UNLOCKING LARGE-SCALE ACCESS TO COMBINED MOBILITY THROUGH A EUROPEAN MAAS NETWORK.

Deliverable D3.4
Data Sharing in MaaS: assessment of opportunities and barriers

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 723314.

Deliverable D3.4

Data Sharing in MaaS: assessment of opportunities and barriers

Due date of deliverable: 30/11/2019

Actual submission date: 20/12/2019

Dissemination Level		
PU	Public	X
CO	Confidential, restricted under conditions set out in Model Grant Agreement	
CI	Classified, information as referred to in Commission Decision 2001/844/EC	

Start date of project: 01/06/2017

Duration: 30 months

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 723314.

Document Control Sheet

Deliverable number:	**D3.4**
Deliverable responsible:	MOSAIC FACTOR
Work package:	WP3
Main editor:	Stefano PERSI

Editor name	Organisation
Stefano PERSI	MOSAIC FACTOR
Burcu KOLBAY	MOSAIC FACTOR
Lidia SALA	MOSAIC FACTOR
Vassilis PSALTOPOULOS	Institute of Communication and Computer Systems
Markus AARFLOT	UBIGO Innovation AB
Steven SARASINI	RISE Research Institutes of Sweden (Viktoria)

Document Revision History			
Modifications Introduced			
Version	**Date**	**Reason**	**Editor**
1.0	01/06/2018	First version	Stefano PERSI
1.1	24/07/2019	Content addition of the document	Vassilis PSALTOPOULOS
1.2	02/09/2019	Chapter 3	Lidia SALA
1.3	03/09/2019	Chapter 9	Vassilis PSALTOPOULOS
1.4	18/11/2019	Integration of partners' feedbacks and Chapter 4, 5, 6 ,7 and 8	Burcu KOLBAY
1.5	22/11/2019	General Content addition	Markus AARFLOT
1.6	27/11/2019	Updating of "Executive Summary" and Chapter 13	Vassilis PSALTOPOULOS
1.7	05/12/2019	Updating of Chapter 12	Steven SARASINI
1.8	05/12/2019	Integration of partners' feedbacks	Burcu KOLBAY

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 723314.

1.9	13/12/2019	Updating of Chapter 13	Stefano PERSI
1.10	18/12/2019	First quality peer review round	Fabio CARTOLANO
1.12	20/12/2019	Deliverable finalization	Stefano PERSI

Legal Disclaimer

Executive Summary

This Deliverable is aiming at summarising the findings related to Task 3.2 “Data Sharing obstacles and opportunities”, which is addressing particularly the non-technical issues concerning data collection, processing and sharing (such as data privacy, security, visibility of internal business, “fears”, trust, regulatory frameworks, etc.). An important issue is to define which is the minimum amount of data that the MaaS actors should share inside the IMOVE ecosystem. In a first stage, the high priority of T3.2 was to identify which are the possible motivations, barriers and triggers for data sharing, and which were the possible data of interest for IMOVE (i.e. classification based on demand, offer and context) and what could help unlocking the data sharing that is currently not happening in most cases, beyond Open Data. Meanwhile, WP1, WP2 and WP4 progressed on the identification of Requirements, Scenarios and High-Level Functionalities, and provided more elements to carry on the data sharing study at a lower and more specific level.

We see from this analysis that in most cases the different actors, both final users and MaaS or (SW and/or Transport) Service providers, are often more interested in elements that goes beyond “selling data for cash”. Understanding their needs, fears and interests will be the key to unlock the full potential of MaaS through (controlled) data sharing.

This deliverable also tackles data privacy and security related matters and explains whether of how data are shared within the project partners or third parties, how the data are legally collected, what are the mechanisms and precautions taken for secure data storage, and how the project’s actions are in complete compliance with the GDPR and previous Privacy and Data protection legislation.

Since IMOVE by nature, is a project that relies on and handles/manages personal data to extract all relevant results and operate, focus is also drawn on the participation of the research subjects, and more precisely on how the latter are informed about the provision of personal data upon their involvement, following certain processes. The project’s consortium ensures the protection of personal data and guarantees that the data subjects and their underlying data provided on the platform are protected with the adoption of specific techniques and principles in full compliance with the GDPR.

The rest of the document is structured as follows: Introduction section aims to frame the document in the wider IMOVE project, providing information on the document purpose. Chapter 1 is to show the motivation of knowing the importance of data. Chapter 2 gives information about the drivers of data sharing, whereas Chapter 3 states the barriers of data sharing. In Chapter 4, the data came from LLs are listed to show the data sharing situations within/between the MaaS services with or without personal data. Chapter 5 is to highlight the data privacy and security issues. Chapter 6 focuses on the personal data and its related issues. Chapter 7 describes the opportunities of data sharing while the focus of Chapter 8 is the possible triggers of data sharing. Finally, Chapter 9 summarizes the main findings and outcomes of the deliverable.

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 723314.

Abbreviations and Acronyms

B2B	Business to Business
D	Deliverable
EC	European Commission
EMT	Empresa Municipal de Transportes de Madrid
ERB	Ethics Review Board
EU	European Union
GDPR	General Data Protection Regulation
ICT	Information and Communication Technology
IPR	Intellectual Property Rights
IR	Internal Report
LL(s)	Living Lab(s)
MaaS	Mobility As A Service
PT	Public Transportation
WP	Work Package

Table of Contents

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 723314.

List of Figures

List of Tables

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 723314.

INTRODUCTION

This deliverable is aiming at focusing, investigating and later, summarising data sharing obstacles and opportunities within the IMOVE project by giving the main attention to the non-technical issues related to data collection, data processing and data sharing in the light of several concepts such as privacy, security, fears, trust and regulatory frameworks. Another focus here is to define the minimum amount of data that the MaaS actors should share inside the IMOVE ecosystem in such a way at the end, that necessary and useful amount of information and knowledge would be reached by data analyses. This issue should be addressed both for the data shared between actors participating in local MaaS operations of the Living Labs and also for the data shared between different MaaS schemes, for example across the Living Labs. Related to that, the task will define which blocks of data are provided by which actor. Using the living labs usage and struggle with data as input to the report, it is able to draw conclusions about how data can be shared to facilitate MaaS development. A special mention is how the user will provide personal data required for the evolvement of the MaaS ecosystem and how protection of personal data should be handled through compliance with all safety standards and requirements. The initial results of this task were reflected in IR3.2 which was available during the early stages of the project, while the final results consist of this document with experience obtained from the Living Labs after experiencing IMOVE.

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 723314.

1 MOTIVATION

A common view is that “data” is important in all parts and functions of society. However, data is often referred to as a general term without detailed specification. With regard to mobility, transport providers are those actors with the most comprehensive access to data. Yet, they can access only the data within their well-defined area of mobility, which is only a fraction of the mobility data that exists. Such structure and lack of mobility data access prevent any actor from gaining a holistic view of transportation systems. In addressing the importance of mobility data, one can distinguish between three different categories:

- **DEMAND-SIDE DATA**: This type of data is related to travel habits, motives, needs and preferences. Demand-side data can be further distinguished in terms of that which refers to customers and that which refers to individual users and travellers, when the two are not the same actor (e.g. with regard to B2B services where the actor which purchases a MaaS service is a public or private organisation). In both cases, demand-side data are useful for a number of reasons. It can be utilised by service providers to further develop and refine their mobility service offers; for marketing and recruitment purposes, and for mobility management within organisations. Demand-side data can also be used within different types of transport system governance, such as the creation of incentives and policy instruments to promote or penalise particular types of travel behaviour, for traffic management, within urban development, to develop parking policies, and so on. Further, there is a palpable need for demand-side data related to the way in which MaaS services influence travel behaviour in order to resolve key uncertainties related to the business case for MaaS, willingness to pay, market demand, travellers’ perceptions of affordability, quality and accessibility, and concomitant sustainability impacts.
- **SERVICE DATA**: this type of data relates to the availability of different mobility services at any given time and in any given location. With regard to public transportation, service data refers to timetables, available seats, accessibility for disabled travellers, trip-planning, ticket pricing and availability, and so on. With regard to other modes such as car clubs and bicycle pools, service data refers to vehicle availability, pick-up and drop-off locations, pricing, and so on. With regard to MaaS, service data are inherently more complex due to the multi- and inter-modality of the service, refers to trip- and route-planning, bookings and payments for different mobility services, modal availability, transfer information, pick-up and drop-off locations, customer support, and potentially information on rewards, games and incentive schemes that aim to promote particular modes of travel.
- **CONTEXTUAL DATA**: This type of data refers to the prevailing traffic conditions in a given area at a certain time, weather conditions, and information regarding any potential bottlenecks or congestion caused by special events (e.g. football matches, concerts, marches and protests).

Data may be further distinguished in terms of whether it is momentaneous or cumulative data. That is, momentaneous data refers to demand-side data, service data and contextual data that are available on an immediate, real-time basis. Immediate demands for particular transport services and modes, the location of travellers and their means of transportation, timetable delays and traffic queues are all examples of momentaneous data. Cumulative data are recorded over a longer period, and refers to things like travel patterns, accumulated travel needs and preferences, average service delays and timetabling, and so on. Cumulative data can be used to gain a more comprehensive overview of demand-side, service and contextual data. Data is as defined dependent on the context and in order to comprehend the challenges related to data sharing it is important to understand the context in which the data are created and used. The concept of a MaaS ecosystem is often used to name the actors involved in different aspects of mobility as a service.

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 723314.

An ecosystem is according to Moore (1996) "An economic community supported by a foundation of interacting organizations and individuals—the organisms of the business world. The economic community produces goods and services of value to customers, who are themselves members of the ecosystem. The member organisms also include suppliers, lead producers, competitors, and other stakeholders. Over time, they coevolve their capabilities and roles, and tend to align themselves with the directions set by one or more central companies" [3].

Developing MaaS services requires the creation of new business ecosystems. A generic MaaS ecosystem can be found in **Errore. L'origine riferimento non è stata trovata.**. Within these ecosystems, existing actor r oles and value chains must be reconfigured in order to deliver MaaS. Generally, a MaaS ecosystem includes: several transport providers, positioned upstream in the value chain; an ICT platform that facilitates bookings, payments and revenue distribution; and a MaaS service provider that bundles and repackages individual mobility services, integrates information and payment services, and delivers these to end users. All of which create different subsets of data and whose collaboration and sharing is critical to the success of MaaS.

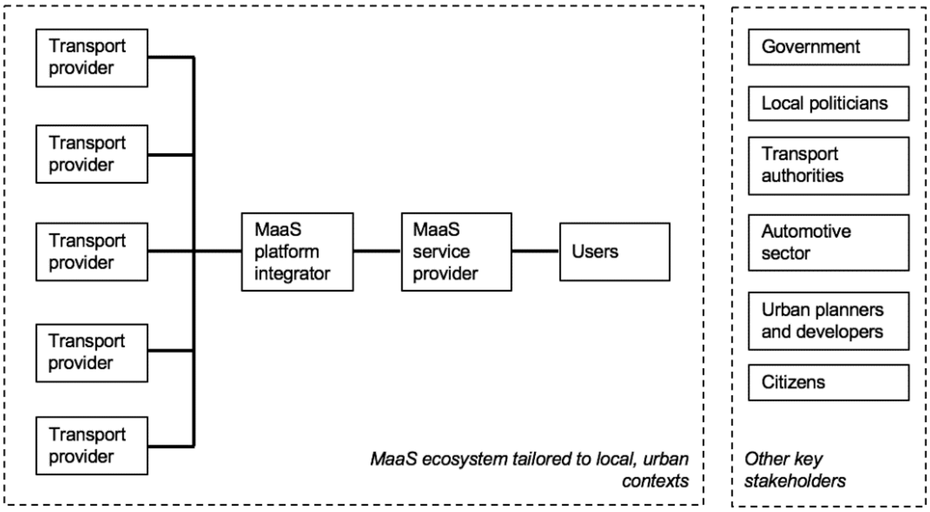

Figure 1. A generic illustration of a MaaS ecosystem.

2 DRIVERS OF DATA SHARING

In many aspects, data sharing is the core foundation of any movement. Consider the momentaneous data required for a person to move from A to B. Data describe the need and the actual usage for a movement, preferred mode of transport, route, time spent on movement, pace and much more. The sharing of this data is what enables a successful movement and adaptation of this movement from A to B from a service perspective. It is also what enables continuous improvements and the development of efficient and adapted transportation services. The drivers of data sharing are thus to an extent an intricate part of any business model logic. Knowledge about customer behaviour both momentaneous and cumulative is one example of central business model information. When designing a successful MaaS business model, data find a place as an essential part of all the key elements. However, the dominating business models in transportation today have all developed in isolated verticals separated per mode of transport. In each vertical, the data sharing is often functional, and it contributes to a more successful business model by generating insights into critical features such like value proposition and functionalities.

With regard to MaaS, opportunities can only be realised through data coordination within the entire ecosystem and among actors that have not previously cooperated. A more nuanced understanding of demand-side, service and contextual data are critical to the development of high quality MaaS services that target user needs in an effective manner and only achievable through new means of data sharing. Data sharing is also crucial for establishing the basis for roaming services across regions and countries. The driver of data sharing from a MaaS perspective is thus the development of attractive and adaptable services for customers. Furthermore, the cumulated data that are shared and compiled can deliver value to public institutions and be useful for medium- to long-term governance and planning activities since MaaS data better than data from vertical mobility services by providing the entire picture of mobility and give a better understanding of what the cumulated mobility demand is. Data are also a prerequisite in a user perspective in order to provide information which can complement the already present mobility information and as such broaden the perspective on mobility. One such example is that when sharing data within the MaaS ecosystem a comprehensive picture of carbon dioxide emission can be used to nudge users into more sustainable habits.

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 723314.

3 BARRIERS OF DATA SHARING

Each driver is balanced by a risk associated with data sharing. Many of the barriers to data sharing within the MaaS ecosystem are synonymous with the creation of the ecosystem itself. Barriers to data sharing may be divided into the following categories:

1. Perceptions of risk
2. Data collection and quality
3. Security and privacy

3.1 PERCEPTIONS OF RISK

Generally, key players' unwillingness to share data within the MaaS ecosystem rests on a set of perceived risks to collaborating within MaaS ecosystems. The three main perceptions centre on:

- Risks of losing customer relationships
- Risks of cannibalisation
- Risks to existing brands

These three risks are interlinked in such a way that each relates to actors' unwillingness to reposition themselves in the value chain, as reflected by public transport operators' reluctance to allow third-party ticket sales, for example. Generally, existing transport service providers such as public transport, car-sharing and taxi firms pride themselves in being customer-facing organisations with close ties to their own customer base, and in having strong brands. The prospect of third-party operators bundling and repackaging their offers is associated with the perceived risk that existing transport service providers will lose control of their customer relationships and thereby their data and insights on customer behaviour. Unwillingness to share data across the ecosystem is a key element of this hesitant stance: existing transport service providers see demand-side data as a source of competitiveness and a safeguard against cannibalisation. In other words, demand-side data (in particular cumulative demand-side data) are seen as an asset that is difficult to replicate and a means to maintain and consolidate existing market positions. Furthermore, the service data are also often perceived as business sensitive, and entering into strategic collaborations within a MaaS ecosystem is perceived to put that data at risk. Entering into strategic alliances to offer a new service is historically associated with the actual risk of loss of sensitive data but it has to be compared with the potential gain of business driving data.

3.2 DATA COLLECTION AND QUALITY

The novelty of the MaaS concept infers barriers to data collection, in that existing datasets based on traditional transport services are insufficient for the analysis of MaaS. The insufficiency depends on several factors. However, one heavy factor originates from the nature of MaaS. The ambition of MaaS is to create a service for all personal mobility and as such data related to MaaS and the success of a MaaS implementation requires more attention than a normal set of data do. Hence in order to collect data, new methods of data collection must be developed, and data collection must target initiatives, trials and pilots of MaaS in order to ensure data relevance. Although several MaaS pilots are currently underway across the globe, there are notable difficulties in collecting data from these pilots, not least organisational shortcomings linked to the coordination of data collection activities among different actors and the competition for pilot data. Another factor for the challenges in data collection is, particularly for MaaS, the duality of data collection which often is extensive, due to the

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 723314.

size of datasets, in MaaS pilots and thus impedes the service. Correspondingly a great service without comprehensive data has high risks of failure.

Generally, pilots are planned with some sort of data collection in mind, in the form of surveys, focus groups, travel diaries, GPS tracking systems or other qualitative assessments. However, data are not typically collected in a standardised fashion, such that resultant datasets lack commensurability. MaaS pilots also require prolonged planning phases, mainly due to the barriers to ecosystem collaboration described in section 2.1, such that gaining access to data is also a lengthy process. An alternative to accessing data on MaaS is to target data on individual mobility services and create datasets that can be used to examine multimodality using a bottom-up approach. However, such datasets may be incomplete (with some attributes missing, or due to irrelevant datasets) or are in some way incompatible (with different parameters and variables, different timeframes, etc.).

3.3 Data Security & Privacy

A final barrier to data sharing is individuals' concerns about data privacy and integrity. Current trends and societal debates on these issues are compounded by new legislation (GDPR) that restricts the storage of personal data in cases where individuals do not allow organisations to collect their personal data. In cases where organisations do collect and store personal data, GDPR could potentially hinder data sharing among organisations. However, GDPR only dictates that the terms of data usage and storage are clearly communicated and does not dictate that data cannot be shared as long as the data privacy regulation is upheld. Generally, public perceptions and misconceptions of the way in which personal data are utilised serve to confound this issue further by creating a generic reluctance to share their personal data. Especially, MaaS is at risk in terms of data privacy since as pointed out earlier, a MaaS dataset is much more comprehensive in its description and detail of personal mobility compared to previous datasets being compiled and stored in different verticals.

From an organisational perspective, data security and privacy are issues that require sophisticated management and governance practices. In addition to complying with legislation on data security and privacy, organisations must create detailed policies on what types of data will be collected and stored, how they will be stored and how they will be used. With regard to data sharing, organisations must create policies that are in some way future-proof. That is, policies are required that inform individuals on key issues linked to data security and privacy at the point in time when data are collected but must account for unforeseen developments (e.g. new forms of data sharing, legislative developments, etc.). Data security is not yet a major threat within MaaS, at least not in the topic of data sharing, since the data relevant for data sharing often does not contain IPR concerning platform and technology which often is the largest antagonist risk from a data security perspective.

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 723314.

4 Living Labs

The data collected from the LLs will be used for making elaborations. The information and knowledge gained through the LL experience are helpful for quality improvement of services and creating tailored offers to the users in the MaaS ecosystem. The data carries the information and knowledge, of course, anonymised and personalised. The surveys consist of personal habits, behaviours, etc. However, thanks to the anonymisation, it is not possible to identify any participant by the information they provided.

In each LL, the datasets for context, pilot and open data are listed with their description in order to show the data sharing situations and to distinguish the ones that hold personal data from the ones do not. It is to point out the level of knowledge having with and without personal data through the analyses in the LLs. The analyses that have been applied and their corresponding results are places in D3.6.

4.1 Berlin Living Lab

In order to meet the ambition of accelerating the deployment and unlock the scalability of MaaS schemes in Europe, and ultimately paving the way for a "roaming" service for MaaS users at the European level, several LLs are chosen. Berlin as the capital and largest city in Germany by both area and population is chosen as one of the LLs, which gives advantage to the project by having URBI in the urban area of the city.

4.1.1 Data shared

The ambitions of the IMOVE project and Berlin LL are also combined with the ambition of private stakeholders who are the vehicle sharing service providers. The data came directly from the providers' services.

4.1.1.1 Data shared in MaaS schemes inside the Living Lab ecosystem

There are two different categories of data shared within the IMOVE ecosystem:

MaaS Platform	Data Analytics

From Berlin LL (Germany), all datasets from different providers are analysed individually and then, the last results are shaped by having the collective analyses in order to reach the bigger and clearer picture that describes the city of Berlin.

In this LL, vehicle sharing services' context data are provided without any service user-related information. The current amount of data was enough to have aggregated level of analyses such as trip flows between the districts/ neighbourhoods in different hours and some area characteristics. However, it was not possible to have user analysis due to the lack of user-related data. Any conclusion that is obtained does not cover any benefit to individual users. In case that the ambition of the cities is to have micro solutions rather than macro, then the user-related data must be shared anonymously. The data coming from the pilot of Berlin do contain the user information, but there is no destination related data for the trips. It is concluded that, the origin and the destination point of the trips, and the anonymised user-related information is the minimum amount to be shared. The trajectory of the trips can be also shared if the further services will be developed such as routing, etc. On the other hand, it is proven during the project that the complementary open datasets such as population data, demographic data for the districts/neighbourhoods and operational areas of vehicle sharing services increase the accuracy and efficiency of the analyses and their results significantly.

The datasets shared for the Data Analytics task are the following ones:

Table 1. Berlin LL – Context Data

Datasets received		Description
Bike-sharing	**Callabike**	Bike-sharing usage records without user and destination point data.
	Nextbike	Bike-sharing usage records without user and partially destination point data (only some records have destination points)
Car sharing	**Car2go**	Car sharing usage records without user and partially destination point data (only some records have destination points)
	Drivenow	Car sharing usage records without user and destination point data.
	Driveby	Car sharing usage records without user and destination point data.
Scooter sharing	**Emio**	Scooter sharing usage records without user and destination point data.
	Coup	Scooter sharing usage records without user and destination point data.

Table 2. Berlin LL – Pilot Data

Pilot data used		Description
Bike-sharing	**Nextbike**	Bike-sharing usage records without destination point data.
Car sharing	**Driveby**	Car sharing usage records without destination point data.

Table 3. Berlin LL – Open Data

Open data used	Description
Population data	Population data by districts.
District data	Geographical district related areas and the borders (Shapefile/polygon).
Vehicle sharing operating area	Operating areas (Shapefile/polygon) of Callabike, Nextbike, Car2go, Drivenow, Driveby, Emio, Coup vehicle sharing services.

4.1.1.2 Data shared between different MaaS schemes and other Living Labs

The Berlin LL is a part of both the theoretical roaming exercise of IMOVE and the practical. By using the software enablers provided in IMOVE Berlin LL users had the possibility of purchasing public transport tickets in Madrid LL. However, it was only on a demonstration scale and not ready for commercial availability. Nonetheless, the related data were shared in order to enable this ticket transaction using the API interface of Madrid LL.

4.2 Turin Living Lab

Turin as a big city and important business and cultural centre in northern Italy, capital of the Piedmont Region, is chosen as one of the LLs. Their MaaS service is not a commercial, but rather publicly driven service, making it easier to overcome some of the barriers to data sharing.

4.2.1 Data shared

The ambition of the Turin LL is also combined with the ambition of the both public and private stakeholders. A part of the data is collected by the PT company in whole the Piedmont region and then by 5T as supporter of the regional administration, whereas the rest came from the vehicle sharing companies.

4.2.1.1 Data shared in MaaS schemes inside the Living Lab ecosystem

There are two different categories of data shared within the IMOVE ecosystem:

MaaS Platform	Data Analytics

From Turin LL (Italy), the datasets from both public and private stakeholders are analysed individually and then, in a complementary way in order to see the bigger and completed picture that describes the city of Turin.

In this LL, the context data of PT ticket validations were without any user-related information and ticket types. All the ticket ids were used to track the validations without knowing the types of the ticket and anything about their users. The other context data which were for the vehicle sharing records was with origin and destination points but without user-related information. Individual and combined results of these two datasets (context data) were enough to have only aggregated level of analyses such as trip flows between the districts/ neighbourhoods in different hours and some area characteristics. However, it was not possible to have user analysis due to the lack of user-related data. Any conclusion that is obtained does not cover any benefit to individual users. In case that the ambition of the cities is to have micro solutions rather than macro, then the user-related data must be shared. The data coming from URBI for the pilot of Turin do contain the anonymised user information and origin-destination related data for the trips. However, the limitation here is the low number of users and their related records. It was possible to combine or validate the results from the context and pilot data. On the other hand, it is proven during the project that the complementary open datasets such as population data, demographic data for the districts/neighbourhoods, valid ticket types and their prices, operational areas and cost of vehicle sharing services increase the accuracy and efficiency of the analyses and their results significantly. The datasets shared for the Data Analytics task are the following ones:

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 723314.

Table 4. Turin LL – Context Data

Datasets received		Description
Vehicle sharing	**Car2go**	Car sharing usage records without user data (only origin and destination points)
	Enjoy	Scooter sharing usage records without user data (only origin and destination points)
	Mobike	The data of Mobike service that is used to find a bike available. Not possible to distinguish the bikes in the service and the bikes available from one-time step before.
PT Ticket validation data		The PT ticket validations (from Turin PT company fully owned by Turin Municipality) without ticket types for bus, tram and metro modes.

Table 5. Turin LL – Pilot Data

Pilot data used	Description
Data from URBI	Data from the mobility service usages via URBI app, which include user id, provider id, origin and destination coordinates with their corresponding start and end date & time. PT usages do not include the end location related data.
Surveys	Two surveys about the user experience and travel behaviours.

Table 6. Turin LL – Open Data

Open data used	Description
Population data	Population data by districts.
District data	Geographical district related areas and the borders (Shape file/polygon).
PT Ticket Types and prices	Description of main types of PT tickets and their prices.
Vehicle sharing operating area	Operation areas (Shapefile/polygon) of Car2go, Enjoy and Mimoto.
Cost of vehicle sharing services	Average Car2go minutely price.

4.3 Gothenburg Living Lab

Gothenburg as the largest city of Västra Götaland Region is chosen as one of the LLs to establish more sustainable MaaS services including public transport (Västtrafik) in the sense of environmental, social and economic perspectives. The pilot is a purely commercial service which was facilitated by the work of Västtrafik and access to tickets of Västtrafik. Two of the pilots in the Gothenburg lab was with publicly owned subsidiaries and thus one called “park and ride” pilot was able to share data. Also, the commercial project was able to share data on their usage and purchases.

4.3.1 Data shared

In the light of the ambition of IMOVE project and the Gothenburg LL, public actors and research partners collaborate to meet the Gothenburg LL goals. The data came from the parking application of Göteborg Stads Parkeringsbloag.

4.3.1.1 Data shared in MaaS schemes inside the Living Lab ecosystem

There are two different categories of data shared within the IMOVE ecosystem:

MaaS Platform	Data Analytics

From Gothenburg LL (Sweden), the dataset comes from the parking that allows the users to have a park and ride activity that includes parking and PT ticket purchase(s), or only PT ticket purchase(s) without any parking activity. In this LL, it was not possible to obtain more personal or non-personal data to have some additional or complementary analyses. Since there was no PT or vehicle sharing data like in other LLs, it was not possible to have flow analysis of the passengers. The only valuable information was the PT stations and parking facilities connection levels. In this LL, the minimum amount of data must be shared for the flow analysis was not provided.

The datasets shared for the Data Analytics task are the following ones:

Table 7. Gothenburg LL – Pilot Data

Datasets received	Description
Park & Ride PT Ticket Purchases	PT ticket purchases happened via the Gothenburg parking application. The purchases can be related to a park and ride activity or only PT ticket purchase without any parking activity.
Survey	A survey about the user experience and travel behaviours.

4.4 Manchester Living Lab

In order to meet the ambition of accelerating the deployment and unlock the scalability of MaaS schemes in Europe, and ultimately paving the way for a "roaming" service for MaaS users at the European level, several LLs are chosen. Greater Manchester is chosen as one of the LLs to encompass a range of transport modes (the Greater Manchester Transport Authority) and operators through an accessible and integrated platform. The site of this LL is based on and through the Manchester Airport, where the residents, visitors and workers of the airport are involved.

4.4.1 Data Shared

In order to meet the IMOVE ambition, the Greater Manchester Transport Authority (TfGM), Manchester City Council and Manchester Airport Group are involved to investigate the potential opportunities of MaaS systems, the barriers to have a sustainable MaaS system, and the effect on the reduced transport emission and congestion levels. The data of this LL are drawn out of the integrated platform to be analysed and investigated.

4.4.1.1 Data shared in MaaS schemes inside the Living Lab ecosystem

There are two different categories of data shared within the IMOVE ecosystem:

MaaS Platform	Data Analytics

From Manchester LL (England), the datasets come from the integrated platform are analysed individually and then relevant comparison analyses are deployed to have clearer insight about different mobility services.

In this LL, scooter sharing and Metrolink context data are provided without any user-related information. Thus, user-related analyses and related results are not possible to be obtained. The results are in an aggregated level more than having a user focus. On the other hand, the pilot data do contain sensible anonymised private user data that include a historic booking log file of the travels around Grand Manchester area, and also personal mobility surveys for the same group of participants. Although user's ids are anonymised the postal codes for users' work place and households are known. However, this information together with the user preferences (extracted from the surveys) and log of the travels would be the minimum amount of data to be shared for deploying a successful tailored mobility service.

The datasets shared for the Data Analytics task are the following ones:

Table 8. Manchester LL – Context Data

Datasets received	Description
Metrolink Trips	Monthly aggregated Origin-Destination matrices per ticket type from all Metrolink stations.
Mobike Trips	Trips provided by Mobike dockless private bike-sharing service that include bike id, origin and destination coordinates with their corresponding start and end date & time. There is no user-related information provided. This service is closed at the end of 2018.
Drakewell Traffic Data	Estimated, non-specified vehicle trips gathered through beacons' Bluetooth signals distributed along the Great Manchester area. The dataset consists of beacon id and location, vehicle id and date & time information. No further information for the vehicle typology is provided.

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 723314.

Table 9. Manchester LL – Pilot Data

Pilot data used	Description
PT Ticket Purchases via booking application	Usage of the application by several users provides a dataset consists of user id, booking id, journey id, origin and destination coordinates with their corresponding start and end date & time, mode of transport, travel time duration and cost.
Survey	A survey from the commuters working at the Manchester Airport to provide their travel characteristics.

Table 10. Manchester LL – Open Data

Open data used	Description
PT information	Routes, stations and schedules of Bus, Metrolink and train.
Division data	Geographical division levels and their related areas and borders (Shapefile/polygon).
City Facilities	Distribution of public/private facilities and stores.
Indicators	In the fields of socioeconomics and demographics.
Local Link Operation Areas	Geographical areas where local link service is available (Shapefile/polygon).

4.5 Madrid Living Lab

In order to meet the ambition of accelerating the deployment and unlock the scalability of MaaS schemes in Europe, and ultimately paving the way for a “roaming” service for MaaS users at the European level, several LLs are chosen. Madrid, the capital and most populous city in Spain, the third largest city In EU is chosen as one of the LLs for this project.

4.5.1 Data Shared

In order to meet the IMOVE ambition, La Empresa Municipal de Madrid (EMT) is involved to investigate of their current services and their usages, and the data are provided by EMT.

4.5.1.1 Data shared in MaaS schemes inside the Living Lab ecosystem

There are two different categories of data shared within the IMOVE ecosystem:

MaaS Platform	Data Analytics

The analytics set up for Madrid are carried out individually but following a similar methodology structure which can vary due the variety of data information type, making not possible to link the different dataset at micro level. Nevertheless, the outcome insights for each service is compared with the rest of them to give a complete picture as possible of movements through the city.
In this LL, several datasets such as public bus, bike-sharing and MaaS Madrid application usage with private anonymised user data are provided. These data with a proper analysis can extract the users' recurrent places that can be used for recommendations for the user-tailored services.

The datasets shared for the Data Analytics task are the following ones:

Table 11. Madrid LL – Context Data

Datasets received	Description
Cable Car	Daily occupancy and incidences for the cable car service.
Electric EMT	Electric vehicle charging sessions for 7 fast charging points at some of the city parking facilities provide a dataset consists of user id, charging point id, energy consumed, paid price and vehicle type.
Bike-sharing	The trip data of public bike-sharing service Bicimad that operates at docking stations consists of user id, bike id, plug and unplug dock/slot id, start and end trip date & time.
GPS location of PT buses	Bus locations with bus service id and line, bus driver id, closest station, date & time and geocoordinates.
PT Bus Ticket Validations	Bus user validation data with ticket id, type of ticket, bus service id and line, date & time and geocoordinates.

Table 12. Madrid LL – Pilot data

Pilot data used	Description
GPS locations of MaaS application users	If the application is activated, then the encrypted user id, geocoordinates and date & time is stored. No further user activity related information is stored.
MaaS app data	It contains the API calls for different services that offer the MaaS Madrid app. The API is divided into MaaS visualizer and MaaS planner. The services to call are air quality, point of interest of mobility services nearby, autocomplete searching planner, planner, close-by transport stops, alerts of arrivals estimation for some mobility services, nearby vehicles information, routes paths, BiciMad station information and information of PT lines.
Survey	A survey about user experience and travel behaviours.

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 723314.

Table 13. Madrid LL – Open data

Open data used	Description
PT information	Routes, stations and schedules of Bus, metro, train and BiciMad.
City Facilities	Registered properties and their activity.
Indicators	In the fields of socioeconomics, demographics and the distribution of public/private facilities and stores.
District and Neighbourhood Data	Geographical district and neighbourhood levels and their related areas and borders (Shapefile/polygon).

4.5.1.2 Data shared between different MaaS schemes and other Living Labs

The Madrid LL is addressing both the theoretical roaming exercise of IMOVE and the practical local MaaS pilot. By using the software enablers provided in IMOVE Berlin LL users had the possibility of purchasing public transport tickets in Madrid LL. However, it was only on a demonstration scale and not ready for commercial availability. Nonetheless, the related data were shared in order to enable this ticket transaction using the API interface of Berlin LL.

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 723314.

5 Data Privacy and Security

Data security and data privacy are two terms that are very closely linked. Therefore, some specific considerations are necessary in the context of data sharing, in terms of challenges and opportunities. Data security refers to protective digital privacy measures that are applied to prevent unauthorized access to computers, databases and websites, and also protects data from corruption. It is an essential aspect of IT for organizations of every size and type. Examples of data security technologies include backups, data masking and data erasure. A key data security technology measure is encryption, where digital data, software/hardware, and hard drives are encrypted and therefore rendered unreadable to unauthorized users and hackers. Data privacy or information privacy is a branch of data security concerned with the proper handling of data – consent, notice and regulatory obligations. More specifically, practical data privacy concerns often revolve around:

1. Whether or how data are shared with third parties.
2. How data are legally collected or stored.
3. Regulatory restrictions such as GDPR.

All data collected in the context of the IMOVE project (deriving from specific Living Labs) and used beyond the scope of the project are declared (in terms of type and criticality for privacy and business) to the National Authorities/Agencies responsible for “Data Protection”. Every entity involved in handling (personal/sensitive) data in some way has to do this at national level, not only the entities collecting data, but for instance also the ones in charge of storing or processing them. Any data that is being shared is in an anonymized form to allow reuse by the research teams and other third parties. These anonymized data do not allow any individuals and/or their institutions to be identified or identifiable.

A specific mention is necessary concerning the storing of data, both for long-term as well for specific real-time operations. In fact, while for the GDPR the main distinction is when the data are crossing or not the European Borders, in many cases the business-related entities (in particular private companies) are still quite concerned about the crossing of the national borders. Being aware that if the data are stored in the cloud by a national entity, it does not mean that data are stored at a national level, unless specifically agreed with the provider. Even if this seems being more a concern than a real threat, this element is important to be considered and dealt with to improve transparency and increase trust.

As mentioned before in this document, GDPR and previous Privacy and Data protection related legislation regulate extensively how personal data may be collected and processed in the EU. As the first approach, the electronic transfer of information (including personal data) between travellers and transportation services providers have existed for years in the EU under strict compliance of these regulations. In this sense, it can be concluded that the introduction of the new platform will not change the situation in this aspect, providing that this platform will act exclusively as an intermediary between all stakeholders so that all responsibilities for data protection compliance would rely on these stakeholders and not on the platform. This would be of course an acceptable approach from the legal perspective, but in certain extent it would be also missing an opportunity to provide all Web of Transportation clients with a technical solution to deal with privacy and data protection issues at platform level.

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 723314.

Basically, this second approach seeks to take advantage of the role played by the platform by implementing mechanisms to guarantee the compliance with data minimization or purpose limitation principles (e.g. the so-called case of the need to reveal all ID card data just to ensure that someone is a legal adult). Recent research on “Attributes Based Credentials” schemes based on trusted entities could be applied, in case this second approach is eventually followed. This approach, if well designed, can also be helpful to keep track of all personal data transactions so that it would facilitate customers to exercise their withdrawal rights in an effective manner.

Under GDPR, EU citizens will benefit from new or stronger rights, such as being informed about how their data are used, the data portability across service providers, the ability to erase or delete their personal information, the access to the personal data an organization holds about them, the ability to correct inaccurate or incomplete information and over automated decisions and profiling.

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 723314.

6 PERSONAL DATA

According to the EC, “personal data” is any information that relates to an identified or identifiable living individual [4]. Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data. Examples of personal data include name and surname, home address, email address such as name.surname@company.com, identification card number, location data (for example the location data function on a mobile phone), Internet Protocol (IP) address, cookie IDs, the advertising identifier of a phone, data held by a hospital or doctor, which could be a symbol that uniquely identifies a person. Data that are not considered as personal data include a company registration number, email addresses such as info@company.com and anonymised data.

Figure 2. Personal Data

In GDPR’s Article 4, the GDPR gives the following definition for “personal data”: ‘Personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person [5].

In the project’s lifecycle we observed that dealing with personal data it's a very complicated and sensitive topic and in no case is it simple and easy. Project participants that act as data subjects need to be respected and protected by all means. In the following two subsections we will describe how personal data were provided to the project and how the project’s team made sure to protect them.

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 723314.

6.1 PROVISION OF PERSONAL DATA

IMOVE is a project that relies on and handles/manages personal data to operate and extract all relevant results. In this context, we need to make it clear how the subjects will be informed about the provision of personal data and how this process will be done.

At first, it should be noted that the project receives (personal) data primarily through various types of electronic devices such as mobile phones, tablets and laptops by using also their embedded sensors (e.g. GPS).

IMOVE makes use of the data acquired that mostly includes personal data, to extract results about mobility habits, preferences, attitudes and actual movements through remotely collated data: GPS tracking, surveillance cameras, traffic counts etc.

Participation in the research and usage of IMOVE's services is voluntary and users can withdraw their consent (i.e. their data being analysed, reported and stored) at any time by contacting a member of the project team.

The project team, prior to users' participation, requests for consent in an intelligible and easily accessible form that clearly declares the purpose of data processing. The team made sure that consent withdrawal is a procedure as easy as it is to give it.

This way, clear consent for users' perspective is granted to the project team, to store and analyse personal information related to mobility behaviour like user preferences, possible personal mobility restrictions, user positions and user trips. Subjects are aware of their rights to disable their activity tracking at any moment with a specific option in the user interface.

In principle, the project collects and uses the personal data of its users only to the extent necessary to provide a functional user-friendly software application in terms of content and services. The collection and usage of users' personal data take place regularly only in the condition of having the consent of the user. An exception applies in cases where prior consent cannot be obtained for factual reasons and processing of the data is permitted by law.

The project uses software for the analysis of the provided personal data. By evaluating this data, IMOVE obtains important information on users' behaviour that helps to improve the quality of its services and produce targeted offers to maximize user experience on the MaaS paradigm.

Data were collected from the Living Labs, were both anonymous and personalised. Anonymous data were collected to assess impacts in Living Labs (travel patterns, attitudes, car ownership)

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 723314.

based on surveys. Any respondents to surveys was handled in a way so that it is not possible to identify individuals or precise locations and all data were stored as anonymous.

The anonymized data from the Living Labs might be used for future publications and other means of disseminating the findings from the research project. Direct quotations (if any) will be attributed to a pseudonym assigned by the researchers.

Primary data are collected for a specific purpose in IMOVE such as original market research while secondary data involve existing data collected by other bodies (e.g. government) to use for multiple purposes (e.g. population census / travel patterns) which will help to establish local contexts and to ensure full cross-sections are surveyed (age, gender, location, socio-economic status).

All data are processed and evaluated for statistical and scientific purposes, while this is exclusively information that does not enable any association with any person, meaning that all data subjects shall and will remain anonymous.

6.2 PROTECTION OF PERSONAL DATA

The project team in order to ensure the protection of personal data and to guarantee that the data subjects and their underlying data provided on the platform are protected has developed specific techniques and adopted principles to achieve these objectives. All relevant actions are in full compliance with the GDPR.

The IMOVE development team has adopted the privacy by design principle prior to the actual development of any software component. Privacy by Design states that any action that involves processing personal data must be done with data protection and privacy in mind at every step. In practice, this means that it must be ensured that privacy is built into the system during its whole life cycle.

Any personal data are collected on a strictly need-to-know basis, solely for the purposes of the IMOVE project and will be destroyed when no longer needed for that purpose. Technical and operational measures are implemented to ensure that users will be able to access, rectify, cancel and oppose the processing and storage of their personal data.

Data anonymization techniques have been developed at the data collection stage as one way of rendering the project GDPR compliant and to protect the users' personal data and other sensitive information. Anonymization takes personal data and makes it anonymous, or not attributable to one specific source or person.

For limited specific purposes, where multiple data samples need to be linked to the same individual, pseudonymization by irreversible cryptographic pseudonyms is employed. The linkage information required for re-identification will be stored in a separate database utilizing server-side encryption and tightly controlled access.

Additionally, the data adapters that were deployed in the data collection phase are developed to transform if needed, personal data, so that they are compliant with the security and privacy standards that the project's ERB has set.

Data are acquired in conformity of security best practices through protected connections and dedicated state-of-the-art IT infrastructures (LANs, protected servers, firewall).

Encrypted communication based on SSL/TLS/SSH is used whenever personal information is transferred between systems. As a general rule, collected sensitive data are stored as close to the point of origin as possible. The data are released for further processing by other software components of the IMOVE system on a strict need-to-know basis.

Only required information will be released and with the minimum amount of detail and specificity required for a given processing task. Whenever aggregated data are sufficient, individual records even made anonymously, will not be used. Secure computation techniques are used to derive required processing results without having to collate all information in a single location. Collected personal data will not be transferred outside the EU.

Access to all data is restricted using appropriate authentication and authorization techniques (passwords, private keys etc.). Only minimum access privileges required to fulfil their roles are granted to individuals and systems. Access to personal information is also monitored and logged. All individuals working with personal data within the project are aware of their responsibilities and obligations with respect to data protection.

7 OPPORTUNITIES

Following on the previous two chapters, the possible opportunities associated to data sharing are both in terms of improving the service quality and the associated revenue/satisfaction. “Data” is the key to a successful implementation. Without data, there is no means for knowing what and if a service creates any value for customers. It is not possible to conclude whether a business model was successful or not. An executed business plan might very well create revenue and be profitable, but there are no real chance in knowing if the customers are satisfied and what could be done to increase the satisfaction and thereby the profitability even more. Business models are all about knowledge and knowledge is all about data. The points below outline some of the other benefits activated by the knowledge of data:

- **Better service and planning**: the data obtained would constitute an important input for better planning the infrastructure and the service operation. It would also improve the real-time information provided to the users and the reaction in case of a disruption.
- **Hidden demand**: better understanding the demand at an individual level, could also increase the chances to reach a potential demand that is currently not renouncing to the individual use of the private cars despite the associated inconvenient (ownership cost, time for parking, contamination, etc.).
- **Business for 3rd parties**: providing new applications or data analytics services.
- **Business model innovation:** developing new MaaS services and ancillary services (e.g. grocery deliveries) based on data that reveal user preferences, needs and travel patterns. Data shared within the MaaS ecosystem is a key enabler of these types of innovation.
- **Awareness**: where data are gathered it is, of course, the provider of the service or the property of the creator, namely the user. Data concerning mobility can become valuable in understanding one’s personal mobility and finding areas of improvement.
- **Automatization**: one of the requirements for automatization is data and gathered mobility data creates the foundation for automatization of personal mobility and autonomous vehicles.

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 723314.

8 POSSIBLE TRIGGERS

It is important to take into account the data available to each actor in the ecosystem and the data producer of each dataset. The different entities within the IMOVE ecosystem, will have often a very diversified access to different types of data, and it has to be clear what every entity (including the user) will have in exchange for their data in order to maximise the sharing of data. The user, for instance, will need to receive some extra functionalities in exchange for their habits' and behaviors' information. This will be requested to be compliant to the GDPR but also to motivate the user to join, in a context where more users start wondering if it is worth and necessary to share certain information. In a similar way, if we consider public and/or private transport operators, MaaS operators, specific service providers and others, all have a piece of information concerning the user and their service and the sharing is essential in order to both improve the knowledge about the user preferences, habits and needs, as well as to improve the global offer and integration of services. But for this, innovative business models and a clear definition of the governance mechanisms are crucial. It should be clear at any moment the concerns of each stakeholder in terms of competitiveness and interests to overcome barriers and create new opportunities, paving the way for roaming across Europe. Main identified triggers:

- **Innovative business models**: not necessarily involving an economic transaction, but rather focusing on identifying the interest of each stakeholder in the MaaS Ecosystem and combining them whenever possible. Every actor should have clear what it will have in exchange and how this will contribute to reaching its goals.
- **Clear rules** on who is sharing which data with whom, in which context, until when, for which purpose, etc. Smart contracts could facilitate this, as well as clear mechanisms on sharing personal information with different groups of people.
- **Data privacy mechanism**: particularly associated with data deletion and data recall when a user/entity decide to withdraw for the service/collaboration.
- **International agreements:** facilitating basic roaming
- **Policies and standards**: it seems likely they will have an important role but it has to be better understood in which terms.

GDPR also includes mandatory breach disclosure that will help consumers to understand serious incident concerning confidentiality and security, and it acts as a transparency mechanism as well as a mechanism to help those affected mitigate any harm.

It is also important to highlight the positive effect that the GDPR will have in the effective opening of borders for the EU travellers as personal data subjects. In an increasingly digitized world, this "European" traveller perspective would be impossible without previously removing all the already existing digital borders between EU countries, and this is exactly the aim of the GDPR, by harmonizing all data protection regulations throughout the EU.

Regarding the role of innovative business models as a potential trigger for data sharing within the MaaS ecosystem – here there is a need for a dynamic perspective that accounts for developments that will evolve over a longer period. MaaS developments within the IMOVE living labs (and elsewhere) are categorised

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 723314.

according to the three different approaches – a market driven approach where commercial actors from the private sector take on the role as MaaS provider; a publicly-driven approach whereby public actors, typical public transport operators, adopt this role; and a public-private approach whereby public-private partnerships facilitate the provision of MaaS services (see D1.4 for a full analysis). These approaches are currently at an embryonic stage and will likely change in any given city over time (the IMOVE living labs have stated this ambition) before consolidating as MaaS services are rolled out on the open market. There are multiple implications of these developments for data sharing. First, it will take considerable time to establish a stable set of rules and governance structures for data sharing with local MaaS ecosystems. Conversely, the consolidation of roles and responsibilities within each ecosystem is a key enabler of data sharing vis-à-vis the types of data required, how they will be used and by whom in the ecosystem. At present, as different actors within the ecosystem are jostling for position and as MaaS business models undergo further iterations, it is difficult to foresee how data sharing practices will play out. The consolidation of MaaS ecosystems and business models will likely aid these processes.

A second, related enabler of data sharing is linked to the notion that MaaS business models can serve as multi-sided platforms. The latter create value by facilitating interactions between multiple customers. Within the digital world, multi-sided platforms typically create linkages between different types of service providers and different types of customers. At present, the majority of MaaS initiatives are focused on delivering different types of mobility-related services involving passengers, and in some cases, movements of goods. The latter is perhaps critical for MaaS to be considered as a genuine alternative to private car ownership and use. Whilst the private car is linked to a range of sustainability problems, it is a wonderful service provider from a traveller's perspective in that cars can be used to transport people and their belongings for commuting, for grocery shopping, for leisure activities, for vacations, and so on. MaaS is thus linked to a range of ancillary services that can facilitate the transportation of goods for individuals within cities. Realised via collaborations with service providers for grocery deliveries, mobility hubs, e-commerce, waste-removal services, sport clubs and associations, hotels, tourist destinations etc. At present, it is difficult to foresee the different types of innovation opportunities associated with mobility as a multi-sided platform. As MaaS business models and ecosystems consolidate and become more structured, it is possible that data sharing via multi-sided platforms may be enabled by the realisation of numerous business and innovation opportunities among platform providers, who may seek to establish platforms similar to those provided by Apple and Google, who outsource certain developments to third-party developers based on the provision of open data and open source code.

A third potential enabler of data sharing within and beyond the MaaS ecosystem is transport system governance and real-time traffic management. As MaaS services proliferate within and around cities, municipal organisations, city administrations and public transport authorities and operators will likely place demands on MaaS providers to provide data that can assist in the management of local transport systems. Both momentaneous and cumulative data are required by these types of public organisations to manage short-term issues such as congestion, demand peaks and system resilience during disruptions and events; and for longer term issues linked to transport planning, licensing, urban development and so on. For this type of enabler to be relevant and effective, MaaS schemes must attract a large number of users within cities.

A final enabler of data sharing within and between MaaS ecosystems is roaming. At present, roaming is relevant among individual transport service providers, where the most widespread case is integrated public

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 723314.

transportation. This can occur in terms of integrating different operators within a given city, or between cities and regions to facilitate seamless intercity travel. In future, as more sophisticated MaaS schemes at higher levels in the MaaS topology [6] proliferate, it is likely that MaaS providers will integrate their services such that users can roam between different MaaS schemes in different locations, nationally and internationally. This type of roaming draws parallels to telecommunications roaming, which took several decades to fully evolve. MaaS roaming will likely be facilitated by commercial and technical agreements between MaaS providers, or via national platforms and access points.

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 723314.

9 LESSONS LEARNED

Regarding data privacy, security and protection of personal data, following constructive and transparent cooperation between the project partners, organisational measures included data and document access policies, restricting access rights appropriately and training on storing passwords securely. On technical measures, vulnerability tests are essential before providing services publicly. Data protection by design and by default application was essential on the software development process with encrypting mechanisms, secure protocols and digital certificates used. Regular checks on the ethics standards set by the Ethics Review Board, played a major role in ensuring that the project is running completely under the proper planning and the GDPR scope.

One of the main challenges was not related to the technology dimension, but the cooperation that requires a specific model to pave the way that existing transport service providers do not lose their customer relationship to the third-party operators. This is related to the competitiveness issues and the fear of possible cannibalism situations that the existing transport service providers try to avoid with the purpose of protecting their market positions. Another crucial aspect here is to guarantee the single transport providers that they will not lose the contact with their customers, both in term of their brand visibility and their data of customer habits, preferences and needs.

One of the barriers of this project in the case that the leading role (MaaS provider towards the users) was assumed by a private company (e.g. Berlin) was buying tickets via an API from public service providers. In some cases, the decision to provide a service operated by the public authority slows down the process of creating such APIs, while in other cases the cost, complexity and uncertainty linked to this have been the barrier. However, as in the case of Open Data, the issue is also linked to political decisions and has already started in some specific countries, particularly in the North of Europe.

Although the city authorities are not in a competition with big companies such as Google, they need to take a proactive role to set a framework for the private operators to avoid that on one side only big players can assume risk and costs of entering a market without clear conditions; on the other side, missing any opportunity for fixing such rules before the big players dominate the market (and resell to the city high value information based on their open data and services.) The cases of Hannover and London are positively inspiring in this sense, given both have been ensuring the setup some clear conditions for collaboration, independently from the decision of being a MaaS provider (e.g. Hannover case) or simply creating a framework for private entities interested in taking such a role.

Through the IMOVE project, it is concluded that the European cities are not completely mature for having large scale MaaS schemes, but neither they can effort to take a completely passive role. And, data are at the core of MaaS, being at the same time linked to several barriers and to possible triggers and opportunities.

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 723314.

REFERENCES

[1] Google. “Privacy Policy”. Privacy & Terms. Last modified 17/04/2017. https://www.google.com/policies/privacy/.

[2] Gorini, Marco, Freixanet, Josep, Fernandez, José. “D7.1 - Project Handbook”. *IMOVE project deliverable*, released 31/08/2017.

[3] Moore, J. F. (1996). *The death of competition: leadership and strategy in the age of business ecosystems* (p. 297). New York: HarperBusiness.

[4] https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en

[5] https://gdpr.eu/article-4-definitions/

[6] Sochor, J., et al., A topological approach to Mobility as a Service: A proposed tool for understanding requirements and effects, and for aiding the integration of societal goals. Research in Transportation Business and Management, 2018. 27: p. 3-14.

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 723314.