with respect to making the data Findable, Accessible, Interoperable and Re-usable (FAIR) and the respective mechanisms to enable data management decisions.

MyCorridor project targets citizens using public and private transport including citizens’ barriers due to low digital literacy and travellers with motor or sensory limitations. MyCorridor will include all potential types of users coming from diverse backgrounds and travel patterns and preferences with the ambition to offer tailored services to each specific user group (i.e. older people, people with disabilities, commuters, business travellers, tourists, etc.) and even to individualise services according to the user profile and the history of use. Thus, the project by definition will in most cases involve users that use public and private transport as travellers and commuters with the general use of those terms. As such, MyCorridor will thereby empower the users for active and independent travelling, but, also, socially included living.

Due to the fact that the project will collect user-related data, the Consortium will fully comply with any laws and regulations in any relevant jurisdiction relating to privacy or the use or processing of data relating to natural persons, including: (a) EU Directives 95/46/EC and 2002/58/EC (as amended by 2009/139/EC) and any legislation implementing or made pursuant to such directives and the Privacy and Electronic Communications (EC Directive) Regulations 2003; (b) from 25 May 2018, the EU General Data Protection Regulation 2016/679 ("GDPR"); and (c) any laws or regulations ratifying, implementing, adopting, supplementing or replacing GDPR; in each case, to the extent in force, and as such are updated, amended or replaced from time to time.

The relevant data management aspects are being analysed in this document and are related to a) the data to be collected for the pilots and the validation activities of MyCorridor, b) consortium decisions with respect to making the data FAIR and the respective mechanisms to support these decisions, c) the security processes to be applied, including data recovery as well as secure storage and transfer of sensitive data and d) the related ethical aspects with respect to e.g., (sensitive) personal data, informed consent, restrictions and constraints of contacting and testing MyCorridor services with real life users, etc., as defined in the ethical framework and policy of MyCorridor [1][2] taking into account European and national/local ethical guidelines and legislation.

In specific, following the EC guidelines for the Data Management Plans [3], **Chapter 1** introduces the purpose and intended audience of the current document as well as the interrelations to other project work items. **Chapter 2** summarises information related to the data to be collected, the template to be used for describing the datasets and the objectives of the project that will be met through these data collection and processing. **Chapter 3** describes the processes and the mechanisms that will be applied for making the data FAIR. **Chapter 4** provides information on the financial aspects of making data FAIR, e.g., costs, how will these be covered, data controller, etc. **Chapters 5** and **6** deal with data security and ethical aspects accordingly. Finally, **Chapter 7** concludes the document.

The current document is a living document and thus it will be updated during the project lifetime as needed, including more detailed information regarding the collected data. The next official updated version will be publicly released as part of D2.2 “MyCorridor interoperable, open and seamless architecture and MyCorridor subsystems and modules specifications”, i.e., M24 (May 2018), providing more information to aspects such as:

- The descriptions of the different datasets, including their reference, file format, standards, methodologies and metadata and repository to be used;
- Institutional repositories in cases where the project partners maintain such a repository;
- More details about license types and rules for the data producers so as to allow the widest reuse possible for the collected data;
- The updated list of services decomposition and their clustering (Annex 1) and the mapping of MyCorridor services to them
- The final DPIA form (Annex 2);
- The name(s) of the data controller(s) and the data manager – see Chapter 4.

Still, intermediate internal (unofficial) versions of the DMP will be released before the two Pilots rounds, on M18 and M28 respectively.

1 Introduction

1.1 Purpose of the document

This is the deliverable report for MyCorridor project Work Package 2, D2.1, Data Management Plan, as required by the project's Grant Agreement number 723384. It defines the data management process that will be followed throughout the project lifetime. As MyCorridor project is currently only in M6, some of these processes may change or more data types may be collected, which we will document in this data management plan as the project evolves. As such, this deliverable is a living document that will be updated when necessary and the updated information will officially be released to the public as part of D2.2 "MyCorridor interoperable, open and seamless architecture and MyCorridor subsystems and modules specifications", i.e., M24 (May 2018).

The scope of the deliverable is a) to demonstrate compliance with all applicable legislation in relation to the data being collected and used, b) to report on the data to be collected for the pilots and the validation activities of MyCorridor, c) present consortium decisions with respect to making the data FAIR and the respective mechanisms to support these decisions, d) discuss on the security processes to be applied, including data recovery as well as secure storage and transfer of sensitive data and e) summarize the related ethical aspects e.g., with respect to (sensitive) personal data, informed consent, restrictions and constraints of contacting and testing MyCorridor services with real life users, etc., as defined in the ethical framework and policy of MyCorridor [1][2] taking into account European and local ethical guidelines and legislation.

1.2 Intended audience

The goal of this manual is to compose a Data Management Guide for all the researchers within MyCorridor and it is considered to be a living document. As the information flow is constant, multi-dynamic and layered, it is imperative to define early in the project the data management processes governing all activities within MyCorridor. Still, being a public document, this Deliverable may also serve as a guide for all researchers working on MaaS initiatives beyond the specific project.

1.3 Interrelations

Data Management aspects are closely related to:
a) Ethics Issues in MyCorridor, especially in the context of collecting, managing and processing (sensitive) personal data from real-life users,
b) Security aspects, e.g., cross-border security issues when dealing with electronic authentication in cross-border solutions, data privacy and protection, data ownership, etc.,
c) Design and development activities in terms of defining the data that need to be collected or re-used so as to offer tailored services (national or cross-border multi-modal corridor suggestions) to each specific user group (i.e., older people, people with disabilities, commuters, business travellers, tourists, etc.) and even to individualise services according to the user profile and the history of use and
d) Legal issues related to personal data (including sensitive personal data), security and privacy.

Therefore, this document will be updated as the work evolves and in close synergy with the following activities:

- A9.3: "Quality Assurance and Ethical Issues";
- WP10: "Ethics requirements";
- A2.3: "Interoperability and cross-border security issues";
- A2.5: "Data management, reliability and QoS";
- A4.1: "Traffic Management Services";
- A4.4: "Added value services";

- A5.1: “Profiling mechanism”;
- A5.3: “Personalisation and inclusiveness mechanism”;
- A5.4: “Development of mobile interfaces for all types of devices”;
- A6.1: Pilot plans and impact framework;
- A6.2: Pilot realisation, and
- A7.4: “Operational, equity and legal issues including security and privacy”.

2 Data Summary

2.1 Data types to be processed, Purpose of the data collection/generation and its relation to MyCorridor objectives

MyCorridor mission is to facilitate sustainable travel in urban and interurban areas and across borders by replacing private vehicle ownership by private vehicle use, as just one element in an integrated/multi-modal MaaS chain. To this end, MyCorridor will provide an innovative platform, based on mature ITS technology that will combine connected traffic management and multi modal services so as to facilitate modal shift. MyCorridor will prove this paradigm change through a number of European sites, which are performing long distance and cross border Pilots in a corridor of 6 European countries; from the South (Greece, Italy) up to Central (Austria, Germany, the Netherlands) and Eastern Europe (Czech Republic). Those sites will develop Mobility Package tokens, purchased through a one-stop-shop and will incorporate the following services: a) Traffic management services b) Services related to MaaS PT interface c) MaaS vehicle related services and d) Horizontal (business related) services.

This mission is then split to the following 3 objectives as included in MyCorridor Grant Agreement:

- **Objective 1:** Integration of MaaS vehicles into a multimodal service chains platform;
- **Objective 2:** Provision of a new business paradigm, actor and model for pan-European cross-border adoption; and
- **Objective 3:** Proof of concept of the new business model and integrated platform by selected UC's and performance of full operational analysis and impact assessment through interconnected Pilots across a European corridor.

In order for MyCorridor to achieve its mission and to meet its objectives, access, collection, process and management of the 4 data types described in chapters 2.1.1-2.1.4 is necessary. Additional data may be appended to this list in case this will be deemed necessary, or as a result of the detailed architectural design that will be delivered at a later stage. These new datasets will be included in the next version of the deliverable. The data collection will comply with all national and EU ethics and legal requirements.

2.1.1 Data from the services that will be aggregated in one-stop-shop

Although the project will provide a holistic "look and feel" of a one-stop-shop (see Objective 1), the purchased services are individual services integrated into the platform and constitute running real world services. Therefore, the data required for each of these services are collected, processed and managed through the respective service and/or data provider. **MyCorridor consortium has no involvement in these data.**

The current services decomposition and clustering of MyCorridor one-stop-shop follows in Annex 1. All services that will be brought in the one-stop-shop will be mapped to one or more of those services types, whereas the specific configuration that will be available in each country per each, will be defined. Updates in both decomposition and clustering may emerge (i.e. more types of services may emerge in some clusters, like the "Added Value" one).

2.1.2 Metadata from the services that will be created by the system to support the system functionalities

These metadata are provided by the service providers that integrate their services in MyCorridor platform. They refer to those data that will be provided so as to describe the services, the cost of the services and any other information required for integrating them in MyCorridor (i.e. API) in terms of including them as part of the overall mobility token offered to the end-user (see objective 1), defining the cost of the token to be offered, paying back the service provider and allowing the user to redeem the services they have bought (see objective 2). In specific, those data will be created/provided as part of the service alignment process – the process where each service will be registered in the one-stop-shop of MyCorridor through the specific interface built by the project. In addition to the info items mentioned above, one of the metadata provided/created will be the associated semantics (i.e. mapping of service to the MyCorridor clustering, tags of the service, etc.) as well as the Business Rules that will be imposed by each service provider.

2.1.3 Subjective data that will be collected during focus groups, surveys and during Pilots

This type of data are collected during all types of qualitative surveys, focus groups (WP1), workshops (WP7) and, of course, pilots (WP6) that will take place in the project. These data are **collected, managed and processed by MyCorridor partners**. Those may be collected from travellers (all types of them), service providers/developers as well as other types of stakeholders. In all cases, they will be anonymised. In the case of **travellers**, subjective data will deal mostly with current travel preferences, habits and needs as well as satisfaction/perceived quality and acceptance (perceived/rated by users) of the services/packages provided to them, acceptance and usability ratings (perceived by users) of the one-stop-shop functionalities and the incentivisation shemes, willingness to have and use the one-stop-shop system (as denoted by the users), reported change in travelling patterns (if any and shared), comfort, pleasure and trust using the system (as perceived by the users).

In the case of the developers/service providers, the performance, complexity, accuracy, security and manageability of the integration will be rated as well as the IT related effort/knowledge up-scale (i.e. how much the developer within everyday working environment has to change their way of working in order to integrate their services to the MyCorridor platform). In addition, the perceived effort for integration, the ease of process and ease of use of integration manual/ process/ steps, the acceptance and added value of the one-stop-shop, etc.

The concrete list of subjective data is/will be specified in the context of each research work item (i.e. D1.1 for the focus groups, D6.1 for the pilots, etc.), as it is necessary that that the research context is first defined for their exhaustive clarification.

2.1.4 Data that will be logged (in the cloud server and the mobile devices) during Pilots

To the project current understanding, the following types of data will be **logged, managed and processed in the MyCorridor system** during performance/real-life Pilots:

- **Personalisation data**: MyCorridor one-stop-shop will offer personalised, context-aware and inclusive User Interfaces (UIs). This is related to objective 1 and involves a) user profiling for personalized recommendations, optimized results, match-making profiles, etc.; b) personalisation and device-oriented adaptation (e.g., to specific types of devices, on different screen sizes and screen resolutions); and c) personalization for guaranteeing accessibility to all user groups, elderly and travellers with disabilities included. This attribute obviously requires the collection, the processing and the management of different types of data, the main categories of which are as follows:
  a) User's profile;
  b) User's preferences and maybe behaviour;
  c) User's usage history of searching and selecting services; and
  d) Device characteristics.

Moreover, MyCorridor will deliver Value-Added Services (VAS) related to e.g., safety and smooth mobility with the help of information and communication technology. In this direction, the platform integrates data (e.g. weather forecasts, road incidents) from all transportation services as well as from other sources of information (e.g. social media). In addition, these added value services are enriched with tourism related data (e.g. points of interest, hotels, shopping centres, restaurants, monuments), as well as information retrieved from the users' profiles (e.g. about the health status of the traveller).

The project involves data collection in the context of user testing and demonstration activities. For this reason, human participants will be involved in the 2nd round of pilots and data will be collected concerning **both the data related to the personalization of the UIs and the data related to the VASs**.

- **Matchmaking data**: Matchmaking takes place in the back-end of the system, receiving as input the travellers profiles and personalisation indices and providing as output the mobility package that considers as the most appropriate for them. This matchmaking input-output – which is associated to the above personalisation part – is perhaps one of the most interesting types of data that will be available in MyCorridor, as they will associate traveller's preferences and profiles to mobility recommendations/outputs/products, the acceptance of which will be later objectively validated through actual usage.

- **Data logged during performance:** Regarding the **travellers**, a series of data will be logged during performance – meaning when using the one-stop-shop – i.e. travel information from route planner(s) to be embedded in one-stop-shop, type and combination of mobility products selected, frequency of selection, validation of the selected products usage through tokens, etc. per user group, MyCorridor platform analytics (i.e. session interaction time, time for completion of a user request, visit times and frequency, no. of registrations, issues and errors reported, etc.). All of them will be anonymised; still, in their aggregated form, after being processed, will feed the impact assessment of A6.4. From the developers' end, the respective MyCorridor platform analytics will be logged (e.g. session duration, visit times and frequency, no. of registrations, issues and errors reported) as well as service registration/integration success, including number of tries, number of times assistance was required, etc.

- **Traveller feedback data**: Through the traveller feedback module of A3.4, an upper level (subjective) evaluation of the system as a whole and its products will be enabled, such as correct service output/result (content accuracy), success in information retrieval (get intended output), etc. Though this is subjective data, it corresponds mainly to the real-life travellers that are going to use the system (in the 2nd pilot round) and the input from them will be automatically logged in the system (through a web-based form). This cluster of data will feed the final user acceptance of the system.

The payment transaction process that will take place in the 2nd pilot round of real-life trials will be protected by the relevant mechanisms and protocols. The transactions data will not be stored in one-stop-shop; will be administrated by VivaWallet, the travellers and the service providers.

2.2 Dataset Description

This chapter provides a preliminary template (see Table 1) to be used for describing the datasets to be produced or collected in MyCorridor project, which has been identified at this stage of the project. As the nature and extent of the datasets can evolve during the project, changes in the template may occur. The descriptions of the different datasets, including their reference, file format, standards, methodologies and metadata and repository to be used will be provided in the next DMP version, i.e., in D2.2 "MyCorridor interoperable, open and seamless architecture and MyCorridor subsystems and modules specifications", M24 – May 2018.

Table 1. Dataset Description template.

Dataset Reference	**MyCorridor_WPX_AX.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 Activity in which it is generated and its version (for example: MyCorridor_WP5_A5.1_01)**
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 Origin	Specify the origin of the data.
Data Size	State the expected size of the 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, etc.)
Re-used existing data	Y/N. If Yes, state the re-used data and how/from where they were retrieved.
Data Utility	Outline to whom the dataset could be useful – potential secondary users.

2.3 Open Access approach

MyCorridor consortium has agreed to follow an "open access" approach (as much as possible depending on the specific data type) following the respective Horizon 2020 guidelines to ensure that the results of the project results provide the greatest impact possible. MyCorridor will ensure the open access[1] 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. Publications and research data made available to third parties will not contain any personal information.

The following lists the minimum fields of metadata that should come with a MyCorridor project-generated scientific publication in a repository:

- The terms: "European Union (EU)", "Horizon 2020"
- Name of the action (Research and Innovation Action)
- Acronym and grant number (MyCorridor, 723384)
- Publication date
- Length of embargo period if applicable
- Persistent identifier

[1] *http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm*

When referencing Open access data, MyCorridor will include at a minimum the following statement demonstrating EU support (with relevant information included into the repository metadata):

“MyCorridor is funded by the European Union within Horizon 2020 research and innovation programme under grant agreement No 723384.”.

More detailed information with respect to the open access process for publications and the respective disclaimers that will be used during MyCorridor lifetime can be found in Chapter 5 “Open Access and obligatory disclaimers” of [4].

The MyCorridor consortium will strive to make many of the collected datasets open access. When this is not the case, the data sharing chapter for that particular dataset will describe why access has been restricted (See Chapter 2.2 – “Data Sharing” field).

In regards to the specific repositories where MyCorridor datasets will be hold during and after the project, they will be noted in the “Archiving and Preservation” field of Table 1 (see Chapter 2.2). In cases where the project partners maintain institutional repositories, these will be listed in the following DMP version (in D2.2 “MyCorridor interoperable, open and seamless architecture and MyCorridor subsystems and modules specifications”, M24 – May 2018). The project scientific publications and in some instances, research data will be deposited on the institutional repository depending primarily on the primary creator of the publication and on the data in question. In cases where the project partners do not operate publicly accessible institutional repositories, they shall use either a domain specific repository or the EU recommended service OpenAIRE (http://www.openaire.eu) as an initial step to finding resources to determine relevant repositories for depositing the scientific publications and the respective data. A good candidate that will be considered for use is the online data repository ZENODO[2], which is 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.

In summary, as a baseline MyCorridor partners shall deposit:

- Scientific publications – on their respective institute repositories in addition (when relevant) to a public MyCorridor repository such as ZENODO
- Research data – to the MyCorridor public repository (e.g., ZENODO) collection (when possible)
- Other project output files – to the MyCorridor public repository (e.g., ZENODO) collection (when relevant)

This version of the DMP does not include the actual metadata about the Research Data being produced in MyCorridor project. Details about technical means and services for building repositories and accessing to this metadata will be provided in the next version of the DMP. The initial template document is provided in Chapter 2.2 and will be used by project partners to provide all requested information.

[2] *https://zenodo.org/*

3 FAIR data

MyCorridor project will in principle participate in the Open Research Data Pilot (ORDP) but data marked as "restricted" or under an "embargo" period will be excluded. To this end, the data that will be generated during and after the project and will be included in ORDP should be 'FAIR', that is findable, accessible, interoperable and reusable. These requirements don't affect implementation choices and don't necessarily suggest any specific technology, standard, or implementation solution.

The FAIR principles were generated to improve the practices for data management and data-curation, and FAIR aims to describe the principles in order to be applied to a wide range of data management purposes, whether it is data collection or data management of larger research projects regardless of scientific disciplines.

With the endorsement of the FAIR principles by H2020 and their implementation in the guidelines for H2020, the FAIR principles serve as a template for lifecycle data management and ensure that the most important components for lifecycle are covered. This is intended as an implementation of the FAIR concept rather than a strict technical implementation of the FAIR principles.

Making data findable, including provisions for metadata

- The datasets will have very rich metadata to facilitate the findability.
- All the datasets will have a Digital Object Identifiers provided by the MyCorridor public repository (e.g., ZENODO).
- The reference used for the dataset will follow this format: MyCorridor_WPX_AX.X_XX, including clear indication of the related WP, activity and version of the dataset.
- The standards for metadata will be defined in the "Standards and metadata" section of the dataset description table (see Table 1 for the current version of the template).

Making data openly accessible

- Datasets openly available are marked as "Open" in the "Data Sharing" section of the dataset description table (see Table 1).
- The repository that each dataset is stores, including Open access datasets, are mentioned in the "Archiving and Preservation" section of the dataset description table (see Table 1). Public repositories such as ZENODO will be one of the considered options.
- "Data sharing" section of the dataset description table (see Table 1) will also include information with respect to the methods or software used to access the data of each dataset.
- Data and their associated metadata will be deposed either in a public repository or in an institutional repository.
- "Data sharing" section of the dataset description table (see Table 1) will outline the rules to access the data if restrictions exist.

Making data interoperable

- Metadata vocabularies, standards and methodologies will depend on the repository to be hosted (incl. public, institutional, etc.) and will be provided in the "Standards and metadata" section of the dataset description table (see Table 1).

Increase data re-use (through clarifying licenses)

- All the data producers will license their data to allow the widest reuse possible. More details about license types and rules will be provided in the next version (D2.2 "MyCorridor interoperable, open and seamless architecture and MyCorridor subsystems and modules specifications", M24 – May 2018).
- "Data Sharing" section of the dataset description table (see Table 1) is the field where the data sharing policy of each dataset is defined. By default, the data will be made available for reuse. If any

constrains exist, an “embargo period” or “restricted flag” will be explicitly raised in this section of Table 1.

- The data producers will make their data available for third-parties within public repositories only for scientific publications validation purposes.

4 Allocation of resources

In order to face the data management challenges efficiently, all MyCorridor partners have to respect the policies set out in this DMP and datasets have to be created, managed and stored appropriately. This Chapter identifies MyCorridor roles related to the management of the data and their responsibilities. These are: a) the data controller, b) the data producer and c) the data manager.

The **data controller** acts as the point of contact for data protection issues and will coordinate the actions required to liaise between different beneficiaries and their affiliates, as well as their respective data protection agencies, in order to ensure that data collection and processing within the scope of MyCorridor, will be carried out according to EU and national legislation. The data controller must ensure that data are shared and easily available.

Given the cross-border and the different pilot sites that are considered in MyCorridor, the consortium investigates who is the most appropriate partner to take the role of the data controller. This role may be undertaken either by the pilot leader of each specific pilot site or the overall data platform integrator (CERTH/ITI) or by each data provider per se depending on the data to be collected and or generated. The name(s) of the data controller(s) will be announced in the next DMP version (in D2.2 "MyCorridor interoperable, open and seamless architecture and MyCorridor subsystems and modules specifications" in M24 – May 2018).

The **data producer** is any entity that produces data within MyCorridor scope. Each data producer is responsible for the integrity and compatibility of its data during the project lifetime. The data producer is responsible for sharing its anonymised datasets through open access repositories, according to the principles and mechanisms defined in the current document. He is in charge of providing the latest version.

Last but not least, the **data manager** (to be announced in the next version of the DMP) will coordinate the actions related to data management, will be responsible for the actual implementation of the DMP successive versions and for the compliance to Open Research Data Pilot (ORDP) guidelines. As the MyCorridor open data will be hosted either by institutional databases or by an open free of charge platform (e.g. Zenodo), no additional costs will be required for hosting the data.

All research entities participating in the MyCorridor project shall ensure that they have entered into an appropriate data sharing agreement prior to any personal data being shared.

5 Data Security

MyCorridor open cloud system will provide out-of-the-box security mechanisms and management procedures so as to a) ensure personal (sensitive) data protection through a strict process of data collection, anonymization, harmonization and integration and b) guarantee data integrity and reliability, ensuring system's high performance operation through the exchange of the necessary information.

The consortium research partners will fully comply at all times with all applicable data protection legislation and regulation during this project, to ensure the security and protection of individuals' personal information in relation to this project. This includes compliance with the General Data Protection Regulation (GDPR), which comes into effect on 25 May 2018. The consortium and research partners acknowledge the various new obligations and the new rights granted to data subjects under the GDPR and are aware of the significant fines that may be imposed should a data breach occur.

In terms of **personal data protection**, personal data will be anonymised and strictly used for project's purposes. Before collecting any personal data, the Local Ethics Representative (see Chapter 3 of [2]) will be responsible for informing the involved pilot users/participants and collecting their informed consents (see Chapter 6.2) that will be maintained and stored based on the Grant Agreement rules and European/local laws. No personal data will be centrally stored, without anonymisation or pseudonymisation. No personal information will be made available by the Local Ethics Representative to the pilot sites, i.e., MyCorridor partners participating in the pilots. Only one person per site (the Local Ethics Representative) will have access to the informed consent form containing the personal information and only that person will be aware of the relation between the participant's unique identifier code and their personal identity, in order to administer the tests. In practice, the Local Ethics Representative will collect those data required for contacting the participants and arranging with them the sequence of the current or future tests. The Local Ethics Representative will then issue a single Test ID (unique identifier code) for each of them. This person (Local Ethics Representative) will not participate in the evaluation and will not know how each user behaved. One month before the end of the project, this reference, i.e., the reference between the Test ID and the real-life contact details of the participant, together with any other personal information held on the participant will be deleted, thus safeguarding full anonymisation of the results.

The stored data will refer to a user's age, gender, nationality and preferences for travelling and commuting but this information will be safeguarded, stored and processed only in accordance with all applicable data protection laws and regulations. The stored data will not contain any other identifier apart from the Test ID. In no circumstances will a participant be asked for information relating to their beliefs, political or sexual preferences. User-related data will be securely and safely stored. Also, data will be scrambled where possible and abstracted to permit its use to achieve project outcomes while ensuring data integrity and security.

Any party which provides any data or information (the "Providing Party") to another party (the "Receiving Party") in connection with the project will not include any personal information relating to an identified or identifiable natural person or data subject. To this end, the Providing Party will anonymise or pseudonymise all data delivered to other parties to an extent sufficient to ensure that a person without prior knowledge of the original data and its collection cannot, from the anonymised or pseudonymised data and any other available information, deduce the personal identity of participants.

Each party shall be solely responsible for the selection of specific database vendors/data collectors/data providers, and for the performance (including any breach) of its contracts between it and such database vendors/data collectors, (to which no other project partner shall be a party, and under which no other partner assumes any obligation or liability) and shall further warrant that it has the authority to disclose the information, if any, which it provides to the other parties, and that where legally required and relevant, it has obtained appropriate informed consents from all the individuals involved.

Partners supplying special data analysis tooling, shall have the right on written notice and without liability to terminate the license that it has granted for such tooling to be used in connection with the project, if the supplying partner knows or has reasonable cause to believe that the processing of particular data through such tooling infringes the rights (including without limitation privacy, publicity, reputation and intellectual property rights) of any third party, including of any individual.

Each pilot site will have its own Ethics Committee and one person will be nominated per site as responsible for following the project's Ethics Management Panel (EMP) recommendations and data protection (see [2] for more).

In terms of **privacy of MyCorridor system**, the following rules apply:

- All required user data will be stored at their profile and be securely protected by the relevant WP2 mechanism. Relevant preferences relate to their transportation modes and everyday mobility and transfer preferences. The user will have the option to set or delete their profile. No identification data will be stored, they will all be anonymised and aggregated and will only serve analysis purposes;
- The user's location and route will be only temporarily stored (i.e., during a trip), in order to assist the user and the system to provide the appropriate mobility product; they will be automatically deleted afterwards, unless the user wishes to store them;
- The user will have the capacity to view, change or delete- as he/she wishes- all stored data by the system (including their profile data, if chosen to be stored).

Moreover, a Data Privacy Impact Assessment (DPIA) will be used to support this DMP. In particular, the DPIA will assist with, and demonstrate, compliance with the GDPR. This will be carried out prior to embarking upon the project and as the project evolves it will demonstrate compliance in the collection and processing of personal information. This DPIA will ensure that collection, handling, use and storage of personal information are reviewed to ensure compliance with all applicable privacy legislation and regulations. The DPIA will:

- assess whether the proposed collection, handling use and storage is legally permissible;
- assess whether the proposal is justified and proportionate;
- detail the actions necessary to ensure collection, handling, use and storage is compliant and to mitigate any risk to personal information; and
- provide a record of the decisions that are made to introduce or change the way in which personal information is collected, handled, used or stored.

An indicative template of the DPIA to be used for MyCorridor can be found in ANNEX 2.

Last but not least, whenever **authorisations for data collection**, processing and management have to be obtained from national bodies, those authorisations shall be considered as documents relevant to MyCorridor. Copies of all relevant authorisations and approvals shall be submitted to the Commission prior to commencement of the relevant part of the research work. The cases where such authorizations are needed will be identified as the pilots become more detailed while a list of the needed authorizations and the partners responsible to obtain each of them will be reported in the next release of MyCorridor DMP.

6 Ethical aspects

6.1 Ethical and legal issues related to data sharing

The project involves data collection in the context of user testing and demonstration activities. For this reason, human participants will be involved in certain aspects of the project and data concerning their profile, their preferences and driving/riding behaviour, their usage history of searching and selecting services will be collected. Given that these data are considered personal (even sensitive in cases such as the health status), the core ethical/legal issues within MyCorridor related to data collection and sharing are:

- Privacy protection and confidentiality;
- Informed consent;
- Incidental findings;
- Transparency of the collected data management by the final system and during its pilots;
- IT-Security and identity management;
- Risk assessment (Insurance);
- Delegation of control; and
- Incentives (Financial inducements, etc.).

The proper management of these issues is carefully investigated and monitored within WP10 and A9.3 led by the Ethics Manager and supported by Ethics Board. All relevant principles and the main procedures regarding privacy, data protection, security, legal issues and ethical challenges are defined in the Project's Ethics Manual [1][2] and will be updated in their upcoming versions. The described procedures have been drafted and will be updated in consultation with the project's Ethics Management Panel (composed of one external member, the Coordinator, the Technical & Innovation Manager and the Quality Manager) that will act as supervisors of the ethical activities of the project and the local ethics committees at each pilot site, in order to take into account both European and national ethical and legal requirements.

6.2 Informed Consent

MyCorridor scenarios will target participants with competence to understand the informed consent information. Pilot sites, i.e., MyCorridor partners participating in the pilots, will receive only anonymised and coded or pseudonymised information. Any recorded data will be available to pilot sites only in anonymised format.

The informed consent form, which each participant will be asked to complete prior to their participation in the pilots, aims at ensuring that the user accepts participation and is informed about all relevant aspects of the research project; it will be collected in written form after the users have been provided with clear and understandable information about their role (including rights and duties), the objectives of the research, the methodology used, the duration of the research, the possibility to withdraw at any time, confidentiality and safety issues, risks and benefits.

The basic elements of the MyCorridor informed consent include:

1. The objective of the study, its duration and procedure
2. Possible risks, discomforts and side-effects
3. Privacy and data protection procedures
4. The possibility to decline the offer and to withdraw at any point of the process (and without consequences)
5. Contact person

All test volunteers will receive detailed oral information. In addition, they will receive in the language of the country conducting the test pilot:

- a commonly understandable written description of the project;
- the project goals;
- the planned project progress;
- the related testing and examination procedures;
- advice on unrestricted disclaimer rights on their agreement.

The informed consent process and forms are described/provided in detail in the project Ethics Manual [2].

7 Conclusions

This deliverable provides an overview of the data that MyCorridor project will produce together with related legal and ethical data processes and requirements that need to be taken into consideration.

Also, it document outlines an overview about the dataset types, defines a set of attributes to be used for describing each dataset, and presents the open access aspects to be followed by the consortium. The datasets will be incrementally enriched along the project lifetime. These descriptions include a detailed description, standards, methodologies, sharing and storage methods.

MyCorridor decisions with respect to making the data FAIR and the respective mechanisms to support these decisions are described in Chapter 3 while the allocation of resources and the data security aspects are presented in Chapters 4 and 5 accordingly.

Last but not least, Chapter 6 summarises the related ethical aspects with respect to e.g., (sensitive) personal data, informed consent, restrictions and constraints of contacting and testing MyCorridor services with real life users, etc., as defined in the ethical framework and policy of MyCorridor [1][2] taking into account European and national/local ethical guidelines and legislation.

The current document is a living document and thus it will be updated during the project lifetime as needed, including more detailed information regarding the collected data. The next official updated version will be publically released as part of D2.2 “MyCorridor interoperable, open and seamless architecture and MyCorridor subsystems and modules specifications”, i.e., M24 (May 2018), providing more information to aspects such as:

- The descriptions of the different datasets, including their reference, file format, standards, methodologies and metadata and repository to be used;
- Institutional repositories in cases where the project partners maintain such a repository;
- More details about license types and rules for the data producers so as to allow the widest reuse possible for the collected data;
- The updated list of services decomposition and their clustering (Annex 1) and the mapping of MyCorridor services to them;
- The final DPIA form (Annex 2);
- The name(s) of the data controller(s) and the data manager – see Chapter 4.

Still, intermediate internal (unofficial) versions of the DMP will be released before the two Pilots rounds, on M18 and M28 respectively.

References

[1] Fairbairn, B. (2017). POPD – Requirement No.1, Deliverable 10.1, MyCorridor (Mobility as a Service in a multimodal European cross-border Corridor) project (G.A.: 723384), http://mycorridor.eu/

[2] Dovinola, G. (2017), MyCorridor Ethics Manual, Deliverable 9.2, MyCorridor (Mobility as a Service in a multimodal European cross-border Corridor) project (G.A.: 723384), http://mycorridor.eu/

[3] European Commission, Directorate-General for Research & Innovation, "Guidelines on FAIR Data Management in Horizon 2020", available at http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf [last accessed: Nov. 29th, 2017]

[4] Franchi, L., Tarle M. (2017). Dissemination strategy and actions (1), Deliverable 8.2, MyCorridor (Mobility as a Service in a multimodal European cross-border Corridor) project (G.A.: 723384), http://mycorridor.eu/

ANNEX 1: Tentative list of services and service clustering

In the context of MyCorridor, there are two main terms used: a) the mobility products and b) MyCorridor services.

Mobility products refer to real life, physical transportation services or transportation management provided by Transportation companies or Authorities; they might be sold to travellers in the form of ticket products, which are based on tariff policy. "Traffic Management" is included although not typical "product"/"service".

MyCorridor services are digital services which will be provided by MyCorridor application and are clustered according to their nature and objectives as follows:

1. **Mobility:** Services related to the on line purchase of Mobility Products, which are available for purchase via the MyCorridor application. They are further divided into the following sub clusters:

   1.1. *Vehicle related:* MyCorridor services supporting purchase of Mobility Products for private cars;

   1.2. *Sharing:* MyCorridor services supporting purchase of sharing Mobility Products;

   1.3. *Public Transport (Para transit):* MyCorridor services supporting purchase of paratransit Mobility Products; and

   1.4. *Public transport:* MyCorridor services supporting purchase of Public Transport Mobility Products.

2. **Traffic management:** Services related to the on line purchase of Traffic Management related Mobility products and/or the use of advanced Traffic Management concepts in the MaaS framework. They are divided into the following sub clusters:

   2.1. *TM2.0:* Relevant TM2.0 use cases incorpoated into the My Corridor application;

   2.2. *Access control & Tolling:* MyCorridor services supporting purchase of traffic/demand management products (such as tolls, urban congestion pricing, zone access control); and

   2.3. *Cooperative Intelligent Transport Systems (C-ITS):* Relevant C-ITS use cases incorporated into MyCorridor application.

3. **Infomobility:** Services related to the information and support of the user in order to plan a trip, decide what mobility product to buy as well as real time support while on trip. They can be related to Mobility Products sold by MyCorridor or for any other Mobility service/product not currently supported by MyCorridor. In the latter case, the user just gets information/guidance without the possibility to buy those Mobility services. They are divided into the following sub clusters:

   3.1. *Multimodal:* MyCorridor service combining multi modal information in a single feedback to the user;

   3.2. *Public Transport:* MyCorridor services supporting use of Public Transport Mobility Products;

   3.3. *Park & Ride:* MyCorridor services supporting use of Park and ride Mobility Products;

3.4. *Traffic:* MyCorridor services supporting use of car; and

3.5. *Other*.

4. **Added Value:** Services giving added value to the user and enhancing user experience. These may include:

4.1. *Touristic/Entertainment:* Services related to the supply of touristic/entertainment information and booking facility to users;

4.2. *Horizontal:* Standard and embedded MyCorridor services which are the same for all MyCorridor aggregators; and

4.3. *Synthetic:* Services that result as a synthesis of independent services through the use of the business rule editor.

Based on the above provided terminology, Table 2 provides a provisional categorization of MyCorridor services correlated with the mobility products to be provided by MyCorridor. Changes may arise as the project evolves and the description of the integrated services become more detailed.

Table 2. Provisional categorization of MyCorridor services.

Service cluster	Sub-cluster	Mobility Products (offered to MyCorridor users)	Services available through MyCorridor Application
Mobility	Vehicle related	Parking	Pre-purchase tickets
			Purchase e-tickets
			Booking parking space
			Parking availability information
	Vehicle related / Public Transport	Park and ride	Pre-purchase combined, and/or parking and/or PT tickets
			Purchase e-tickets
			Booking parking space
			Parking availability information
			PT scheduled information
			PT real time information
	Vehicle related / Sharing	Car sharing	Booking shared car
			Pre-purchase shared car ticket
			Purchase shared car e-tickets
		Car-pooling/ Ride sharing	Ride sharing apply and book
	Public Transport (paratransit)	Taxi	Taxi apply and book
	Public transport	Urban PT	Pre-purchase tickets
			Purchase e-tickets
			PT scheduled information
			PT real time information

		Interurban PT (train, maritime, bus)	Pre-purchase tickets
			Purchase e-tickets
			PT scheduled information
			PT real time information
	Vehicle related/ Public Transport	Ferry boat booking/ticketing	Route planning
			Booking
			Purchase tickets
			PT scheduled information
	Sharing	Bike sharing	Booking shared bike
			Pre-purchase shared bicycle ticket
			Purchase shared bicycle e-tickets
Traffic management	TM2.0	Adaptive real-time traffic management	Real time traffic state and forecast
			Event Management
			Advanced Traffic Forecasting provision
	Access control & Tolling	Zone access control	Zone access control information provision
	C-ITS		GLOSA, Traffic light status, traffic light forecast
			Traffic events
Infomobility	Multimodal		Multi modal journey planner
	Public Transport		Multi-modal service real time information (Urban PT, Ferry boat, Train, interurban bus)
	Park & Ride		Real time information for parking availability and PT estimated time of arrival
	Traffic		Traffic Flow
			Traffic Incident
			Adaptive navigation
	Other		"Smart Mobility'
Added Value	Touristic/ Entertainment		POIs
			Car rentals
			Hotel
			Events (concerts etc)
	Horizontal		Cross border travel palnning
			Mapping
			eco driving

	Synthetic		Discounts based on proposed business rules
			Synthesis of new mobility products
			Loyalty

ANNEX 2: Tentative DPIA form

What is a PIA?

A PIA will be required under Article 35 of the General Data Protection Regulation (EU) 2016/679. A PIA is a process which helps assess privacy risks to individuals in the collection, use and disclosure of information. PIAs help identify privacy risks, foresee problems and bring forward solutions.

Why should I do a PIA?	When should I start a PIA?
• To identify privacy risks to individuals. • To identify privacy and data protection compliance liabilities for your organisation. • To protect your reputation. • To instil public trust and confidence in your project/product. • To avoid expensive, inadequate “bolt-on” solutions. • To inform your communications strategy.	PIAs are most effective when they are started at an early stage of a project, when: • the project is being designed; • you know what you want to do and how you're going to do it; • you know who else is involved. But ideally it should be started before: • decisions are set in stone; • you have procured systems; and • you have signed contracts/ MOUs/agreements.

Do I have to do a PIA?

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?
- Will the project compel individuals to provide information about themselves?
- Will information about individuals be disclosed to organisations or people who have not previously had routine access to the information?
- Are you using information about individuals for a purpose it is not currently used for, or in a way it is not currently used?
- 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.
- Will the project result in you making decisions or taking action against individuals in ways which can have a significant impact on them?
- 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.
- Will the project require you to contact individuals in ways which they may find intrusive?

Carrying out a PIA

This template is an example of how you can record the PIA process and results. You can start to fill in details from the beginning of the project, after the screening questions have identified the need for a PIA. You can adapt the process and this template to produce something which allows your organisation to conduct effective PIAs integrated with your project management processes.

1. Identify the need for a PIA

2.1. Explain what the project aims to achieve, what the benefits will be to the organisation, to individuals and to other parties.

2.2. You may find it helpful to link to other relevant documents related to the project, for example a project proposal.

2.3. Also summarise why the need for a PIA was identified (this can draw on your answers to the screening questions).

Please fill this out in light of the questions you answered 'yes' to above and provide further information.

E.g. The project involves [X] sharing personal data with [X]. [X] will also share personal data about [X individuals] with [X] organisation. The overarching purpose is [XYZ]. The benefits of collecting and processing the personal information is [X].

The relationship between [X] and [X organisation] is [X] and [explain the role each party is playing and their responsibilities e.g. X organisation is delivering an IT system or [X] is providing research services].....etc.

2. Describe the information flows

2.1. The collection, use and deletion of personal data should be described here (e.g. where you are getting the data from, where it will be stored and where it could be transferred to). You should also say how many individuals are likely to be affected by the project. It may also be useful to refer to a flow diagram or another way of explaining data flows.

E.g. Data will be collected from research participants by [X] via [online] forms

↓

Data will be stored encrypted on departmental drives

↓

Pseudonymised dataset will be provided to [Department X] etc.

3. Consultation requirements

3.1. Explain what practical steps you will take to ensure that you identify and address privacy risks. Who should be consulted, internally and externally? How will you carry out the consultation? You should link this to the relevant stages of your project management process.

3.2. Consultation can be used at any stage of the PIA process.

E.g. Discussed storage with Information Security Team.

4. Identify the privacy and related risks

4.1. Identify the key privacy risks and the associated compliance and corporate risks. Larger-scale PIAs might record this information on a more formal risk register.

Privacy issue	**Risk to individuals**	**Compliance risk**	**Associated organisation / corporate risk**
E.g. *1. Risk that the security of the data is compromised.*	*Risk that sensitive personal data is lost or stolen or destroyed causing distress or damage to the data.*	*Risk of breach of data protection legislation.*	*Risk of reputational damage to entity/entities involved and of enforcement action being brought. Risk to delivery of research objectives both current and in the future. Risk of complaints or litigation from affected individuals.*
2. Risk that personal data is retained for longer than is necessary.	*Risk that individual's data is held for longer than is required and that security and other organisational methods applied to the personal data lapse.*	*Risk of breach of data protection legislation.*	*As above.*

5. Identify privacy solutions

5.1. Describe the actions you could take to reduce the risks, and any future steps which would be necessary (e.g. the production of new guidance or future security testing for systems).

Risk	**Solution(s)**	**Result:** is the risk eliminated, reduced or accepted?
Risk '1' above.	*Encryption measures are used.*	

Risk '2' above.	*Appropriate retention periods have been agreed.*	

6. Sign off and record the PIA outcomes

6.1. Who has approved the privacy risks involved in the project? What solutions need to be implemented?

Risk	Approved solution	Approved by
E.g. Risk 1	*Data will be deleted when it is no longer necessary to retain such data.*	*E.g. Data Protection Officer.*

7. Integrate the PIA outcomes back into the project plan

7.1. Who is responsible for integrating the PIA outcomes back into the project plan and updating any project management paperwork? Who is responsible for implementing the solutions that have been approved? Who is the contact for any privacy concerns which may arise in the future?

Action to be taken	Date for completion of actions	Responsibility for action
Data to be deleted.	*Insert date/description of when.*	*E.g. Data Protection Officer.*

Contact point for future privacy concerns
E.g. Data Protection Officer's details.

Next steps

It is recommended that the PIA is signed off at a senior level internally.

There is no strict requirement to file or be able to produce a PIA report but, if privacy concerns arise, it is good practice to be able to do so. We would recommend that the PIA is filed and stored internally.

Any questions?

If you are still unsure about whether you need to carry out a PIA or have any questions about the guidance above your first point of contract should always be your Data Protection Officer:

Name	
Position	Data Protection Officer
Telephone	
E-mail	