MOBIDATALAB

Labs for prototyping future mobility data sharing solutions in the cloud

D2.2 Recommendations on Data Sharing Legal Frameworks

01/02/2024

Author(s): Aliki BENMAYOR (KUL), Alberto BLANCO-JUSTICIA (URV)

MobiDataLab is funded by the EU under the H2020 Research and Innovation Programme (grant agreement No 101006879).

Summary sheet

Deliverable Number	D2.2
Deliverable Name	Recommendations on Data Sharing Legal Frameworks
Full Project Title	MobiDataLab, Labs for prototyping future Mobility Data sharing cloud solutions
Responsible Author(s)	Aliki BENMAYOR (KUL)
Contributing Partner(s)	Alberto BLANCO-JUSTICIA (URV), Benjamin PLOT (HOVE), Hiba MECHYAKHA (HOVE)
Peer Review	ICOOR, POLIS
Contractual Delivery Date	31-12-2023
Actual Delivery Date	22-12-2023
Status	Final
Dissemination level	Public
Version	V1.0
No. of Pages	60
WP/Task related to the deliverable	WP2/T2.1
WP/Task responsible	AKKODIS/KUL
Document ID	MobiDataLab-D2.2-RecommendationsOnDataSharingLegalFramworks.v1.0
Abstract	This deliverable analyses the most important legal gaps that need to be addressed for the smooth operation of transport/mobility data spaces as identified in D2.1. It also provides corresponding recommendations, where relevant.

Legal Disclaimer

MOBIDATALAB (Grant Agreement No 101006879) is a Research and Innovation Actions project funded by the EU Framework Programme for Research and Innovation Horizon 2020. This document contains information on MOBIDATALAB core activities, findings, and outcomes. The content of this publication is the sole responsibility of the MOBIDATALAB consortium and cannot be considered to reflect the views of the European Commission.

Project partners

Organisation	Country	Abbreviation
CONSORZIO INTERUNIVERSITARIO PER L'OTTIMIZZAZIONE E LA RICERCA OPERATIVA	Italy	ICOOR
HOVE	France	HOVE
KATHOLIEKE UNIVERSITEIT LEUVEN	Belgium	KUL
POLIS - PROMOTION OF OPERATIONAL LINKS WITH INTEGRATED SERVICES	Belgium	POLIS
UNIVERSITAT ROVIRA I VIRGILI	Spain	URV

Document history

Version	Date	Organisation	Main area of changes	Comments
0.1	13/11/2023	KUL	All	Draft version
0.2	17/11/2023	ICOOR, POLIS	All	Peer review
0.3	19/12/2023	KUL	All	Rework + TL Quality Check
0.4	20-22/12/2023	AKKODIS	All	Coordinator Quality Check
1.0	22/12/2023	AKKODIS	All	Submission

Executive Summary

Data are increasingly viewed as a commodity to be traded in their own rights. This has created a growing interest from policymakers in the creation of so-called “data markets” and “data spaces”, viewed as a mean to foster data sharing.

However, the legal framework is not quite well-aligned with this new pattern towards the commodification of data. On the one hand, it is generally agreed that there is no “ownership right” or general exclusive rights on data in the acquis of the European Union (EU). Thus, data cannot be legally ‘sold’ like a physical object. On the other hand, a variety of legal frameworks may apply to data and data transactions.

For example, data protection law is applicable when personal data are processed, competition law may apply as well if data is shared between or pooled by competitors (actual or potential) resulting in the distortion of competition in the relevant markets. In addition, mobility sectoral regulations lay down various types of obligations relating to data directly or indirectly, such as an obligation to provide access for third parties to re-use the data. The legal frameworks have indeed accumulated over time, with various rationales.

In the previous version of this report (D2.1, “Legal and Regulatory Data Sharing Gap Analysis”), we analyzed the current EU legal and regulatory frameworks for data sharing and re-use in the transport sector (covering all transport modes) to identify legal and regulatory gaps. In this report, we dive deeper in a selection of these gaps (those that can be effectively addressed) and provide relevant recommendations. As with D2.1, a separate chapter is dedicated to the use case of Mobility-as-a-Service. Through this analysis, several potential legal gaps are addressed for the smooth operation of transport/mobility data spaces.

Funded by the European Union

Table of contents

Funded by the European Union

List of figures

List of tables

Funded by the European Union

Abbreviations and acronyms

Abbreviation	Meaning
API	Application Programming Interfaces
CJEU	Court of Justice of the EU
DGA	Data Governance Act
DMA	Digital Markets Act
DSA	Digital Services Act
DSS	Data Sharing Services
EC	European Commission
ECD	E-Commerce Directive
ECHR	European Convention of Human Rights
ECJ	European Court of Justice
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
EEA	European Economic Area
EU	European Union
GDPR	General Data Protection Regulation
MaaS	Mobility-as-a-Service
MMTIS	Multimodal Travel Information Services [Delegated Regulation]
NAP	National Access Point
PTA	Public Transport Authority
PTO	Public Transport Operator

RTTI	Real-Time Traffic Information Services [Delegated Regulation]
TFEU	Treaty on the Functioning of the European Union

1. Introduction

1.1. Project overview

There has been an explosion of mobility services and data sharing in recent years. Building on this, the EU-funded MobiDataLab project works to foster the sharing of data amongst transport authorities, operators and other mobility stakeholders in Europe. MobiDataLab develops knowledge as well as a cloud solution aimed at easing the sharing of data. Specifically, the project is based on a continuous co-development of knowledge and technical solutions. It collects and analyses the advice and recommendations of experts and supporting cities, regions, clusters and associations. These actions are assisted by the incremental construction of a cross-thematic knowledge base and a cloud-based service platform, which will improve access and usage of data sharing resources.

1.2. Purpose of the deliverable

This deliverable focuses on a legal and regulatory data sharing gap analysis. The following chapters analyse a selection of these gaps, namely those concerning Data protection and the application of the General Data Protection Regulation ("GDPR"), Competition law as well as those related to MaaS. The main criterion for determining the scope of the analysis was whether the gap identified could be effectively addressed or not, based on the current state of play of the legislation and the available research.

1.3. Structure of the deliverable and its relationship with other work packages/deliverables

This deliverable is organised as follows. Section 1 analyses the gaps identified under the GDPR. Section 2 deals with Competition law. Section 3 analyses the gaps under MaaS. Conclusions and recommendations are provided at the end of each section. Final conclusions and a summary of all recommendations is provided in Section 4.

Funded by the European Union

2. Horizontal legal and regulatory gap analysis and recommendations

The first version of this report (D2.1 “Legal and Regulatory Data Sharing Gap Analysis”) contained an analysis of the current EU legal and regulatory frameworks for data sharing and re-use in the transport sector (covering all transport modes). It analysed the following horizontal and sector-specific legislation:

- Privacy and Data Protection (the General Data Protection Regulation, the e-Privacy Directive and forthcoming Regulation);
- Competition law;
- The Public Sector Information Directives (including the 2019 Open Data and Public Sector Information Directive);
- The Regulation on the free flow of non-personal data;
- Legislation concerning digital platforms and/or intermediaries (the e-Commerce Directive, the Platform-to-Business Regulation and the recent Digital Services Act Package);
- The proposal for a Data Governance Act.
- The Intelligent Transport Systems Directive (including its Delegated Regulations) (sectorial legislation)

A separate chapter was dedicated to the use case of MaaS. Through this analysis, several potential legal gaps were identified to be addressed for the smooth operation of transport/mobility data spaces. The gaps were identified in the following areas:

- The application of the GDPR;
- Competition law;
- The Open Data & Public Sector Information Directive;
- The proposal for a Data Governance Act;
- The Intelligent Transport Systems Directive & Delegated Regulations (namely with regard to their interface with the other horizontal legal frameworks).

An overview of the legal and regulatory gaps can be found below:

Table 1 - Overview of legal and regulatory gaps

Legal framework	Gaps			
GDPR	How to ensure GDPR compliance while benefiting from the information personal data can provide (e.g. when identified or pseudonymous	Identifying the legal basis under which data processing can take place if consent is withdrawn or rendered invalid	Potential difficulty in distinguishing between what constitutes personal and what non-personal data	Characterising the role of actors in the data stakeholder framework under the GDPR

Funded by the European Union

	data may be necessary to understand mobility patterns)			
Competition law	Defining 'data markets' to assess dominance (and establish potential abuses)			
Open Data and Public Sector Information Directive	Lack of guidance on the aspect of protection of personal data under PSI (e.g. in terms of limits to anonymisation, opportunity to carry out data protection impact assessments).	Divergence of implementation of the 2013 PSI Directive in EU Member States which may also lead to divergence of the 2019 Open Data Directive		
Regulation on the free flow of non-personal data	Difficulty to qualify data as 'non-personal' (which may result in an unnecessary application of the GDPR)			
Data Governance Act	Risk of overlap in the scope of application of the DGA proposal with the Open Data Directive which results in a lack of clarity on which obligation(s) is(are) concretely applicable to Public Sector Bodies			
ITS Directive and Delegated Regulations	Lack of clarity on the interface with the Open Data and Public Sector Directive, particularly following the 2019 revision.	No obligation under the Delegated Regulation 2017/1926 on multimodal travel information services to make dynamic travel and traffic data accessible though the National Access Point	Lack of clarity on the application of the GDPR	Different standards set by Delegated Regulation 2017/1926 on the provision of EU-wide multimodal travel information services and Delegated Regulation 2015/962 the provision of EU-wide real-time traffic information services in terms of (i) data protection principles and (ii)

Funded by the European Union

				whether anonymisation is required, or other privacy preserving mechanisms would suffice

The following chapters will focus on the analysis of a selection of these gaps, namely those concerning data protection and the application of the GDPR, Competition law as well as those related to MaaS. The main criterion for determining the scope of the analysis was whether the gap identified can be effectively addressed or not through recommendations, taking into account the current state of play of the legislation and of the available research.

2.1. GDPR

2.1.1. The necessity of using personal mobility data

2.1.1.1. The notion of ‘personal data’

Personal data is broadly defined as any type of information that relates to an identified or identifiable natural person (“data subject”). The Court of Justice of the EU (“CJEU”) has clarified that personal data is not limited to sensitive or private information, but potentially encompasses all types of information, both subjective and objective provided that it relates to the data subject. The possible extent of “personal data” was clarified by the CJEU in the Breyer case, which concerned IP addresses. The CJEU clarified that a piece of information can be considered personal data whenever additional information can be sought from third parties to identify a data subject. The European Convention of Human Rights (“ECHR”) has also interpreted the term “personal data” as not being limited to matters of the private sphere of an individual.

The Working Party 29 (WP29), predecessor of the current European Data Protection Board (EDPB) has provided guidance as to how the different elements of this definition should be interpreted in its 2007 opinion on the concept of personal data.[1] Its opinion focuses on four elements of the definition (“any information”, “relating to”, “identified or identifiable” and “natural person”) and provides with guidance as to what these different elements entail. While this opinion of often still referred too, it should be noted that it has as such not been explicitly endorsed by the EDPB.[2]

The information about a person can be clear, *i.e.*, directly identifying an individual (*e.g.*, name, surname), or it can indirectly allow for the individual to be identified (*e.g.*, by combining information on the specific hour a ticket is validated and footage from surveillance cameras).

[1] Article 29 Working Party, Opinion 04/2007 on the concept of personal data, WP 136, 30 June 2007.
[2] https://edpb.europa.eu/our-work-tools/general-guidance/endorsed-wp29-guidelines_en.

Funded by the European Union

Personal data is further categorised as *volunteered*, *observed*, and *inferred* data. *Volunteered* (provided) data originate from direct actions of the data subject, in full awareness of the consequences that result with the disclosure of his/her personal data. Examples of volunteered data include data disclosed in the context of a loan application, credit card use or shared (actively) via online social networks.

Observed data such as IP addresses, meta-data, device ID, browser information or interaction data is either captured for a purpose as a result of a deliberate measurement or it is simply the *exhaust* data which comes into being as a by-product of ICT systems (e.g., firewalls, load balancers, routers, switches, etc.) deployed for other purposes.

Inferred (a.k.a. *derived*) data is the output of data processing as an aggregate. It is the data/information resulting from a subsequent analysis of the raw data either provided (volunteered) by the data subject or actively observed by the data controller. Inferred data could include user profiles, spending habits, peak hours of a commercial establishment or an assessment of one's physical condition based on the data collected by a smartphone application.[3]

2.1.1.2. The distinction between personal and non-personal data

If a data controller or another person wants to assess whether the processing of a given dataset is subject to the GDPR, an assessment is required of the dataset in light of the criteria put forth by recital 26 of the GDPR. If the data relates to an identified or identifiable natural person, the data will be considered personal data and the data controller will need to comply with the provisions of the GDPR. In order to determine whether or not an individual is identifiable, recital 26 of the GDPR states that "*account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly."*

In order to assess what is *'reasonably likely'*, one has to consider "*objective factors, such as the cost of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments"*. Moreover, the possibility of identification has to be assessed taking into account technological developments during the period for which the data will be processed, keeping in mind that identification that may not be possible today given the current state of technology, may be possible in the future.[4] As 29WP has noted, a mere hypothetical possibility to single out the individual is not enough to consider a person 'identifiable'.[5]

[3] For an argument for the inclusion of derived data in the scope of data portability, see Bertin Martens and others, 'Business to Business Data Sharing: An Economic and Legal Analysis' (Digital Economy Working Paper 2020-05, European Commission, Seville, 2020 JRC121336) 4 https://ec.europa.eu/jrc/sites/default/files/jrc121336.pdf, accessed 14 January 2022.

[4] *Ibid*

[5] Article 29 Working Party Opinion on the concept of personal data (n 1), p.15.

Funded by the European Union

Furthermore, according to the CJEU interpreting the concept of “means reasonably like to be used” in the context of dynamic IP addresses (as mentioned above) it is not necessary that all the information for the identification of the data subject is in the hands of one person, i.e. the data controller.

However, in the case where additional data is required to identify the individual, what matters is the means reasonably likely to be used to access and combine such additional data.[6] In its recent judgement in in the case *SRB v. EDPS (Case T-557/20),* the General Court confirmed this approach.[7]

In essence, recital 26 of the GDPR prescribes a risk assessment to differentiate between personal and non-personal or anonymous data. This entails that, where there is merely a negligent or hypothetical possibility to single out an individual, this individual should not be considered as ‘*identifiable’.*[8] Recital 26 thus instates a ‘*risk-based approach’* to qualify information as either personal or non-personal.[9] The ‘*reasonably likely’* part of recital 26 thereby acts as a “*brake that keeps the concept of ‘personal data’ from spinning out of control”*.[10] Recital 26 makes the GDPR concept of ‘personal data’ suitable for an assessment of the (non-)personal nature of data that is tailored and context-specific.[11] However, it should be noted that there is disagreement on the level of risk implied by recital 26. Some argue that it dictates a “zero-risk” approach, i.e. any risk of identification should be close to zero by considering all the possible means for re-identification, while others adopt a more “acceptable-risk” approach where re-identification is minimised by considering all probable means for re-identification.[12]

In addition to these concepts, the GDPR has introduced the notion and definition of ‘pseudonymisation’. More specifically, pseudonymisation refers to the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.[13]

[6] CJEU 19 October 2016 C582/14 Patrick Breyer v Bundesrepublik Deutschland ECLI:EU:C:2016:779 (‘Breyer case’) para 43-45.

[7] https://www.law.kuleuven.be/citip/blog/anonymous-vs-pseudonymous-data-the-cjeu-reaffirms-the-relative-approach-to-the-concept-of-personal-data/.

[8] Article 29 Working Party Opinion on the concept of personal data (n 1), 15; FINCK, M. and PALLAS, F., "They who must not be identified – distinguishing personal from non-personal data under the GDPR", International Data Privacy Law 2020, vol. 10, 14; PURTOVA, N., "The law of everything. Broad concepts of personal data and future of EU data protection law", Law, Innovation and Technology 2018, 46; STORMS, S., "Identify me if you can – Identifiability and anonymisation, 16 January 2018, https://www.law.kuleuven.be/citip/blog/identify-me-if-you-can-identifiability-and-anonymisation/.

[9] FINCK, M. and PALLAS, F., "They who must not be identified – distinguishing personal from non-personal data under the GDPR", *International Data Privacy Law* 2020, vol. 10, 14.

[10] STORMS, S., "Identify me if you can – Identifiability and anonymisation, 16 January 2018, https://www.law.kuleuven.be/citip/blog/identify-me-if-you-can-identifiability-and-anonymisation/.

[11] PURTOVA, N., "The law of everything. Broad concepts of personal data and future of EU data protection law*", Law, Innovation and Technology* 2018, 44, in reference to SCHWARTZ, P. and SOLOVE, D., "The PII Problem: Privacy and a New Concept of Personally Identifiable Information", *New York University Law Review* 2011, 1814-1894.

[12] ‘On the legal nature of synthetic data’, César Augusto Fontanillo López, Abdullah Elbi, p.7 and the sources mentioned therein. https://openreview.net/pdf?id=M0KMbGL2yr.

[13] GDPR, Art. 4 (5).

Funded by the European Union

Pseudonymisation is commonly perceived as a data security measure that reduces linkability by replacing any identifying characteristic or attribute by another identifier, a pseudonym.[14] The reason why this new concept is important is because according to the GDPR, pseudonymised data are personal data.[15]

2.1.1.3. Mobility data

Mobility data in its simpler form are data about individuals that include their locations at specific times. Sources of real-time raw individual location data include but are not limited to, cell towers, Wi-Fi access points, radio-frequency identification (RFID) tag readers, location-based services, or credit card payments. Historical location data, in the form of datasets in which each of the records corresponds to an individual and includes their location data for some periods are referred to as trajectory microdata sets. Such trajectory microdata sets are often of interest to transport authorities, operators, and other stakeholders to evaluate and improve their services, the state of the traffic, etc. and thus are often publicly released or shared. Sharing of mobility data is occasionally shared as aggregates (e.g., heat maps) instead of at an individual level.[16]

Whichever the form of these mobility data, they all share some statistical characteristics that make their sharing a potential privacy risk. Mobility data are highly unique and regular. Unicity refers to the data of different individuals to be easily differentiable, particularly at some specific locations. The starting and ending locations of users' trajectories are often their home and work locations which, again are highly unique and can lead to reidentification. Studies show that user full trajectories can be uniquely recovered with the knowledge of only two locations. The regularity of trajectories means that for single individuals, their data follows periodic patterns. Namely, individuals tend to follow the same trajectories during workdays—home to work and back to home.[17]

The easiness with which one activity can include personal data can also be demonstrated in smart mobility systems where there is the risk that users could be identified when mobility information is matched with data from other sources.[18]

For example, the provision of additional non-transport functionalities and services through smart mobility cards may reveal users' social relations and activities. This is the case with travel cards that also provide access to parking areas or grant discounts at public services (museums, concert halls, etc.) or commercial services. Mobility information may reveal that two different users travelled from their home/office to reach the same place, attended the same play in the same theatre, then made the same journey, and had dinner in the same restaurant, which adopts discount rates for travel cardholders. Social relationships and related interactions can be better monitored when mobility data are coupled with publicly available information (e.g. Twitter postings, blog entries).[19]

[14] Article 29 Working Party, Opinion 05/2014 on Anonymization Techniques, WP 216, p. 20.
[15] GDPR, Rec. 26.
[16] See MobiDataLab, D2.3 "State of the art on Mobility and Transport data protection technologies".
[17] *Ibid*
[18] Alessandro Mantelero, 'Data Protection, E-Ticketing, and Intelligent Systems for Public Transport' (2015), International Data Privacy Law, pages 309-320, Available at SSRN: https://ssrn.com/abstract=2659732.
[19] *Ibid*

Similar concerns can arise when tourist information is integrated with mobility information. For example, let's imagine a scenario where local authorities that organise tourism in their area (e.g. tourist offices) provide an application that integrates all available mobility services, if possible, in real time, with their data - especially to improve the tourist information they provide (e.g. with car parks, public transport services, tourist buses, and even data from bike sharing companies).[20]

This app may collect information about a user's location, the means of transport they have used, the exact sightseeing route he/she has taken, the museums visited and even perhaps the places where they had lunch. Even if some of these data are pseudonymised, their collective reading may lead to the user being identifiable.

Personal data in connected vehicles

In its Guidelines on the processing of personal data in the context of connected vehicles and mobility-related applications, the European Data Protection Board ("EDPB") notes that connected vehicles are generating increasing amounts of data, most of which can be considered personal data since they will relate to drivers or passengers.[21]

Even if the data collected by a car are not directly linked to a name, but to technical aspects and features of the vehicle, it will concern the driver or the passengers of the car. For example, data relating to the driving style or the distance covered, data relating to the wear and tear on vehicle parts or data collected by cameras may concern behaviour as well as information about other people who could be inside or outside the vehicle. Personal data in connected vehicles can also include metadata, such as vehicle maintenance status.

Geolocation data & connected vehicles

In the same Guidelines mentioned in example 1, the EDPB makes a specific reference to geolocation data as a category that warrants special attention.[22] The EDPB notes that geolocation data are particularly revealing of one's life habits. The journeys carried out are very characteristic and can reveal private details about a person's life (e.g. residence, places of leisure, places of worship, etc.). Vehicle and equipment manufacturers need therefore to be particularly vigilant not to collect location data except if doing so is absolutely necessary for processing. The EDPB further mentions several principles that need to be complied with when collecting geolocation data.

In D2.3, an overview of location data was provided and whether they constitute personal and non-personal data. It is evident from the table below that the distinction between the two is not always evident.

[20] See MobiDataLab D2.9, Use case for research, 3.3, Transport data sharing within the Linked Open Data vision.
[21] EDPB, 'Guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications', 28 January 2020, paras 3, 27-28 https://edpb.europa.eu/sites/default/files/consultation/edpb_guidelines_202001_connectedvehicles.pdf
[22] *Ibid*, paras 60-61.

Funded by the European Union

Arguably, anonymous mobility data can be used to understand mobility patterns. But it has been raised that without identified or pseudonymous data, it is not possible to get an accurate representation of the most valuable information for mobility planning.[23]

Table 2 - Categorisation of location data as personal data

Type of data	Description and sourcing	Personal
Driver location data	Driver location data in real-time for route planning, ETA, emission reporting, and other location-based services. These data are directly sourced from the users, via their mobile phones or GPS-enabled vehicles.	YES
GPS traces	GPS traces, also known as trajectory microdata in this document, consist of databases in which each record corresponds to an individual and contains the location data from this individual in each period of time. These traces are obtained by collecting and merging user-provided data.	YES
Static map data or transport network topology	Static map data or transport network topology, corresponding to maps and roads. These data are offered by public entities such as municipalities and departments of transport. Organisations such as Google and OpenStreetMap provide access to this kind of data	NO
Real-time traffic data	Real-time traffic data, including disruptions alerts and planned events. This information is also typically provided by public authorities in real-time, although it is common for service providers to offer real-time or aggregated traffic information obtained through crowdsourcing. Service providers that continuously monitor the location of their users (*e.g.,* Google via Android phones), can obtain traffic information aggregating the speed, location, and density of their users, and enrich this data with publicly available information.	YES/NO
Points of interest	Points of interest can be obtained from public authorities, such as departments of tourism, by companies announcing themselves or by crowdsourcing. PoIs can also be obtained by analysing highly visited locations in trajectory datasets.	YES/NO
Vehicle data	Vehicle data, fuel type and load are useful information for emissions reporting. These data must be supplied by users or companies.	YES/NO
Public transport data	Public transport data is offered by public transport operators, although they can also be obtained through crowdsourcing from public transport users, in real-time or not.	YES/NO

[23] *Mantelero* (n 18).

Funded by the European Union

2.1.1.4. Data anonymisation as a solution

The term "anonymous" is not defined in article 4 of the GDPR, but it is referred to in recital 26 to the GDPR. This recital states that the principles of data protection should “not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes”.[24]

Anonymisation mechanisms modify the original data to prevent disclosure of personal information, both of identities and their related confidential attributes. This modification may hurt the utility of the data. Anonymisation methods aim to protect the privacy of the respondents while producing as little as possible effects on the utility of the data. As utility, we may refer to the differences between the original and the modified data or to the differences in the results obtained from some processing on the original and the modified data (e.g., the accuracy of a machine learning model trained on the original data versus the accuracy of the model trained on the anonymised data).[25]

At the same time, the threshold for anonymisation has been set high. According to the Article 29 Working Party[26], “an effective anonymisation solution should prevent all parties from singling out an individual in a dataset, from linking two records within a dataset (or between two separate datasets) and from inferring any information in such dataset. Therefore, removing directly identifying elements is not enough to ensure that identification of the data subject is no longer possible.”[27] This means that if a controller de-identifies a dataset but at the same time keeps the original (identifiable) dataset – the de-identified dataset is unlikely to be considered truly anonymous but would still qualify as personal data.[28]

The Article 29 Working Party assessed the strengths and weaknesses of several anonymisation (and other) techniques[29] by focusing on the following three risk factors:

- singling out, that is “the possibility to isolate some or all records which identify an individual in the dataset”;
- linkability, that is the “ability to link at least two records concerning the same data subject or a group of data subjects (either in the same database or in two different databases). […]”;
- inference, which is “the possibility to deduce with significant probability the value of an attribute from the values of a set of other attributes”.

According to the Article 29 Working Party, an anonymisation technique that provides a solution against these three risks would be sufficiently robust against re-identification by the data controller or a third party considering the means they are reasonably likely to employ.[30]

[24] GDPR, recital 26.
[25] See MobiDataLab, D2.3 (n 16), p. 34.
[26] This is the supervisory authority that was replaced by the Data Protection Supervisory Board and essentially ensures the same function (uniform interpretation of the GDPR).
[27] Article 29 Data Protection Working Party Opinion on Anonymization (n 14), p.9.
[28] *Ibid*.
[29] *Ibid*, section 3, p. 11.
[30] *Ibid*., p. 23-24.

Funded by the European Union

It does however point out that careful engineering may allow combining several techniques to enhance the robustness of the anonymisation outcome. The optimal solution should ultimately be decided on a case-by-case basis.[31] National data protection authorities have also issued guidance on anonymisation.[32]

2.1.2. Consent as the legal basis for data transactions

According to the GDPR, personal data can be lawfully processed only based on the following[33]:

i) the consent of the subject;
ii) contractual necessity;
iii) legitimate interests of the data controller or a third party;
iv) compliance of the data controller with a legal obligation;
v) protecting the vital interests of a data subject or another person;
vi) necessity arising out of the performance of a task carried out in the public interest or the exercise of official authority vested in the controller.

Data-driven activities heavily rely on consent to collect and process personal data in a lawful manner.[34] Consent must be "given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data".[35] Usually, this is gathered by a user accepting the service's terms and conditions of the privacy policy.

Consent is not characterised by the anonymizationperson to whom it is provided (the controller) but concerns each act of personal data processing. It must be a) freely given, b) specific, c) informed and d) unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her. Generally, consent can only be an appropriate lawful basis if a data subject is offered control and is offered *a genuine choice* about accepting or declining the terms offered or declining them without detriment.

When asking for consent, a controller must assess whether it will meet all the requirements to obtain valid consent. If obtained in full compliance with the GDPR, consent is a tool that gives data subjects control over whether personal data concerning them will be processed. If not, the data subject's control becomes illusory and consent will be an invalid basis for processing, rendering the processing activity unlawful.[36]

[31] *Ibid*.
[32] See for example the Guidance of the ICO, the Irish Data Protection Authority and CNIL: https://ico.org.uk/media/1061/anonymisation-code.pdf, https://www.dataprotection.ie/sites/default/files/uploads/2022-04/Anonymisation%20and%20Pseudonymisation%20-%20latest%20April%202022.pdf, https://www.cnil.fr/fr/lanonymisation-de-donnees-personnelles.
[33] GDPR, Article 6.
[34] Eleni Kosta, *Consent in European Data Protection Law* (Martinus NIJHOFF Publishers 2013).
[35] GDPR, Article 4(11).
[36] EDPB, 'Guidelines 05/2020 on consent under Regulation 2016/679', 4 May 2020, https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf.

Funded by the European Union

Furthermore, as the requirement of informed consent is to ensure that data subjects will not be deceived or coerced and thereby wronged, it could facilitate data transactions by contributing to the building of trust between businesses and data subjects.[37]

Before obtaining valid consent, the controller(s) need to define the specific, explicit and legitimate purpose for the intended processing activity.[38] This acts as a safeguard against possible widening or blurring of purposes for which data is processed after a data subject has agreed to the initial collection of the data (so-called "function creep").[39] It is also important that consent is specific to the purpose of processing. A controller cannot seek one consent to cover different operations if these operations do not serve the same purpose (see also the principle of purpose limitation under 3.1.3 below in that regard). If a controller wants to use the personal data he has collected and is processing for another purpose, a compatibility assessment needs to be carried out.[40]

Providing information to data subjects before obtaining their consent is essential to enable them to make informed decisions, understand what they are agreeing to and exercise their right to withdraw their consent. If the controller does not provide accessible information, user control becomes illusory and consent will be an invalid basis for processing.[41] A controller needs to provide to data subjects at least the following information: i) the controller's identity, ii) the purpose of each of the processing operations for which consent is sought, iii) what (type of) data will be collected and used and iv) the existence of the right to withdraw consent.[42]

According to Article 7(3) of the GDPR, the data subject has the right to withdraw his or her consent at any time. The withdrawal must be as easy as providing consent. For example, if consent was provided by clicking "I agree" on the privacy policy of a service, it should be as easy to withdraw consent.

As mentioned above, in case of joint controllership, the data subject can exercise his/her rights against each of the controllers. This right makes consent the "weak spot" of data transactions as they can be prone to invalidation. Withdrawal of consent results in an ex-post invalidation of data processing while keeping processing made prior to withdrawal valid. If there is no other lawful basis justifying the processing of the data, they should also be deleted by the controller.[43] The possibility for the data subject to exercise his/her right to withdrawal creates significant uncertainty for data sharing as it may result in the collapse of a chain of data sharing transactions. Similar considerations apply if it is found that consent did not fulfil the conditions analysed above.

While data controllers typically try to rely on consent as a means to stipulate the lawfulness of their data processing operations, other grounds can also be called upon. Data processing can be justified without the data subject's consent where alternative criteria specified in Article 6(1) GDPR apply.

[37] Laurens Naudts, 'The Right Not to Be Subject to Automated Decision-Making: The Role of Explicit Consent.' (*CITIP Blog*, 2 August 2016) <https://www.law.kuleuven.be/citip/blog/the-right-not-to-be-subject-to-automated-decision-making-the-role-of-explicit-consent/> accessed 14 January 2022; Aurelia Tamò-Larrieux , 'Privacy and Data Protection Regulation in Europe' in Aurelia Tamò-Larrieux, *Designing for Privacy and its Legal Framework*, vol 40 (Springer International Publishing 2018).
[38] GDPR, Article 5(1)(b) – "purpose limitation".
[39] EDPB Guidelines on consent (n 36), para. 56.
[40] GDPR, Article 6(4).
[41] EDPB Guidelines on consent (n 36), para. 62.
[42] *Ibid*, paras 64-65.
[43] GDPR, Article 17(1)(b) and (3).

Funded by the European Union

Contractual necessity as laid out in the GDPR could, in limited circumstances, provide the ground for a data transaction. Nevertheless, this is limited to specific contracts and does not constitute a ground for more general processing of personal data. Where certain processing is an indispensable part of a performance or formation of a contract, the data controller may process personal data within this capacity. Under the provision, the performance of a contract may not be made dependent upon the consent to process further personal data which is not needed for the performance of that contract.[44] Other grounds— legitimate interest of the controller or third parties, protection of "vital interests of the data subject" or the necessity of performance of a task in the public interest—may occasionally be the legal basis for processing personal data without the data subject's consent. Cumulatively, the *legal* grounds for processing reflect the understanding that lawfulness of data processing requires a balancing of interests among the data subject, data controller and public at large.[45]

2.1.3. The ecosystem of mobility actors and attribution of roles and responsibilities

2.1.3.1. Controller

Under the GDPR, the controller is "any natural or legal person, public authority agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data. Where the purposes and means of processing are determined by national or Community laws or regulations, the controller or the specific criteria for his nomination may be designed by national or Community law".[46] The controller is, essentially, the entity having control over the personal data processed, is responsible for the data processing and for ensuring that such processing – including any processing carried out by a third party (i.e. processor) - complies with the GDPR.

The assessment of controllership should be made based on the facts of a particular case. The key criterion to assess who is a controller is to designate the person who determines the "purposes" (the "why") and the "means" (the "how") of the processing of personal data.[47] It seems however the purpose may take precedence over the means. As such, determining the purpose of the processing, in any case, leads to a qualification as a controller. Determining the means would lead to control only when it concerns the essential means, such as which data is processed, the duration of the processing, which third parties have access to the data.[48] The determination of the technical and organisational elements of the means (e.g., which hardware or software to use) does not necessarily imply control and can hence be done exclusively by the processor.[49]

[44] W29 Party, «Opinion 06/2014 on the Notion of Legitimate Interests of the Data Controller under Article 7 of Directive 95/46/EC (WP 217),» 2014.
[45] A. Tamo-Larrieux, «Privacy and Data Protection Regulation in Europe,» de Designing for Privacy and its Legal Framework, Springer International Publishing, 2018.
[46] GDPR, Article 4 (7).
[47] EDPB, 'Guidelines 7/2020 on the concepts of controller and processor in the GDPR', 2 September 2020, section 2.1.4, p. 13 https://edpb.europa.eu/sites/default/files/consultation/edpb_guidelines_202007_controllerprocessor_en.pdf.
[48] *Ibid*. p.14.
[49] *Ibid*

Funded by the European Union

Under the GDPR, both natural, legal persons and a public authority can be considered as a controller. But normally it would be the company or a body that would qualify as a controller, rather than a specific individual within the company or the body.[50]

Joint controllership

When two or more parties jointly determine the purpose and means of processing, they are considered joint controllers.[51] "Jointly" must be interpreted as meaning "together with" or "not alone", in different forms and combinations.[52] Joint controllership exists when the parties decide together to process data for the same or common purpose. Joint controllership also requires that two or more entities have exerted influence over the means of the processing.

Joint participation

Joint participation through a *common decision* means deciding together and involves a common intention following the most common understanding of the term "jointly" referred to in Article 26 of the GDPR.[53] However, joint controllers could also adopt converging decisions. But they would need to complement each other and be necessary for the processing to take place in such a manner that they have a tangible impact on the determination of the purposes and means of the processing, where, for example, processing by each party is inextricably linked.[54]

The CJEU has broadened the scope of joint controllership in several cases. In *Jehovah's Witnesses*, the CJEU even considered that the entire community was considered a controller jointly with its members as the community participated in the determination of the purposes and means by organizing and coordinating the activities of its members, which helped to achieve the objective of the entire community.[55]

The Court also confirmed that the fact that one of the parties does not have access to the personal data processed is not enough to exclude joint controllership. For example, in *Jehovah's Witnesses*, the CJEU considered that it was not necessary that the community had access to the data in question, or to establish that that community had given its members written guidelines or instructions about the data processing.[56] The community participated in the determination of purposes and means and knew on a general level of the fact that such processing was carried out to spread its faith.[57]

[50] *Ibid*, section 2.1.1, p.9.
[51] GDPR, Articles 4(7) and 26.
[52] EDPB Guidelines on the concepts of controller and processor (n 47), p. 17.
[53] *Ibid*, para.52.
[54] *Ibid*, para.53.
[55] Case C-25/17, *Jehovan todistajat* [2018], ECLI:EU:C:2018:551, para.71.
[56] *Ibid*, para.75.
[57] *Ibid*, para.71.

Funded by the European Union

Jointly determined purpose

In *Fashion ID*[58] and *Wirtschaftsakademie*[59], the CJEU suggested that even when the entities do not have the same purpose for the processing, they may still be considered joint controllers, if their purposes are closely linked or complementary, for example, when there is a mutual benefit, provided that each of the entities involved participates in the determination of the purposes and means of the relevant processing operation. Yet, the mere existence of a mutual benefit (e.g., commercial) arising from a processing activity does not give rise to joint controllership. If the entity involved in the processing does not pursue any purpose(s) of its own in relation to the processing activity but is merely being paid for the services rendered, it is acting as a processor rather than as a joint controller.[60]

Jointly determined means

For joint controllership to exist, each entity involved does not need to determine all the means in each case. Indeed, as clarified by the CJEU, different entities may be involved at different stages of that processing and to different degrees.[61] It may also be the case that one of the entities involved provides the means of processing and makes it available for personal data processing activities by other entities. This scenario can notably arise in the case of platforms, stanstandardizedls, or other infrastructure allowing the parties to process the same personal data and which have been set up in a certain way by one of the parties to be used by others that can also decide how to set it up.[62]

Also, the use of a common data processing system or infrastructure will not in all cases lead to qualify the parties involved as joint controllers, in particular where the processing they carry out is separable and could be performed by one party without intervention from the other or where the provider is a processor in the absence of any purpose of its own (the existence of a mere commercial benefit for the parties involved is not sufficient to qualify as a purpose of processing).[63]

Joint controllers must determine and agree on their respective responsibilities on how to comply with the obligations under the GDPR, namely concerning the exercise of data subjects' rights and the duties to provide information (e.g. on the identity and contact details of the controller, the purposes and the legal basis for processing, any data recipients, etc.[64]).[65] In addition to this, the distribution of responsibilities should cover other controller obligations such as regarding the general data protection principles, legal basis, security measures, data breach notification obligation, data protection impact assessments, the use of processors, third country transfers and contacts with data subjects and supervisory authorities.[66] In short, they need to decide "who does what". But they have flexibility in distributing and allocating obligations amongst them as long as they ensure full compliance with the GDPR with respect of the given processing.[67]

[58] Case C-40/17, *Fashion ID* [2019], ECLI:EU:C:2019:629.
[59] Case C-210/16, *Wirtschaftsakademie Schleswig-Holstein* [2018], ECLI:EU:C:2018:388.
[60] EDPB Guidelines on the concepts of controller and processor (n 47), para. 60.
[61] *Ibid*, p. 19.
[62] *Ibid*, paras 62-63.
[63] EDPB Guidelines on the concepts of controller and processor (n 47), para. 66.
[64] The information obligations are set out in Articles 13 and 14 of the GDPR.
[65]GDPR, Article 26(1).
[66] EDPB Guidelines on the concepts of controller and processor (n 47), p.4.
[67] *Ibid*, para. 165.

Funded by the European Union

There is no obligation for the joint controllers to have a written contract, but the EDPB recommends drafting a legally binding document to ensure legal certainty. In any event, the main points ("essence") of the arrangement made on each controller's role and responsibilities needs to be made available to data subjects so that they know which of the controllers is responsible for what.[68] For efficiency purposes, joint controllers can designate in the arrangement a contact point for handling data subjects' requests.[69] But data subjects are not bound by this and remain free to contact either of the joint controllers to exercise their rights under the GDPR.[70]

Finally, each joint controller must ensure that they have a legal basis for the processing and that the data are not further processed in a manner that is incompatible with the purposes for which they were originally collected.[71]

2.1.3.2. Processor

The processor is a natural or legal person, public authority, agency, or other body which processes personal data on behalf of the controller.[72] In principle, there is no limitation as to which type of actor might assume the role of a processor. It might be an organisation, but it might also be an individual. Two characteristics define who can qualify as a processor: i) being a separate entity in relation to the controller, and ii) processing personal data on the controller's behalf.[73]

However, it has been noted that not every service provider that processes personal data while delivering a service is a "processor" within the meaning of the GDPR. The role of a processor is granted not from the mere processing of data but its concrete activities in a specific context. It is the nature of the service that will determine whether the processing activity amounts to the processing of personal data on behalf of the controller within the meaning of the GDPR.[74]

2.1.3.3. The mobility ecosystem

Given the multitude of actors active in a data-sharing ecosystem in the transport sector (see Figure 2 below), the correct characterisation of each actor's role under the GDPR can be quite challenging. This becomes crucial as otherwise, it is not possible to define the obligations for each actor and comply with the GDPR provisions to ensure lawful data sharing.

[68] GDPR, Article 26 (2).
[69] GDPR, Article 26 (2); EDPB Guidelines on the concepts of controller and processor (n 47), paras 180-183.
[70] EDPB Guidelines on the concepts of controller and processor (n 47), paras 184-187.
[71] *Ibid*, p.4.
[72] GDPR, Article 4(8).
[73] EDPB Guidelines on the concepts of controller and processor (n 47), section 4, p. 24.
[74] *Ibid*

Funded by the European Union

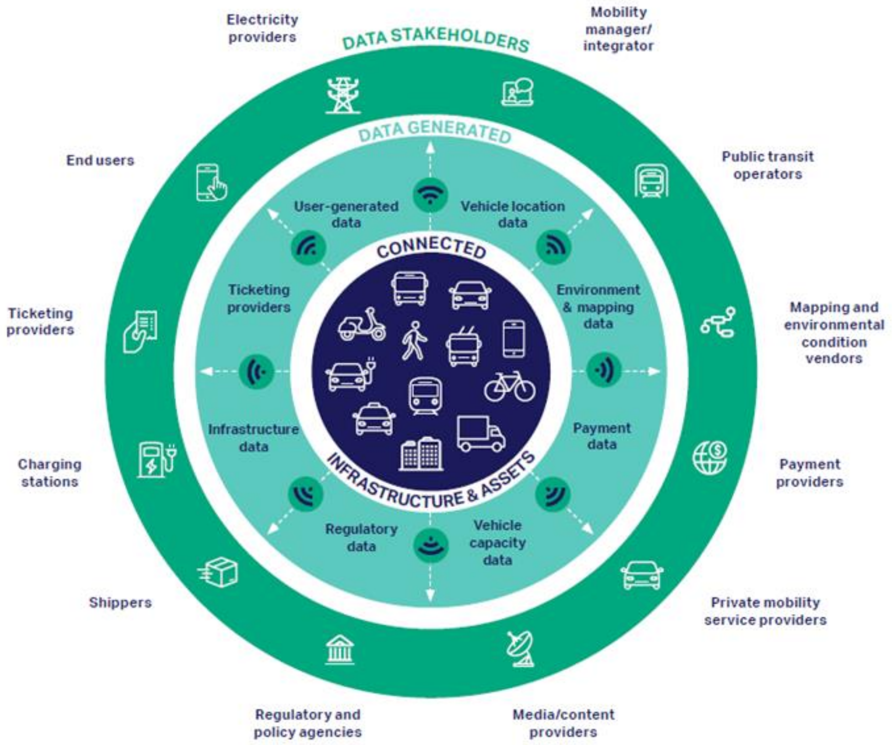


Figure 1 - Data stakeholder framework (WBCSD, 2020)

It is important to remember that different forms of collaboration are possible with diverging legal relationships which may be difficult to recognise and each bringing differing requirements[75]:

- A controller may “share” data with its service provider, who is bound to the controller by a written agreement. If the service provider acts only on behalf of the controller with the latter entrusting certain contractually defined processing activities to it, this would qualify as a controller - data processor relationship, provided that the processor is selected carefully and that an appropriate written agreement has been implemented (as required under Article 28 of the GDPR);
- A controller may share data with another controller, where that second data controller will use the data for entirely separate purposes and using separate means than the first one. This constitutes a controller-to-controller relationship. This type of interaction presents a series of unique challenges, notably in ensuring that there is a clear legal basis for the transfer and the further processing and that the further processing is compatible with the initial purposes of processing;
- A more complex case is that of joint controllership where multiple legal entities are jointly responsible for a common (shared) data processing activity as it requires the joint controllers to implement appropriate arrangements - habitually but not necessarily taking the form of contracts –to ensure that the GDPR is complied with.

[75] European Commission Support Centre for Data Sharing, ‘Analytical report on EU law applicable to sharing of non-personal data’, 24 January 2020, p.14, https://eudatasharing.eu/sites/default/files/2020-02/EN_AR%20on%20EU%20law%20applicable%20to%20sharing%20of%20non-personal%20data.pdf

Funded by the European Union

At the same time, the fact that several actors are involved in the same data processing operations does not mean that they are necessarily acting as joint controllers of such processing. Not all kinds of partnerships, cooperation or collaboration imply qualification as joint controllership as such qualification requires a case-by-case analysis of each processing at stake and the precise role of each entity with respect to each processing.[76]

2.1.4. Conclusions and recommendations

The above analysis demonstrates that there is an element of uncertainty concerning the application of some GDPR aspects, starting from the very concept of personal data. Some authors state that the attempts by EU data protection regulators to clarify the dichotomy between 'anonymous information' and 'personal data' have partly failed because of its "*implicit adoption of a static approach, which tends to assume that once the data is anonymized, not only can the initial data controller forget about it, but also the recipients of the dataset are free from any obligation or duty because the transformed dataset always lies outside the scope of data protection law*."[77] They argue that the WP29's "*near-zero probability"* understanding of acceptable re-identification risk constitutes "*an idealistic and impractical standard that cannot be guaranteed in a big data era".*[78] In reality, the line between personal data and non-personal data is fluid and evolves over time.[79]

In the same vein, anonymisation can arguably never be absolute. As such, "if the law were to insist that it must be, the only logical conclusion would be that data that once was personal data can only ever be pseudonymised but never anonymised. […] Any information that was ever in the scope of the GDPR would need to be presumed to be forever within that scope”.[80] In the mobility context, this can likely be interpreted as suggesting that the application of the GDPR will most likely be the default scenario. Therefore, if processing personal data is necessary to attain specific mobility-related purposes, then the focus should be on picking the right tools from the GDPR toolbox to achieve such processing in a legally compliant manner.

Based on the above analysis, we provide the following recommendations:

1. **Consider the available toolbox under the GDPR to ensure lawful processing of personal mobility data**

The GDPR already offers the necessary toolbox to ensure lawful processing of personal (mobility) data, for example, privacy by design, abidance with the principles of data processing (e.g. purpose limitation, data minimisation) and the application of privacy enhancing techniques. In this respect, consider the analysis carried out under D2.3. Special attention should be given to data that are considered particularly sensitive, such as geolocation data.

[76] EDPB Guidelines on the concepts of controller and processor (n 47), para.67.
[77] STALLA-BOURDILLON, S. and KNIGHT, A., "Anonymous v. Personal Data – False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data", *Wisconsin International Law Journal* 2016, 287.
[78] *Ibid*, 298.
[79] *Ibid*, 318.
[80] FINCK, M. and PALLAS, F. (n 9), 36.

Funded by the European Union

2. Perform a risk-benefit analysis of using anonymous vs personal data and explore the use of synthetic data

As mentioned above, it is not possible to draw a definite line between anonymous and personal data. A comparison between the potential risks from the use of personal data and the utility of anonymous data must be determined on a case-by-case basis, taking into consideration the risk of re-identification of anonymous data based on the means available and likely to be used at a given time.

In addition, an alternative to data anonymisation and privacy-preserving techniques has gained ground in the past years: generating synthetic data.[81] Synthetic data are data that are artificially created, typically using mathematical models or algorithms, to resemble real-world data. They share similar statistical properties, patterns, and characteristics with real data but do not (in theory) contain any information about specific individuals or entities.[82] They can be used for various applications, such as developing and testing analytics processes, training and testing machine learning models, and performing simulations, while ensuring data privacy.

The use of synthetic data is motivated by several constraints. First is missing data and data availability in general. Data processing, and in particular machine learning, demands high-quality data in large quantities. Often, synthetic data are used as a tool to complete missing values or to *augment* the available data.[83] When completing missing values, one can use the mean of the sample for particular attributes, regression based on available values, or, in the case of mobility data, interpolation between GPS points. For data *augmentation*, i.e., increasing the amount of available data, transformation of existing data (for example, scaling and rotations in image data) and generation of new data (for example, with generative ML models) techniques are often used.

For years, synthetic data has been used to protect the privacy of respondents[84] and, in recent years, improvements in generation mechanisms based on machine learning and growing privacy concerns and related regulations have increased the interest of synthetic data generation. The idea is that randomly generated data, even if their structure or statistical measures are preserved, do not belong to any individual. Thus, re-identification or any other privacy violations are avoided.

This protection is rather nuanced since synthetic data, especially high-quality synthetic data, may unintentionally contain data that matches that of some real individual. It has also been argued that even synthetic data imply a certain risk of re-identification, which can, to a large extent, be controlled by design.[85] From a legal perspective, there is no consensus on the legal nature of synthetic data. They can be considered personal or anonymous data based on whether the identifiability threshold set by the GDPR is met.

[81] Farrell, E.; Minghini, M.;Kotsev, A.; Soler-Garrido, J.; Tapsall, B.; Micheli, M.; Posada, M.; Signorelli, S.; Tartaro, A.; Bernal, J.; Vespe, M.; Di Leo, M.; Carballa-Smichowski, B.; Smith, R.; Schade, S.; Pogorzelska K.; Gabrielli, L.; De Marchi, D., *European Data Spaces: Scientific insights into data sharing and utilisation at scale*, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/400188, JRC129900.

[82] Rubin, D.B., 1993. Statistical disclosure limitation. Journal of official Statistics, 9(2), pp.461-468.

[83] Reiter, J.P., 2004. Simultaneous use of multiple imputation for missing data and disclosure limitation. Survey Methodology, 30(2), pp.235-242. Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J. and Greenspan, H., 2018, April. Synthetic data augmentation using GAN for improved liver lesion classification. In 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) (pp. 289-293). IEEE.

[84] Rubin, D.B., 1993. Statistical disclosure limitation. Journal of official Statistics, 9(2), pp.461-468.

[85] Farrell, E.; Minghini, M.;Kotsev, A.; Soler-Garrido, J.; Tapsall, B.; Micheli, M.; Posada, M.; Signorelli, S.; Tartaro, A.; Bernal, J.; Vespe, M.; Di Leo, M.; Carballa-Smichowski, B.; Smith, R.; Schade, S.; Pogorzelska K.; Gabrielli, L.; De Marchi, D., *European Data Spaces: Scientific insights into data sharing and utilisation at scale*, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/400188, JRC129900.

Funded by the European Union

The considerations of “zero-risk” vs “acceptable risk” also come into play in this context. Supporters of synthetic data as anonymous data rather support an “acceptable-risk” approach, whereas those that oppose them rather identify themselves with the “zero-risk” approach.[86]

3. Determine GDPR responsibilities based on each scenario

It is evident from the above analysis that there are many determinants in identifying the exact role and relationships in a data sharing mobility scenario. As such scenarios vary depending on the envisaged aim, it is impossible to *a priori* set a roadmap that will determine who is responsible for GDPR compliance. Nevertheless, companies should use the current CJEU case law as guidance to determine whether they fit the tests set by the EU Courts.

4. Ensure accountability: documenting the entire data lifecycle

In line with the accountability principle underlying data protection law, a data controller will need to be able to justify and provide documentation on the processing of personal data and compliance with data protection legislation.[87] This also includes the very process of anonymisation itself, as this needs to be considered a processing of personal data. It is more importantly a processing that aims to create new data that is no longer personal, thereby placing this data outside of the scope of data protection law. Consequently, the data controllers should provide clear documentation, containing a full assessment of the anonymisation process, that supports the claim that the result of the anonymisation process indeed creates data that is no longer considered to be personal data under data protection law.

2.2. Competition law

2.2.1. Delineating a “market for data”

Article 102 TFEU deals with the unilateral conduct of a firm that holds a dominant position and acts in a manner that abuses that position. It provides that: “*Any abuse by one or more undertakings of a dominant position within the internal market or in a substantial part of it shall be prohibited as incompatible with the internal market in so far as it may affect trade between Member States.*

Such abuse may, in particular, consist in:

(a) Directly or indirectly imposing unfair purchase or selling prices or other unfair trading conditions;

(b) Limiting production, markets or technical development to the prejudice of consumers;

[86] On the legal nature of synthetic data’, César Augusto Fontanillo López, Abdullah Elbi, p.7 and the sources mentioned therein. https://openreview.net/pdf?id=M0KMbGL2yr.
[87] GDPR, Article 5(2).

Funded by the European Union

(c) Applying dissimilar conditions to equivalent transactions with other trading parties, thereby placing them at a competitive disadvantage;

(d) Making the conclusion of contracts subject to acceptance by the other parties of supplementary obligations which, by their nature or according to commercial usage, have no connection with the subject of such contracts".

Article 102 applies only where one undertaking has a "dominant position", or where two or more undertakings are "collectively dominant".[88] According to the ECJ in *United Brands*, a dominant position is "*a position of economic strength enjoyed by an undertaking which enables it to prevent effective competition being maintained on the relevant market by affording it the power to behave to an appreciable extent independently of its competitors, customers and ultimately of its consumers*".[89]

Before assessing dominance, the relevant product market and the geographic market need to be defined[90]**:**

- **Product market**: the relevant product market is made of all products/services which the consumer considers to be a substitute for each other due to their characteristics, their prices and their intended use. To determine this the so called SSNIP test is deployed (Small but Significant Non-transitory Increase in Price Test), i.e. if there is a 5% price increase in a product/service, would customers move to the other product/service under consideration?
- **Geographic market**: the relevant geographic market is an area in which the conditions of competition for a given product are homogenous.

The Court has found in *AKZO v Commission* that an undertaking with a market share of 50% or more will be presumed dominant.[91] .[92] In the Guidance, the Commission also notes that market shares provide a useful first indication of the market structure and of the relative importance of the various undertakings active on the market, but the Commission will interpret market shares in the light of the relevant market conditions, and in particular of the dynamics of the market and of the extent to which products are differentiated.[93]

Looking at data, any test that relies on price in practically inapplicable. As far as the Commission is concerned, it has not yet had to define a market for personal data or for any of its particular usages.[94] In its *Facebook/WhatsApp* merger decision, the Commission explicitly stated that it had not investigated any possible market definition concerning the provision of data or data analytics services, since neither of the parties involved was active in any such potential markets.[95]

[88] Richard Wish, David Bailey, *Competition Law* (8th edn, Oxford University Press 2015), p.190.
[89] Case C-27/76, *United Brands v Commission* [1978], ECLI:EU:C:1978:22, para. 65.
[90] https://ec.europa.eu/competition-policy/system/files/2021-05/antitrust_procedures_102_en.pdf
[91] Case C-62/86, *AKZO v Commission* [1991], ECLI:EU:C:1991:286, para. 60.
[92] *Whish, Bailey* (n 88), p.194.
[93] Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings (n 164), para. 13.
[94] Graef, Inge, 'Market Definition and Market Power in Data: The Case of Online Platforms' (2015), World Competition: Law and Economics Review, Vol. 38, No. 4 (2015), p. 489., Available at SSRN: <https://ssrn.com/abstract=2657732> or <http://dx.doi.org/10.2139/ssrn.2657732>,
[95] European Commission Case No COMP/M.7217 – Facebook/WhatsApp, para. 72.

Under current competition law standards, a correct market definition requires the existence of supply and demand for the product or service.[96] Since all online platforms that have been under scrutiny by the Commission do not trade data, a relevant market was not possibly identified. The Commission delineated the relevant market for online platforms around the services or functionalities offered (e.g. web search, online search advertising intermediation[97]).

But once data are established as a tradeable good, which the Commission pursues through the creation of an "internal market for data", a 'real' data market may arise. There are, however, some examples where the Commission identified a service of "selling data to consumers" as the relevant market.[98] For example, in *TomTom/Tele Atlas*, the Commission identified the provision of navigable digital map databases as a relevant product market. In *Publicis/Omnicom*, a market for marketing data services was defined, further segmented into marketing information services, market research services and media measurement services. In *IMS Health*, the ECJ accepted the definition of a "potential" market for data ("brick structure") which was at the same time an input for the product sold ("studies").[99]

The question is how the existence of a dominant position in a market for data can be measured and in particular how value can be attributed to data.[100] It has been argued that the amount or quality of data that an undertaking controls do not seem to constitute adequate indicators for market power because the datasets of different providers cannot be easily compared in this regard. It may be hard, if not impossible, to distinguish different pieces of information and assign value to each of them individually. A more objective way to measure the competitive strength of providers active in a market for data would be to look at their ability to monetise the collected information. The revenue gained by a provider through licensing of data to third parties, delivering targeted advertising services, or offering other paid products and services to customers having data as input indicates how successful it is in the market.[101]

Graef provides the following non-exhaustive conditions that may point towards a potential market power in a data-related market: "(1) data is a significant input into the end products or services delivered on online platforms; (2) the incumbent relies on contracts or on intellectual property and trade secret law to protect its dataset as a result of which competitors cannot freely access the necessary data; (3) there are few or no actual substitutes readily available on the market for the specific information needed to compete on equal footing with an incumbent; (4) it is not viable for a potential competitor to collect data itself to develop a new dataset with a comparable scope to that of the incumbent (for example due to network effects or economies of scale and scope)".[102]

[96] Commission Notice on the definition of relevant market for the purposes of Community competition law, 97/C 372 /03, paras 13-23.
[97] https://ec.europa.eu/commission/presscorner/detail/en/IP_13_371.
[98] Édouard Bruc (2019) Data as an essential facility in European law: how to define the "target" market and divert the data pipeline?, European Competition Journal, 15:2-3, 177-224, DOI: 10.1080/17441056.2019.1644576, p.187.
[99] *Ibid*, p.180.
[100] *Graef* (n 94), pp.501-502.
[101] *Ibid*
[102] *Ibid*, pp.504.

Bruc, on the other hand, argues that quality is also a variable of competition and must be subsumed in the analysis along with other variables such as innovation, with the cross-network effects (of platforms) determining the boundaries of the market.[103]

He suggests four possibilities: (i) stretching the SSNIP test to the paying side of the platform; (ii) focusing on the cost of the operation and considering data as a currency; (iii) measuring the substitutability through quality and; (iv) relying less on market shares and focus analysis on market power through the market dynamics.[104] He then suggests moving towards a “holistic inter-dynamic analysis test” which will adopt a dynamic rather than a static approach of market analysis, looking at both sides of a platform, both supply side and demand side.[105]

2.2.2. Conclusions and recommendations

It is evident from the above analysis that the debate is not mature enough in the literature currently to explore how to define a “market for data” where data itself is the commodity. This debate is likely to become more pertinent once the Data Spaces discussion evolves in the coming years. Therefore, the recommendation to be provided is to continue monitoring for European Commission competition law cases that may develop such market analysis and anything relevant analyzed in the literature.

[103] *Bruc* (n 98), p.187.
[104] *Ibid*, p.189.
[105] *Ibid*, p.189-194.

Funded by the European Union

3. Mobility-as-a-Service (MaaS)

This section is inspired by MobiDataLab's use case on Mobility as a Service as analysed in deliverable D2.9. As mentioned in D2.9, mobility takes new forms, with new behaviours and new services –especially in urban areas. Mobility becomes a mixture of different modes of transport, both individual transport solutions (either cars, bicycles, e-scooters, etc.) and new forms of public transport like ride pooling and ridesharing. In order to motivate users to use and mix these different modes of transport, it is important to offer them a complete end to end solution that will not only allow them to plan their journey, but also to book and pay their ticket for the complete journey. Given the importance of this use case, we have decided to also examine it from a legal perspective.

In deliverable D2.1, we identified the following regulatory gaps concerning MaaS:

Table 3 - Overview of MaaS legal and regulatory gaps

MaaS - legal and regulatory gaps			
Lack of EU wide definition of MaaS that sets out the characteristics of the service.	Lack of understanding of the different actors that participate in a MaaS ecosystem, their precise role and the corresponding obligations under the GDPR.	Lack of clarity on whether the "essential facilities doctrine" under competition law could be applicable to oblige private operators to share data – and if so, how.	The ITS Delegated Regulations do not include in the scope of data that need to be made available via NAPs fare data

The following sections seeks to analyse each gap identified and provide relevant recommendations.

3.1. Lack of EU-wide MaaS definition

3.1.1. Defining MaaS

MaaS is a concept that has arisen in the past years seeking to transform traditional transport provision with a view to improving travellers' experience, addressing transport challenges (primarily at local and urban level) and responding to sustainability and 'greener mobility' demands. It allows users to plan, book and pay for multiple types of transport services using a single interface (typically a mobile application). The objective is to offer to MaaS users mobility solutions that are ideally suited to their individual needs, while letting the MaaS operator the opportunity to promote alternative transport modes that users may not have considered otherwise. This bundling of different mobility means seeks to create a shift away from an ownership-based transport system toward an access-based one, where travellers will substitute their private cars with other means.[106]

[106] Jittrapirom, Peraphan & Caiati, Valeria & Feneri, Anna Maria & Ebrahimigharehbaghi, Shima & Alonso Gonzalez, Maria & Narayan, Jishnu. (2017). Mobility as a Service: A Critical Review of Definitions, Assessments of Schemes, and Key Challenges. Urban Planning. 2. 10.17645/up.v2i2.931.

However, to date, a commonly accepted definition does not yet exist.[107] There are many discussions and disagreements around what MaaS is, what are its core components with academics and stakeholders focusing on different aspects. The first comprehensive definition was provided by Hietanen (owner of the company MaaS Global). He described MaaS as a mobility distribution model delivering users' transport needs through a single interface of a service provider bundling services to a package, similar to a monthly mobile phone contract. He also stressed the importance of integrations between transport data, data infrastructure and physical transport infrastructure.[108]

Holmberg, Collado, Sarasini and Williander noted the element of subscription in MaaS, requiring travellers to register to use the service and thereby giving them the possibility to plan their journey, in terms of booking and paying the several transport modes that might be required, all in one service. Subscription can further result in personalisation of offers, thereby leading to an experience suited to each traveller's needs. Other options include 'pay-as-you-go' or pre/post pay, considering their registration and a monthly subscription.[109]

Kamargianni et al. continued to narrow the scope and specify that MaaS refers to the offer of bundled services. In this case, the planning, booking, and payment of different forms of mobility through one platform, as well as the availability of mobility packages, were named as important criteria.[110] Atkins defines MaaS as a new way to provide transport, which facilitates the users to get from A to B by combining available mobility options and presenting them in a completely integrated manner. Thus, it is possible to consider MaaS as mobility service that is flexible, personalized and on-demand. Giesecke et al. conceptualize MaaS as a socio-technical phenomenon with sustainability aims at the core of the concept.[111]

Hensher summarized MaaS with "three Bs"—bundle, budget, and broker, considering also the innovation to be the networking aspect and the novelties presented by providers. The packages bought are linked to services such as the flexible selection of starting locations and starting times and are offered to the customer in a personalized way, e.g., depending on age, location, or passenger volume.[112] Another critical aspect of MaaS relates to the integration of on-demand transportation services on top of public transport, providing solutions for the first/last miles.[113] That said, public transport plays a central – if not the most important - role in MaaS.

Mobility stakeholders have provided their own definition of MaaS, as shown in the table below:

Table 4 - Overview of different stakeholder MaaS definitions

Overview of different stakeholder MaaS definitions	
MaaS Alliance	The integration of various forms of transport services into a single mobility service accessible on demand. To meet a customer's request, a MaaS operator facilitates a diverse menu of transport options, be they public transport, ride-, car-, or bike-sharing, taxi or car rental/lease, or a combination

[107] Maas B. Literature Review of Mobility as a Service. *Sustainability*. 2022; 14(14):8962. https://doi.org/10.3390/su14148962
[108] https://silo.tips/download/sampo-hietanen-ceo-its-finland.
[109] *Jittrapirom et al* (n 106).
[110] *Maas B* (n 107).
[111] *Jittrapirom et al* (n 106).
[112] *Maas B* (n 107).
[113] *Ibid.*

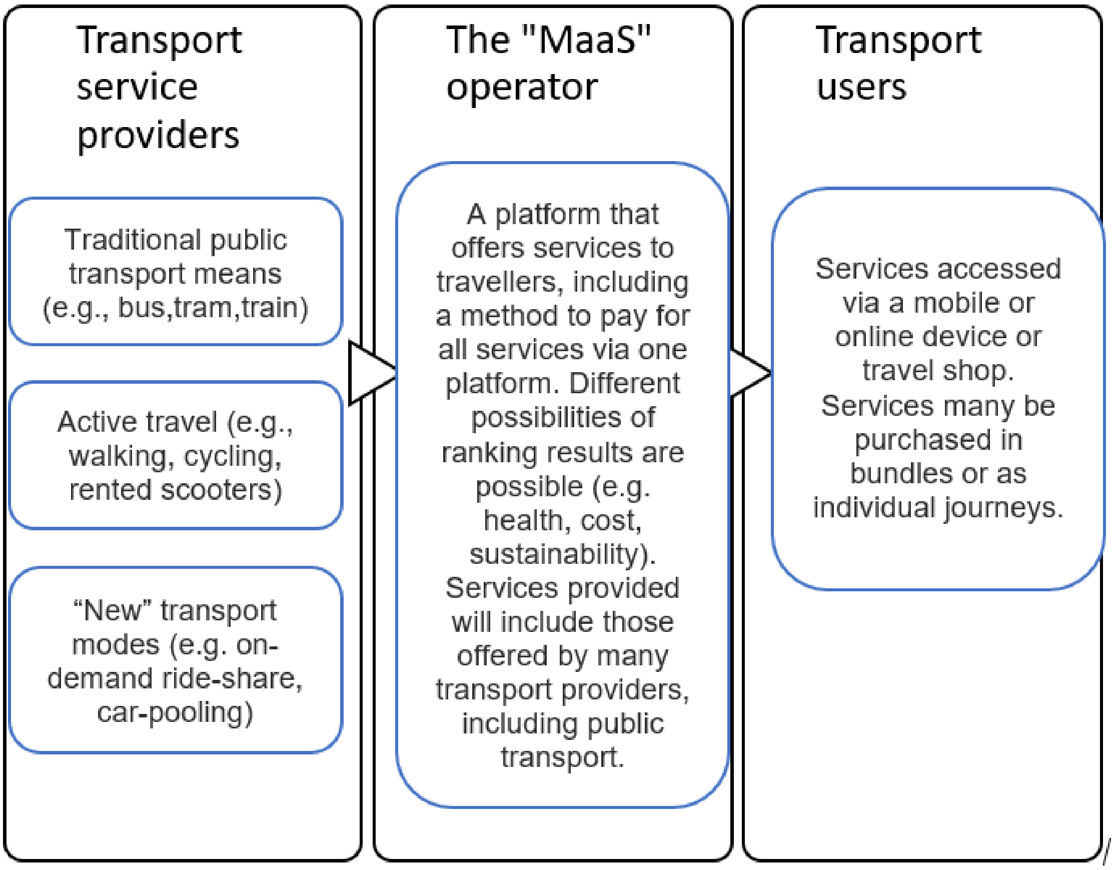
	thereof. For the user, MaaS can offer added value through use of a single application to provide access to mobility, with a single payment channel instead of multiple ticketing and payment operations. For its users, MaaS should be the best value proposition by helping them meet their mobility needs and solve the inconvenient parts of individual journeys and the entire system of mobility services".[114]
UITP	The integration of different transport services (such as public transport, ridesharing, car-sharing, bike-sharing, scooter-sharing, taxi, car rental, ride-hailing and so on) in one single digital mobility offer**,** with active mobility and an efficient public transport system as its basis.[115]
International Transport Forum (ITF)	Mobility as a Service (MaaS) is a distribution model for mobility services that uses shared data and a digital interface to efficiently source and manage the provision of transport related services into a seamless offer. It is typically delivered via a MaaS app, which is a single, digital, customer interface that sources and manages travel related services and improves the ease of planning, booking and making journeys in a region. MaaS joins different transport, information and payment services into a smooth and reliable digital customer experience. It enables the integration of public transport (PT) modes, commercial transport services such as ridesourcing, bike and carsharing, and taxis into a comprehensive mobility offer.[116]

3.1.2. Core characteristics

In its simplest form, MaaS works as follows[117]:

[114] MaaS Alliance, "What is MaaS?", https://maas-alliance.eu/homepage/what-is-maas/.
[115] https://www.uitp.org/trainings/mobility-as-a-service-maas/.
[116] International Transport Forum (2021), *Developing Innovative Mobility Solutions in the Brussels-Capital Region*, https://www.itf-oecd.org/developing-innovative-mobility-brussels-capital-region.
[117] Brown, C.; Hardman, M.; Davies, N.; Armitage, R. Mobility as a Service: Defining a Transport Utopia. Future Transp. 2022, 2, 300–309. https://doi.org/10.3390/futuretransp2010016

Funded by the European Union


Transport service providers
Traditional public transport means (e.g., bus,tram,train)
Active travel (e.g., walking, cycling, rented scooters)
“New” transport modes (e.g. on-demand ride-share, car-pooling)
The "MaaS" operator
A platform that offers services to travellers, including a method to pay for all services via one platform. Different possibilities of ranking results are possible (e.g. health, cost, sustainability). Services provided will include those offered by many transport providers, including public transport.
Transport users
Services accessed via a mobile or online device or travel shop. Services many be purchased in bundles or as individual journeys.

Jittrapirom et all have identified the following as key characteristics of MaaS[118]:

Table 5 - Key MaaS characteristic

Core characteristic	Description
Integration of transport modes	A goal of MaaS schemes is to encourage the use of public transport services, by bringing together multi-modal transportation and allowing the users to choose and facilitating them in their intermodal trips. Following transport modes may be included: public transport, taxi, car-sharing, ride-sharing, bike-sharing, car-rental, on-demand bus services. Envisioning a service beyond the urban boundaries, it will embrace also long-distance buses and trains, flights, and ferries.
Tariff option	MaaS platform offers users two types of tariffs in accessing its mobility services: “mobility package” and “pay-as-you-go”. The package offers bundles of various transport modes and includes a certain amount of km/minutes/points that can be utilized in exchange for a monthly payment. The pay-as-you-go charges users according to the effective use of the service.
One platform	MaaS relies on a digital platform (mobile app or web page) through which the end-users can access to all the necessary services for their trips: trip planning, booking, ticketing, payment, and real-time information. Users might also access to other useful services, such as weather forecasting, synchronization with personal activity calendar, travel history report, invoicing, and feedback.

[118] *Jittrapirom et al* (n 106).

Funded by the European Union

Multiple actors	MaaS ecosystem is built on interactions between different groups of actors through a digital platform: demanders of mobility (e.g. private customer or business customer), a supplier of transport services (e.g. public or private) and platform owners (e.g. third party, PT provider, authority). Other actors can also cooperate to enable the functioning of the service and improve its efficiency: local authorities, payment clearing, telecommunication and data management companies.
Use of technologies	Different technologies are combined to enable MaaS: devices, such as mobile computers and smartphones; a reliable mobile internet network (WiFi, 3G, 4G, LTE); GPS; e-ticketing and e-payment system; database management system and integrated infrastructure of technologies (i.e. IoT).
Demand orientation	MaaS is a user-centric paradigm. It seeks to offer a transport solution that is best from customer's perspective to be made via multimodal trip planning feature and inclusion of demand-responsive services, such as taxi.
Registration requirement	The end-user is required to join the platform to access available services. An account can be valid for a single individual or, in certain cases, an entire household. The subscription not only facilitates the use of the services but also enables the service personalisation.
Personalisation	Personalisation ensures end users' requirements and expectations are met more effectively and efficiently by considering the uniqueness of each customer. The system provides the end-user with specific recommendations and tailor-made solutions on the basis of her/his profile, expressed preferences, and past behaviors (e.g. travel history). Additionally, they may connect their social network profiles with their MaaS account.
Customisation	Customisation enables end users to modify the offered service option according to their preferences. This can increase MaaS' attractiveness among travellers and its customers' satisfaction and loyalty. They may freely compose a specified chained trip or build their mobility package with a different volume of usage of certain transport modes to better achieve their preferred travel experiences.

3.1.3. Legal definition and the legal status of the MaaS provider

From a legal perspective, the only definition of integrated mobility services currently available at the EU level can be found in the 2017 Finnish Act on Transport Services which deals specifically with MaaS platform providers. According to the Act, integrated mobility services refer to the "formation of travel chains and other service packages in return for remuneration by combining the mobility services offered by different service providers, excluding travel packages or combined travel arrangements falling within the scope of the Act on Travel Service Combinations".[119]

The European Commission has introduced the term "Multimodal Digital Mobility Services (MDMS)" in its proposal for a new Regulation[120] that seems to incorporate but takes a broader scope than MaaS services. The EC defines these services as "systems providing information about, inter alia, the location of transport facilities, schedules, availability and fares, of more than one transport provider, with or without facilities to make reservations, payments or issue tickets" clarifying that it could include route-planners, MaaS, online ticket vendors and ticket intermediaries. It clarifies that they help both passengers and/or other intermediaries compare different travel options, choices and prices, and can facilitate the sale and re-sale of mobility products from different operators, whether they are private or public, within one mode or across modes.

[119] Björn Lundqvist and Erion Murati, 'Collaborative Platforms and Data Pools for Smart Urban Societies and Mobility as a Service from a Competition Law Perspective' in Michèle Finck, Matthias Lamping, Valentina Moscon, Heiko Richter (eds) *Smart Urban Mobility, Law, Regulation and Policy* (Springer 2020), p.192.

[120] https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13133-Multimodal-digital-mobility-services_en

It should be taken into consideration that the definition of the service and the characteristics entailed thereof can have consequences on the legal status of the MaaS provider and the contractual liability towards users. In principle, such platforms would qualify as "intermediaries" under the e-Commerce Directive[121] ("ECD") which regulates legal aspects of information society services, i.e. any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services[122]. One issue under this Directive is that under certain conditions, they can escape liability (under civil, criminal or administrative national laws) – the so-called "safe harbour" provisions. Generally, this takes place when they act as "mere conduits", i.e. they do not curate the data passing through their services nor do they intervene in some way.

In its 2017 judgment in the case of Asociación Profesional Elite Taxi v Uber Systems Spain, SL, the CJEU ruled that "an intermediation service such as that at issue in the main proceedings [i.e. Uber], the purpose of which is to connect, by means of a smartphone application and for remuneration, non-professional drivers using their own vehicle with persons who wish to make urban journeys, must be regarded as being inherently linked to a transport service and, accordingly, must be classified as 'a service in the field of transport".[123] The justification was that Uber did more than just be an intermediary between a driver and a user: "in that regard, it follows from the information before the Court that the intermediation service provided by Uber is based on the selection of non-professional drivers using their own vehicle, to whom the company provides an application without which (i) those drivers would not be led to provide transport services and (ii) persons who wish to make an urban journey would not use the services provided by those drivers.

In addition, Uber exercises decisive influence over the conditions under which that service is provided by those drivers. On the latter point, it appears, inter alia, that Uber determines at least the maximum fare by means of the eponymous application, that the company receives that amount from the client before paying part of it to the non-professional driver of the vehicle, and that it exercises a certain control over the quality of the vehicles, the drivers and their conduct, which can, in some circumstances, result in their exclusion". The Court concluded that Uber must thus be regarded as forming an integral part of an overall service whose main component is a transport service and, accordingly, must be classified not as 'an information society service' under the ECD, but as 'a service in the field of transport' within the meaning of Article 2(2)(d) of Directive 2006/123.[124]

In another case, Star Taxi App, the CJEU concluded that the app was an information society service. Star Taxi App was a smartphone application linking customers and professional taxi drivers. When customers connect to the app, they receive information about taxi drivers available, their prices as well as comments about drivers left by other users. In order to be referenced by Star Taxi App, taxi drivers concluded contracts with the business, without the latter making any particular selection, and pay a monthly fee to the business.

[121] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000, OJ L 178, 17.7.2000, p. 1–16.
[122] ECD, recitals 17-18, Article 2(a).
[123] https://curia.europa.eu/juris/document/document.jsf?text=&docid=198047&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=1108338.
[124] C-434/15 - Asociación Profesional Elite Taxi, Judgment of the Court (Grand Chamber) of 20 December 2017, paras 39-40.

The Court considered that "intermediation service which consists in putting persons wishing to make urban journeys in touch, using a smartphone application and in exchange for remuneration, with authorised taxi drivers, for the purposes of which the service provider has entered into contracts for the provision of services with those drivers, in consideration of the payment of a monthly subscription fee, but does not forward the bookings to them, does not determine the fare for the journey or collect it from the passengers, who pay it directly to the taxi driver, and exercises no control over the quality of the vehicles or their drivers, or over the conduct of the drivers, constitutes an 'information society service' within the meaning of those provisions".[125]

The Court considered that the Star Taxi App case was different from the Uber one in the following ways[126]:

- Star Taxi App connected customers and professional drivers who were already registered. So the app was just an extra means of attracting customers. Conversely, in Uber, the application allowed non-professional drivers previously absent from the market to provide a service via a single channel.
- When concluding contracts between drivers and Star Taxi App, the latter had no word in the details and did not exert any special control. The drivers were free to set their own prices, which they collected directly (without going through the platform). In Uber, the CJEU recognised the existence of a real economic and functional dependence between the drivers and Uber, since Uber had created an innovative collective urban transport service involving a selection of drivers.

On that basis, a MaaS provider may hypothetically undertake one of the following legal roles[127]:

- A comprehensive service provider acts as a principal and is responsible towards passengers for ensuring that transport in the travel chain or replacement transport is carried out. Here the platform delivers the digital and the underlying service and plays a predominant role in the defining and/or delivering of the material service.
- An intermediary, which means it would offer a platform for connecting service providers and customers, mainly regulated by the ECD as an information society service. Here MaaS provider concludes intermediation contracts, on one hand, with the passenger for the digital services, on the other hand, with the transport providers for the promotion and advertisement the said service. A third contract is concluded, for the provision of the underlying service, between the transport carrier and the passenger. An intermediary is not a party to the contract and is not responsible for its performance.
- A ticket vendor, which means under bus and coach regulation (EU) no. 181/2011 (Bus Regulation) "any intermediary concluding transport contracts on behalf of a carrier".
- A travel agent, which means under Bus Regulation "any intermediary acting on behalf of a passenger for the conclusion of transport contracts".

[125] https://curia.europa.eu/juris/document/document.jsf?text=&docid=234921&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=1110648.

[126] https://www.stibbe.com/publications-and-insights/after-the-uber-case-and-the-airbnb-case-the-star-taxi-app-case-focus-on.

[127] As identified by Erion Murati in "Mobility-as-a-service (MaaS) digital marketplace impact on EU passengers' rights", European Transport Research Review (2020) 12:62.

Funded by the European Union

Depending on the nature of its operations, the role of a MaaS provider may also be a hybrid between those discussed here. However, the main difference between these typologies relates to their applicable legal regime, users' rights and MaaS provider's liabilities. In principle, in the position of intermediary, ticket vendor, travel agent and provider of a linked travel arrangement, MaaS provider does not assume any responsibility for the underperformance/breach of contract of the material service.[128]

3.2. Understanding the MaaS ecosystem actors

The MaaS ecosystem can comprise several actors, as demonstrated in the figure below:

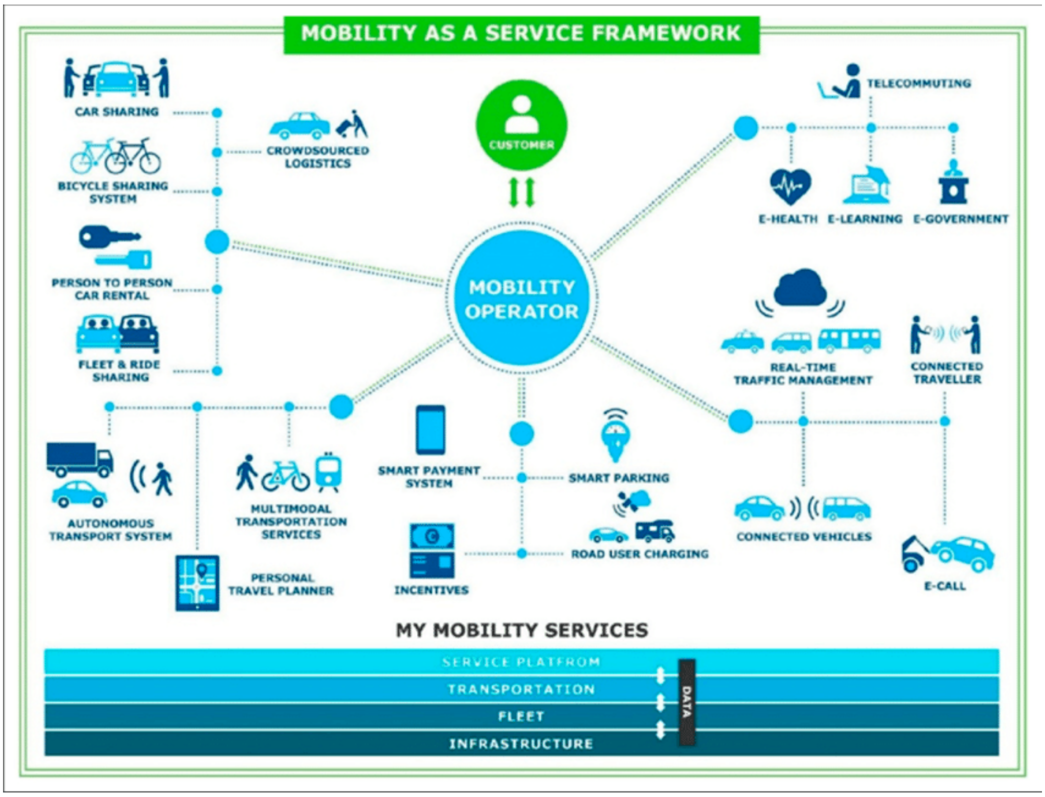


Figure 2 - The MaaS framework (Reproduced from Kivimäki et al.) [129]

[128] *Ibid*.

[129] https://www.researchgate.net/publication/338156972_State_of_the_Art_of_Mobility_as_a_Service_MaaS_Ecosystems_and_Architectures-An_Overview_of_and_a_Definition_Ecosystem_and_System_Architecture_for_Electric_Mobility_as_a_Service_eMaaS,

Funded by the European Union

3.2.1. Business roles

A 2022 UITP Handbook[130] identifies the following business roles in a MaaS ecosystem:

- **Mobility Service Provider (MSP)**: the transit providers (bus, metro, rail), taxi, shared mobility providers (bike sharing, car sharing, scooter sharing, etc), on demand transport, car rental. They can be public or private.
- **MaaS Operator**: the digital operator that, through a single front-end, conveys multiple mobility services to end users in a seamless multimodal way. The MaaS Operator handles end-user requests for mobility, providing customer support, handling payment and managing how to distribute the fees across all MSPs through a compensation engine.
- **MaaS Integrator**: a transport software integrator, with the mission of putting to system various transport services for one or more MaaS Operators. The MaaS Integrator integrates data and services from all MSPs and creates multimodal value-added services for information, planning, booking, payment, and travel support, making them available to MaaS operators via suitable, ideally open and standard interfaces.
- **Transport Authority's role**

A 2022 OECD background note on Competition Policy identifies actors at three different levels: the public authority regulating public transport and possibly the MaaS platform, transport service providers (public and private) and an aggregator digital platform, whether provided by one of the service providers or a third party.[131]

3.2.2. Business models

There is no universal MaaS model and configurations can vary depending on the different business models adopted. According to an ITF/OECD report, business models for MaaS are nascent, involve interactions between settled service delivery models and emerging ones, and are developing under unclear longer-term market dynamics and regulatory frameworks.[132] The report suggests that early MaaS business models focused either on business-to-customer (B2C) or business-to-government-to-citizen (B2G2C) relationships.

In B2C markets, MaaS providers seek to create, customise and market services to the public such that remunerative margins can be achieved. UbiGo (www.ubigo.me) and MaaS Global (Whim – whimapp.com) are examples of this approach. In B2G2C markets, a public entity serves as the MaaS service provider for the whole market. The MaaS platforms in Berlin and Madrid are examples of this approach.

[130] UITP, 'Ticketing in Mobility as a Service', Handbook, July 2022, https://www.uitp.org/publications/ticketing-in-mobility-as-a-service/.

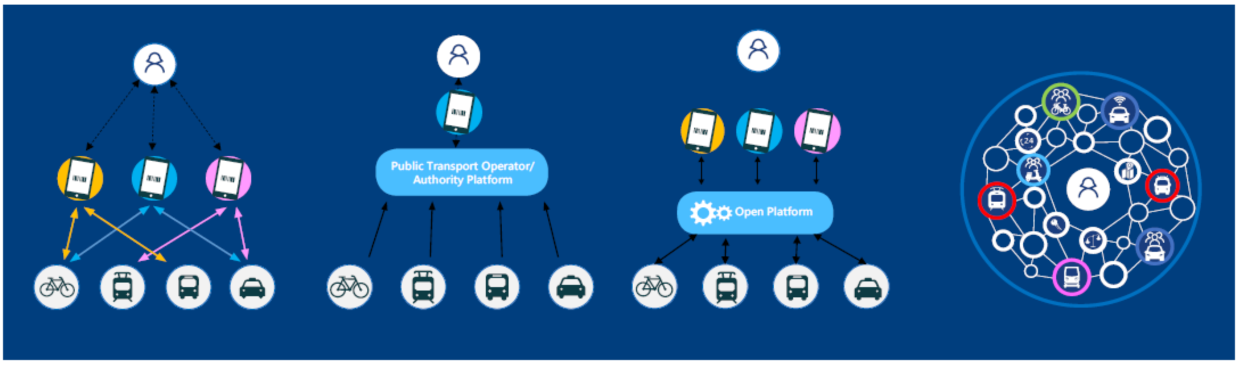
[131] OECD (2022), Competition and Regulation in the Provision of Local Transportation Services, OECD Competition Policy Roundtable Background Note, www.oecd.org/daf/competition/competition-and-regulation-in-the-provision-of-local-transportation-services-2022.pdf.

[132] 'The Innovative Mobility Landscape: the case of Mobility as a Service", the International Transport Forum, 2021.

Funded by the European Union

B2G2C MaaS offers build-on open platforms where affiliation is conditioned on certain (minimal) requirements that ensure alignment with public policy objectives. B2G2C models are seen as inevitable and a natural end-point, where proponents see parallels between the organisation of MaaS markets and the market organisation of public transport. Another form of MaaS market organisation targets the business-to-business (B2B) market. Some B2B MaaS offers aim to help employers manage their employees' work travel and provide added value through management and administrative services to ease the burden on the employer.[133]

The report identifies the four most common business models. These include the following: Walled Gardens, Public MaaS, Regulated Utility MaaS and "Mesh-y-Maas". These are demonstrated in the figure below:



Source: ITF (2020), based on UITP (2019).

Figure 3 - Different Maas business models

Walled gardens occur when a primary mobility service operator, for example, a ride-sourcing company, keeps its customer relationship/interface and integrates other modes or services, including digital wallets and payment services, that may interest their clients. This model is particularly prevalent in markets where public transport faces quality or convenience challenges and where new mobility services have been quick to capitalise on addressing latent travel demand.[134] This model could result in the creation of a monopolistic player if mobility service providers consistently work with just one platform (EMTA, 2019) with that player gathering all the data in its platform, limiting the potential to use it in support of improved planning (EMTA, 2019).[135]

In the ***Public MaaS*** model, the public transport operator or the public transport authority takes the role of MaaS aggregator. The public entity acts as a gatekeeper to the MaaS ecosystem and sets the terms for the integration of other mobility services onto the platform (and thus with public transport services). In practice, the actual operation of the platform may be undertaken in-house (likely for larger and better-funded public transport operators/authorities) or outsourced under contract.[136]

[133] *Ibid*, p.70.
[134] *Ibid*, p.71.
[135] *Ibid*, 2021, p.71.
[136] *Ibid*, p.71.

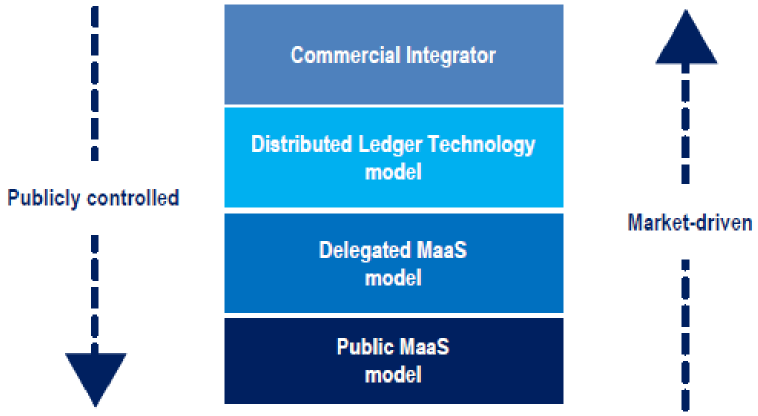
Funded by the European Union

The concern here is that the public entity may exert monopoly power or otherwise unduly favour its services at the expense of others, including new players. Any solutions developed are also likely to be limited to the local transport setting given the entity's local remit, restricting the possibility to develop on a wider scale (EMTA, 2019).[137]

In the ***Regulated utility MaaS*** model, the provision (and oversight) of public transport services and the public provision of a MaaS aggregation platform are divided. Oversight of the platform, including the setting and enforcement of platform access rules, is entrusted to the public entity with a broader remit than simply the provision of public transport. This model is characterised by a shared back-office aggregation platform that is treated as public infrastructure and that can be used by private MaaS providers who develop their own customer interfaces and apps. The MaaS app that the customers interact with can be owned by a separate entity from the MaaS platform, and multiple MaaS apps can operate off the one MaaS platform. The public MaaS platform is managed as a utility and provides common information and transaction integration across all MaaS market actors.[138] In practice, this integration can be delivered via a centralised platform or the adoption of common application programming interfaces (APIs) by all mobility service providers – the two are functionally equivalent in terms of market operation (but not in terms of public access to historical data).

Finally, the ***Mesh-y MaaS*** model builds on the distributed API model described above but integrates automated transaction processing, vetting and clearing based on distributed ledger technology (DLT) and automated contracts (ITF, 2018). In this model, the role of the aggregator is rendered obsolete through the execution of smart contracts directly between operators.[139]

The OECD note on Competition Policy presents the following models for the provision of MaaS:



Source: OECD Secretariat.

Figure 4 - Models for MaaS provision

[137] *Ibid*, p.71.
[138] *Ibid*, p.72. Upstream in Vienna uses this model and it also deployed in Lyon and is envisaged for the Ile de France region, with the regional mobility agencies housing and managing the platform.
[139] *Ibid*, p.73. An early DLT pilot in support of MaaS has been carried out as part of the official national MaaS trials programme in the Netherlands

Funded by the European Union

The Commercial Integrator approach seems to match the *Walled gardens* approach mentioned above, according to which the MaaS platform is run by a private operator who contracts with the different mobility service providers. Whim, the MaaS application in Helsinki, follows this approach.[140] Besides direct access to public transport, it integrates access to bike-sharing, car-sharing and even taxis and conventional rental cars. It also offers the possibility to book and pay directly via the application. In addition, the service offers several multimodal package offers, including a non-subscription formula, a basic subscription with fixed maximum charges per journey and a full-range subscription with unlimited access to all integrated transport services.[141] The Distributed Ledger Technology model seems to match the Mesh-y MaaS model while the Delegated MaaS model may coincide with the Regulated Utility MaaS.

In Vienna, there is a mix of Public Utility and Public MaaS. The local transport authority set up a database with data from various mobility services. Operators can use these data to develop their own apps. Today, the local transport operator Wiener Linien has developed its own MaaS WienMobil Service, after establishing a subsidiary (Upstream Mobility) specifically dedicated to MaaS in 2016. The MaaS app today shows all means of transport in Vienna, including electric scooters, bike-sharing, car-sharing and taxis.[142] But theoretically, other private operators could develop their own MaaS offering based on the public database.

Smith et al. argue that public mobility providers are instrumental in realising MaaS.[143] They propose the idea of MaaS integrators to facilitate the development and implementation of MaaS. So-called intermediary MaaS integrators (IMIs) collect offerings from transport service providers (TSPs) and distribute them to the operating MaaS providers. By focusing on the technical aspects and the handling of data, the use of IMIs reduces some of the main challenges involved in collaboration between transport and MaaS providers. Still, the use of IMIs is not limited to the technical aspect; rather, they support the overall development of MaaS and reduce the presence of barriers in the development process.[144] This proposition resembles the *Regulated utility MaaS* model proposed by the ITF/OECD report mentioned above.

3.3. Existence of an obligation to share data with MaaS providers under Competition law

The MaaS model is designed for the pooling of data and as a data-driven business model, it can trigger network effects (i.e. the more users make use of one application providing the MaaS service the more value it gains) and potentially cause the market to tip in favour of one dominant platform.[145] To arrive at that point, however, full integration of mobility providers is likely to be required. The more integration in the MaaS service, the better multimodal solution will be provided, as users will arguably have more choices and offers. But to provide that solution, access to data is necessary.

[140] OECD note on Competition and Regulation in the Provision of Local Transportation Services (n 131), p.19.
[141] *Ibid*, p.19.
[142] *Ibid*, p.19.
[143] Smith, G.; Sochor, J.; Karlsson, I. Intermediary MaaS Integrators: A case study on hopes and fears. Transp. Res. Part A Policy Pract. 2020, 131, 163–177.
[144] *Maas B* (n 107).
[145] *Lundqvist, Murati* (n 119), p.205.

Funded by the European Union

While offering services, MaaS gathers data from hundreds of billions of public and private transport journeys per year. There are multiple sources of data coming into play, including private, passive, community and self-quantification data. These data are typically held by governments, governmental organisations and local communities and include sensor, transport data and energy use figures. Private data may include proprietary information held by private firms or individuals.[146]

Looking at the different data categories in more detail:

a) Public transport data (provided by PTOs)*:*

- Static network description (lines, stop points etc): although static, the network description may be updated frequently***;***
- Real-time data (network disruptions, next departures, vehicle occupancy, vehicle position …): this information is updated continuously and will be refreshed very frequently (e.g. every 30 seconds)***;***
- Road traffic: when made available by PTO or PTA.

b) Geographical data*:*

- Cartography: could be provided by Open Street Map or other actors (e.g. Google Maps)***;***
- Addresses: national addresses databases are usually openly available (e.g. BANO in France)***;***
- Points of interests: could be provided by Open Street Map (user contributions), by MaaS main operator or by MaaS integrated transport operators***.***

c) Other transport data (provided by private transport operators)*:* free-floating, ride-sharing and road traffic data***.***

d) Booking and payment data:

- Static fares: a fare table provided by transport operators, or by the MaaS main operator if it has an agreement with transport operators to sell transport tickets to a different price. The price of each transport section should be displayed;
- Dynamic fares: the fare is calculated by the MaaS operator, eventually using the transport operators own fare systems, depending on various parameters (departure date, expected occupancy …). The price could be displayed only for the whole journey.

e) Ticketing data:

- e-Tickets: provided by transport operators to the MaaS operator, which in turn will create a single e-ticket or m-ticket for MaaS user.

[146] *Ibid.*

Funded by the European Union

f) User input data:

- User location: provided by the MaaS user, if he/she accepts to share his current location;
- Journey planning: preferred departures and arrivals, preferred transport modes, etc.;
- Personal details such as name, email, postal address (required for registration, booking and/or payment).

While static data are essential for information and planning purposes before the trip, dynamic travel data enable platforms to provide the best solution at any point in time to ensure time-saving. However, if MaaS platforms cannot access such data in a machine-readable and interoperable form, it becomes impossible to provide cost-effective and efficient services. This risk is particularly serious under a commercial integrator or delegated model, in which a vertically integrated public transport operator develops the MaaS platform and may have an incentive to deny access to its data to competing MaaS providers. Similarly, even in the absence of vertical integration, the public transport operator may deny access to its static and dynamic data to the extent that it perceives MaaS platforms and new mobility services as a competitive threat.[147] The seriousness of this concern was also highlighted by respondents to a recent call for evidence launched in the UK in the framework of the *Future of Transport Regulatory Review* (Department for Transport, 2020, p. 58). In particular, respondents stressed the danger that refusal to share data would have for a level playing field and open market entry.[148]

3.3.1. Refusal to supply essential facilities

One aspect more pertinent to the analysis is the case where the MaaS operator would like to access and use data of a given transport operator (for example, to make the app more attractive to users). Normally the two companies would enter into negotiations privately, but it may be the case that the (private) transport operator refuses to provide the data. Competition law could provide a solution by imposing a data access obligation if it could be considered that the transport operator – who must hold a dominant position in the relevant market – abused that dominant position by refusing to share data that is 'essential' or 'indispensable' for the other operator to develop the MaaS service.

In principle, companies are free to choose their contractual counterparties (generally known as contractual freedom). However, a dominant undertaking may be prohibited, in the absence of objective justification, to refuse to grant access to 'essential facilities' on a non-discriminatory basis to new customers, at least in circumstances where a refusal would eliminate effective competition on the downstream market.[149]

[147] OECD note on Competition and Regulation in the Provision of Local Transportation Services (n 131), p.19.
[148] OECD note on Competition and Regulation in the Provision of Local Transportation Services (n 131), p.19.
[149] Bellamy & Child, *European Union Law of Competition* (8th edn, Oxford University Press 2018), para. 10.149.

Funded by the European Union

The meaning of *essential facility* or *indispensability* is a fact-specific issue that depends upon the presence of technical, legal or even economic obstacles preventing the would-be user of the 'facilities' from competing in the relevant market.[150] The Advocate General Jacobs has explained what could constitute an essential facility in *Bronner*[151]:

"*An essential facility can be a product such as a raw material or a service, including provision of access to a place such as a harbour or airport or to a distribution system such as a telecommunications network. In many cases the relationship is vertical in the sense that the dominant undertaking reserves the product or service to, or discriminates in favour of, its own downstream operation at the expense of competitors on the downstream market. It may however also be horizontal in the sense of tying sales of related but distinct products or services*".

Under EU competition law, the 'essential facilities doctrine' has been developed in a long line of cases dealing with access to physical infrastructure as well as licensing of intellectual property rights. However, to our knowledge, there is no case at the EU level (yet) that has led to the obligation to provide access to data.[152] Several cases at the EU level can, however, be interpreted as relating to information assets more broadly. In *Magill*[153], the ECJ concluded that the refusal by three Irish broadcasting companies to provide the publishing company Magill with a copyright license for the weekly listings of their television programmes was abusive.

In the *Ladbroke* decision, the case concerned a refusal to grant a transmission licence (over i*ntellectual property rights on the sound and pictures of horse races)* by the horse race organiser to *Ladbroke* who needs the content for its betting services. The General Court held that because the race organisers and betting services were not rivals, refusal to grant a license did not interfere with the competition in the relevant market.[154] Where the data is also protected as intellectual property, European Courts further evaluate whether the benefits of compulsory sharing outweigh the expected benefits of IP protection (in consideration of the legitimate interests of the parties).[155]

In *IMS Health*, the ECJ found an abuse in the context of a refusal of IMS, a company active in providing data on regional sales of pharmaceutical products in Germany, to grant a license to its competitor NDC for the use of the copyrighted brick structure that IMS had developed and that had become a de facto standard. In *Microsoft*, the General Court held Microsoft's refusal to provide rivals with interoperability information necessary for non-Microsoft work group server operating systems to communicate with Microsoft's dominant client PC operating system Windows to be abusive.[156]

But the CJEU in its case law (*Magill, Bronner, IMS Health, Microsoft*) has set strict criteria for refusal to supply to be considered an abuse of a dominant position:

[150] *Ibid.*
[151] Case C-7/97, *Bronner*, Opinion of Advocate General Jacobs, 28 May 1998, para.50.
[152] Van Gorp, N., de Bijl, P., Graef, I., Molnar, G., Peeters, R., & Regeczi, D. (2020). *Exploring data sharing obligations in the technology sector*, p.37 https://www.government.nl/documents/reports/2020/11/30/exploring-data-sharing-obligations-in-the-technology-sector, p. 38.
[153] Joined cases C-241/91 and C-242/91, *Magill* [1995], ECLI:EU:C:1995:98.
[154] Case T-504/93 *Tiercé Ladbroke v. Commission* [1997] ECLI:EU:T:1997:84, para 133.
[155] In the *Microsoft* case, the defendant (Microsoft) argued that its refusal to supply was objectively justified on the grounds of: (i) the intellectual property rights, (ii) the fact that the technology which it was required to disclose to its competitors was secret, (iii) the circumstance that forced disclosure of the necessary protocols would adversely affect its incentives to innovate. Yet, Microsoft failed to persuade the Court of Justice in all of these grounds. See *Microsoft* Case. para 666.
[156] *Van Gorp N., de Bijl P., Graef I., Molnar G., Peeters R. & Regeczi D.* (n 152), p. 38.

Funded by the European Union

- The input or assets is indispensable for producing the downstream service (where the dominant firm is also active);
- The denial of access leads to exclusion of effective competition in the downstream market,
- And the dominant firm does not have an objective justification for denying access.[157]

The type of data has implications on the ability of potential competitors to gather or obtain the same dataset independently. Depending on the specificities of a given market, the different forms of collection and use of data such as individual-level data, anonymized data, aggregated-level data, and historical or real-time data are also of significance when deciding about the availability of substitute data. On the question of indispensability, the Report *Competition policy for the digital era* refers to the renowned categorisation of volunteered, observed, and inferred data as reflecting the capacity of competitors to access or obtain the same information or data independently.[158]

3.3.2. Data portability as a fall-back option for data access

GDPR

The right to data portability provided under Article 20 of the GDPR allows individuals to receive their personal data which they provided to the data controller based on consent (under GDPR Article 6(1)(a)) or a contract (under GDPR Article 6(1)(b)), in a format that is structured, commonly used and machine-readable, to make it easier to transmit it to another controller if they desire. When an individual, a data subject wishes to transmit the data they obtained using their right to data portability, the controller which gave this data to the data subject cannot hinder this transmission. In cases where it is technically feasible, the individual/data subject may also request their data to be transferred directly to another controller (GDPR Art 20(2)).

The right to data portability in Article 20 is an important tool, since it has the potential to empower individuals with more control over their personal data, as long as businesses/data controllers implement its requirements and regulators are keen on effectively enforcing the provision.[159] This mechanism enables competition by directly empowering the data subject to move their profile and data contained in one platform/app to another one.

Digital Markets Act

Provisions in the Digital Markets Act ("DMA") follow in the same spirit.[160] The DMA was initially intended for the biggest US digital platforms but its provisions may have a wider impact in the future on all platforms that are identified as "gatekeepers"[161].

[157] *Ibid.*
[158] J. Crémer, Y.-A. de Montjoye, and H. Schweitzer, "Competition policy for the digital era," Publications Office of the EU, 2019. Accessed: Oct. 29, 2021. [Online]. Available: https://data.europa.eu/doi/10.2763/407537.
[159] I. Graef, "Paving the Way Forward for Data Governance: a Story of Checks and Balances," *Technol. Regul.*, pp. 24-28 Pages, Jul. 2020, doi: 10.26116/TECHREG.2020.003.
[160] Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), OJ L 265, 12.10.2022, p. 1–66.
[161] The conditions for designating a digital platform a "gatekeeper" are laid out in Article 3 of the DMA.

Funded by the European Union

According to the DMA, they will have to provide business users with continuous and real-time data portability. Recital 9 of the DMA states that "*[t]he gatekeeper shall provide end users and third parties authorised by an end user, at their request and free of charge, with effective portability of data provided by the end-user or generated through the activity of the end user in the context of the use of the relevant core platform service, including by providing, free of charge, tools to facilitate the effective exercise of such data portability, and including by the provision of continuous and real-time access to such data"*.

The data portability provisions under the DMA seem to have a broader scope than the GDPR's right to data portability, ensuring additional forms of portability, including portability of non-personal data for business users. These obligations aim to limit gatekeepers' exclusive control over the data and allow business users to access and re-use data that they or end-users generate on the platform. However, it has been argued that the scope and implementation of the rules on data portability, data sharing and interoperability are not clear enough and could be further specified.[162]

While the right to data portability as enshrined in the GDPR and the DMA does give users in certain circumstances the possibility to transfer their data between different apps/platforms/operators, it does not entitle those operators to claim access to that data.[163] So it does not seem to be a viable option for operators to get access to certain datasets.

3.4. Lack of legislative obligation to share fare data

Under Delegated Regulation (EU) 2017/1926 on the provision of EU-wide multimodal travel information services ("MMTIS"), Member States are required to set up National Access Points (NAPs) to facilitate the exchange and re-use of data. The implementation of NAPs is important to allow data to be shared and is a prerequisite to facilitate the wider development of ITS services. The MMTIS Regulation provides that transport authorities[164], transport operators[165], infrastructure managers and transport on-demand service providers[166] should make the (static and historic) travel and traffic data, corresponding metadata and information on the quality of the data, including data updates, accessible to users through a NAP.[167]

The type of data to be made accessible through the NAPs are detailed in Annex I of the Delegated Regulation.

[162] European Parliamentary Research Service and T. Madiega, "Briefing - EU Legislation in Progress - Digital Markets Act." May 2021. Accessed: Oct. 29, 2021. [Online]. Available: https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/690589/EPRS_BRI(2021)690589_EN.pdf

[163] 'Part II: Data as an Essential Facility', in Inge Graef, EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility, International Competition Law Series, Volume 68, pp. 123-124.

[164] Defined as 'any public authority responsible for the traffic management or the planning, control or management of a given transport network or modes of transport, or both, falling within its territorial competence'. MMTIS Delegated Regulation, Article 2(9).

[165] Defined as 'any public or private entity that is responsible for the maintenance and management of the transport service'. MMTIS Delegated Regulation, Article 2(10).

[166] Defined as 'any public or private provider of transport on demand service to users and end-users, including travel and traffic information thereof;'. MMTIS Delegated Regulation, Article 2(18).

[167] MMTIS Delegated Regulation, recital 10, Article 3.

Funded by the European Union

Static travel and traffic data are deemed essential information for planning purposes during the pre-trip phase, hence the sharing obligation. However, for dynamic travel and traffic data, Member States are only encouraged to include these types of data (listed in the Annex) through the NAP.[168]

The use of static and dynamic data for travel information services involves data from different actors across the value chain. In many cases, the original data from transport authorities, transport operators, infrastructure managers or transport on-demand service providers will be used by a travel information service provider. In this instance, the original source, the date and time of the last static update are indicated when used.[169]

The MMTIS Regulation is currently subject to revision, with an updated proposal shared for feedback in May 2023, the consultation closed on 28 June 2023.[170] The basic principles of the regulation, such as the fact that data need to be provided through NAPs remain the same. However, some changes should be highlighted: there is a new data category of historic and observed data, and an obligation to share dynamic data and new data sets. The explanatory Memorandum of the Act states that "to support the development of EU-wide multimodal travel information services, mandating the accessibility of dynamic data sets is essential for all modes including transport on demand. To allow for more accurate and accessible multimodal travel information services, additional static, historic, observed and dynamic data types are required. This includes data on parking, accessibility for persons with disabilities and persons with reduced mobility, and capacity for bicycles on-board scheduled transport".

The draft proposal continues to include some fare data as part of the categories of static, historic and observed travel and traffic data (e.g. basic common standard fares, common fare products, special fare products)[171], as is the case currently with the MMTIS. The fare structure describes the basis and scope (origin-destination pairs, zones, etc.) and access rights (single, multiple travel, class of use, etc.). The fare products assemble these as permitted combinations with specific usage and commercial conditions attached and assign a monetary cost. Fare distribution channels and payment methods may also be described in data describing the tariff structures of a network.

However, some aspects, such as prices (and seat availability) may be dynamic data[172] and are not included at the moment within scope. Sharing fare data (static and dynamic) is essential for the success of EU-wide integrated ticketing, which is essential to achieve MaaS. Static fare data are defined as fare data which do not change, or do not change often. Dynamic fare data refers to how the fare (or ticket price) is calculated in real time, depending on a set of variables that determine the price of a ticket in real time. According to the analysis carried out, while the static fare data is often open to the public and other transport operators, real time fare data is not.[173]

[168] *Ibid*, recital 12.
[169] *Ibid*, recital 21.
[170] European Commission, *EU-wide multimodal travel – new specifications for information services*, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12912-EU-wide-multimodal-travel-new-specifications-for-information-services_en
[171] European Commission, *EU-wide multimodal travel – new specifications for information services*, Annex https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12912-EU-wide-multimodal-travel-new-specifications-for-information-services_en
[172] European Commission, 'Remaining challenges for EU-wide integrated ticketing and payment systems', July 2019, https://op.europa.eu/en/publication-detail/-/publication/af05b3eb-df43-11e9-9c4e-01aa75ed71a1.
[173] *Ibid.*

Funded by the European Union

Interestingly, in a presentation in March 2023, a UITP representative refers to a new Article on “data enabling the payment and booking for transport services”[174], but this does not seem to have been included in the draft EC text.

3.5. National examples and lessons learnt

The French mobility orientation law (La loi d’orientation des mobilités – “LOM”)[175]

France also adopted a new legal framework to enhance the use of new mobility means. The LOM aims, amongst others, to accelerate the opening up of data and development of digital services. The LOM affirms, steps up, and facilitates the move towards open data. All the organising authorities (“AOs”) (‘métropoles’, community municipalities or intermunicipalities) are required to open up data on all existing modes.

The Regions are charged with coordinating the opening up of data and the transport regulatory authority (ART) with monitoring and settling disputes. Data must be disseminated either statically (a file) or dynamically via an API on the NAP, transport.data.gouv.fr. This NAP collects the data and shares it with re-users.

Moreover, with the LOM, *AO*s must now ensure MaaS exists to facilitate intermodality across all transport modes. They are also required to provide a digital interface accessible to parties offering MaaS services. Also new in the *LOM* is the definition of two legal categories of digital services for information, reservations and selling mobility services:
– the ‘contact platform’, which simply allows parties to deliver their own fare products;
– the ‘distributor’, which can, if the AO agrees, set the price a) for selling its own fare products and b) for reselling those of the transport operator. This means the ‘distributor’ could even create a different price.

SNCF announced the operation of SNCF Connect as of January 2022, an application which will not only be a real-time travel information services provider, journey planner and ticketing platform but will also integrate intermodal door-to-door journey comparison including the companies’ own services (train) but also green (bicycle, kick-scooter) and shared (taxi, ride-sharing, carpooling) mobility offerings. CO2 emissions of the suggested journeys will also be included.[176]

Finnish legislation on MaaS[177]

Finland is the first country to use legislation in such a way as to mesh together all different transport modes from taxis and city trams to long-distance trains and bike shares so that users can get

[174] Anika Degen, *The bigger picture – Latest news from the EU regarding MMTIS and MDMS*, https://kollektivtrafikk.no/app/uploads/2023/03/2023-03-23_EU-data-and-ticketing-initiatives_UITP.pdf
[175] https://futuramobility.org/en/lets-talk-lom-french-mobility-orientation-law/.
[176] https://www.sncf.com/sites/default/files/press_release/CP_NR_SNCF_Connect_19112021.pdf.
[177] When the going gets easier, Harri Pursiainen, Permanent Secretary, Ministry of Transport and Communications of Finland, 2 March 2020, OECD library, https://www.oecd-ilibrary.org/docserver/34521141-en.pdf?expires=1637682530&id=id&accname=guest&checksum=B181F99FB3D2F7BF81AA6B2EF93A3852.

Funded by the European Union

around and transport goods from A to B as frictionlessly as possible. The Finnish Act on Transport Services regards the entire transport system as a single entity.

It requires all transport service providers to open up their essential data, such as information on routes, stops, timetables, prices, availability and accessibility in a machine-readable form via open interfaces. By sharing data, service providers can use their transportation fleet more effectively in moving goods and passengers.

The Act also requires transport service providers to have compatible systems and grant each other access to their ticket and payment system interfaces. The government has given service providers an incentive to do this by making interoperability a criterion for public procurement. Service providers can sell customers tickets for other transport modes—a train vendor can sell you a train as well as the bus ticket you need to get to your destination from the train station, for instance. This makes going from A to B as easy and user-friendly as possible.

In keeping with the data regulation in the Act, mobility service providers have opened up a large number of interfaces for exchanging essential data and the opening of sales interfaces has also started.

In the first stage of the legislation in Finland, passenger service providers were obliged to open their sales APIs for regular single tickets. In the second stage, the sales APIs were required to offer the ability to act on behalf of the user and to allow the mobility or integrated mobility service provider with a right of access to purchase ticket products on the user's behalf using the identification and user information of the service user's existing user account with the mobility service provider. Acting on behalf of a user is considered an enabling element and catalyst for MaaS offerings by private and commercial providers.[178]

Looking namely at the Finnish legislation which is considered to have pioneered the adoption of a new regulatory regime, two key aspects merit discussion and that are currently lacking in EU legislation and thus in other Member States:

1. **The obligation of all mobility service providers to open essential data in a digital machine-readable format.** Before opening up their ticketing APIs, transport operators had to renew their ticketing system from (physical) card/ticket-based ones to account-based ticketing. In cases where operators have lacked their own computerised systems, the State has provided the necessary tools and financing to digitise information.[179]

   As mentioned above, the ITS Delegated Regulation 1926/2017 on multimodal travel information services stipulates the opening up of data. The Regulation requires private and public transport operators to make travel and traffic data accessible for re-use through the NAP. However, access to fare data, as granted by the Finnish laws is not envisioned by the Delegated Regulation, nor by the proposed revision.

[178] Inga Margrete Ydersbond, Heidi Auvinen, Anu Tuominen, Nils Fearnley, Jørgen Aarhaug, 'Nordic Experiences with Smart Mobility: Emerging Services and Regulatory Frameworks', Transportation Research Procedia,Volume 49, 2020, Pages 130-144, ISSN 2352-1465, https://doi.org/10.1016/j.trpro.2020.09.012.
[179] EUI Policy Brief, 'Towards EU-wide Intermodal Ticketing', September 2022, p.8.

Funded by the European Union

2. **The obligation of transport operators to open their ticket and payment system interfaces for third-party service providers**. The Finnish legislation provides that there must be a well-justified reason for refusing to negotiate or enter into an agreement dealing with opening ticket interfaces.[180] The law does not, however, regulate the third-party sale commission or revenue sharing, which in turn, is left to the transport operators and companies involved to negotiate.[181] The French law also provides that MaaS platforms are entitled to access the mobility service providers' selling channels with no right for them to refuse, in this case on the condition that the mobility service provider has a digital sales channel.[182]

The 2022 UITP handbook refers to ticketing as a "MaaS enabler" and trust builder. Indeed, all involved partners in MaaS can have an actual insight into their information flows, can share information in (near) real-time according to agreements with their partners and can flexibly adapt tariffs and other parameters while making optimum use of back-office systems that, in an inter-operable network, can be enabled to communicate with other back-office systems with mobility partners.[183]

According to the handbook, accessing all mobility services via one access point is a key requirement for MaaS. Interconnecting ticketing systems means enabling access to all mobility services via one access point. However, ticketing systems cannot be opened as transport data (timetable, real-time information) because it implies many aspects that must be contracted: fees, relation with Transport Users, assistance, refund procedures, claims management, ticketing technology, controlling, fines management, etc.[184]

3.6. Conclusions and recommendations

A consensus found in the literature is the formulation and specification of a roadmap for the development of MaaS. However, the focus lies mostly on local authorities, not at national or EU level.[185] This is probably because transport is primarily organised at a local level by public authorities.[186] This need for a more urban focus was also reiterated in the 2023 EC report of the Multimodal Passenger Mobility Forum.[187] Indeed, the daily mobility of citizens takes place in the urban area. Metropolitan and inter-metropolitan ecosystems are different, both concerning market size, the level of multimodal integration and the nature of the trips undertaken.[188]

[180] European Commission, Multimodal Passenger Mobility Forum, Report from the Expert Group, February 2023: https://transport.ec.europa.eu/news-events/news/multimodal-passenger-mobility-forum-final-report-2023-02-02_en.
[181] EUI Policy Brief, 'Towards EU-wide Intermodal Ticketing', September 2022, p.8.
[182] European Commission, Multimodal Passenger Mobility Forum Report (n 180), para 3.3.
[183] UITP Handbook (n 130), p.11.
[184] UITP Handbook (n 130), p.14.
[185] *Maas B* (n 107).
[186] *Maas B* (n 107).
[187] European Commission, Multimodal Passenger Mobility Forum Report (n 180).
[188] European Commission, Multimodal Passenger Mobility Forum Report (n 180), sections 2.1 and 4.1.

Funded by the European Union

According to the report, local authorities are responsible for public space management and infrastructure development according to policy goals, alongside public transport authorities that are responsible for contracting public transport services.[189]

It should also be noted that there are many differences in the transport systems of the different EU countries. Thus, it may not be possible to formulate one definition that will capture all possible implementations. It has been argued that determining a MaaS definition may be best left to local authorities depending on what MaaS means to a given city based on the preferences and aspirations of the city's inhabitants.[190] At the same time, multiple MaaS configurations exist and it is too early to tell what final configurations each market may have. The most appropriate model will only be understood as MaaS systems mature. Nonetheless, business models will be influenced by how the public authority regulates the MaaS market and the extent to which it plans to be involved in it.[191]

Based on the above analysis, we provide the following recommendations:

1. MaaS definition

Looking at EU-level regulation, it should be examined and clarified whether the EC's upcoming Regulation on "Multimodal Digital Mobility Services" seeks to regulate MaaS as described above, or additional platforms that do not qualify as MaaS providers. If the former, the aim should be to avoid inconsistency between the terms used to avoid confusion on the applicable regime. This practice was evident in the case of platform regulation, where, due to the many different legislation enacted, a certain platform could fall under many different definitions at the same time, for example, "intermediary", "video sharing platform" or "online intermediation service".

Regulators should also consider the following:

- how the service is defined and the characteristics entailed thereof can have consequences on the legal status of the MaaS provider and its contractual accountability towards users. Qualifying this service correctly from a legal perspective is essential since it determines which regulation applies and therefore which obligations, rights and procedures must be respected.
- Different modes of transport are regulated differently and any horizontal legislation should align with mode-specific legislation.

2. Definition of MaaS roles and responsibilities

The analysis has demonstrated the variety of MaaS business models and the large number of actors involved in the design of a MaaS system amongst which public transport companies, transport authorities, private on-demand players, data providers, technical service and IT providers, ICT infrastructure providers and customers. Coordination and alignment among actors is crucial for the success of MaaS implementation and this is one of the biggest challenges.

[189] European Commission, Multimodal Passenger Mobility Forum Report (n 180), section 2.1.

[190] Brown, C.; Hardman, M.; Davies, N.; Armitage, R. Mobility as a Service: Defining a Transport Utopia. Future Transp. 2022, 2, 300–309. https://doi.org/10.3390/futuretransp2010016

[191] ITF report (n 132), p.73.

Funded by the European Union

This suggests that it is not possible to determine *a priori* roles and responsibilities but to decide those based on the guidelines of the given legislation (e.g., GDPR) and on a case-by-case basis.

3. Access to data under competition law

In the absence of specific case law, the application of the essential facilities doctrine for MaaS operators to access data requires further refinement and case-by-case analysis, e.g. how a market will be defined (i.e. whether “MaaS” constitutes a separate downstream market or a complementary market). Further research could also evaluate the potential benefit of invoking the GDPR portability provisions.

4. Access to fare data and integrating ticketing systems

The above analysis demonstrated that a full MaaS offering cannot be realised without integrated ticket services and access to the relevant data. However, there are inherent risks in sharing potentially competitive sensitive information such as fares and those need to be examined. As there is no proposal for legislation that envisages similar provisions at the EU level, efforts should focus on implementing obligations under current legislation (i.e. sharing of data through NAPs under the ITS Directive).

Funded by the European Union

4. Conclusion and summary of Recommendations

Data sharing relies on the existence and availability of data. Data are heterogeneous, even in the mobility sector. For example, we have referred in our analysis to (real-time and historic) traffic data, passenger data, geolocation data or machine-generated data (in the context of connected cars). This heterogeneity can also be reflected in the legal environment, as there is not one legislation governing data. Conversely, a complex and fragmented legal landscape regulates data, and by extension, data sharing. This is mainly caused by the lack of, firstly, a uniform legal definition of data and secondly, a commonly accepted legal status of data. Indeed, a significant part of the data do not have a default legal status as intangible assets.

Despite the Commission's push for data-specific legislation, the reality remains that several regulatory gaps exist today, imposing barriers to data sharing. Our analysis in the first version of this deliverable has shown that such gaps span across different legal regimes and are particularly pertinent when we look at the interface of two different pieces of legislation, for example, the GDPR and the ITS Directive, or the Open Data and the ITS Directives. The picture becomes even more complicated if we consider that during the time of drafting of both of these reports, a wave of legislative interventions has taken place (e.g. the DMA, the Data Act, the updated ITS Directive), while others are still foreseen (e.g. initiative on Multimodal Digital Mobility Services, update of the MMTIS Delegated Regulation under the ITS Directive). This means that the current legal landscape on data sharing is constantly changing, which in turn makes guidance for legal aspects arising in the MobiDataLab project equally challenging.

In the present deliverable, we have focused our analysis on gaps identified in two legal regimes – the GDPR and Competition law – that are deemed the most complex and where (legislative or other) solutions could be offered to address those gaps. For GDPR, the analysis demonstrates the necessity of using personal data in mobility scenarios and the difficulties of anonymising the data to potentially escape the GDPR provisions. Similarly, consent withdrawal as a legal basis for processing personal data could jeopardise data transactions, necessitating potential reliance on other legal bases. Mobility ecosystems are complex involving several actors, constituting the identification of roles and responsibilities under the GDPR often too complicated. For Competition law, there is still little guidance on how to approach the issue of market definition for data markets.

Section 3 dealt with MaaS, a prominent use case both for MobiDataLab, but also for mobility stakeholders, analysing four identified gaps that could be resolved to make MaaS a reality. The analysis portrayed several core elements of a MaaS service and how those elements could affect its legal categorisation under the current EU law jurisprudence. It demonstrated four different MaaS business facets that could impact the determination of responsibilities amongst actors. Looking through the lens of Competition law to enhance access to data that could be critical for MaaS services, it demonstrated that the current rules do not suffice to enable data access. Data portability provisions do not appear a sound alternative either.

Funded by the European Union

Finally, it examined the ITS Directive and the two most prominent examples of national legislation – in France and Finland, to understand what is still missing in other EU Member States to advance MaaS. Open points relate to sharing ticketing and fare data as well as integrating interfaces.

Below is a compilation of the recommendations provided in Sections 2 and 3.

4.1. GDPR

1. **Consider the available toolbox under the GDPR to ensure lawful processing of personal mobility data**

The GDPR already offers the necessary toolbox to ensure lawful processing of personal (mobility) data, for example, privacy by design, abidance with the principles of data processing (e.g. purpose limitation, data minimisation) and the application of privacy-enhancing techniques. In this respect, consider the analysis carried out under D2.3. Special attention should be given to data that are considered particularly sensitive, such as geolocation data.

2. **Perform a risk-benefit analysis of using anonymous vs personal data and explore the use of synthetic data**

As mentioned above, it is not possible to draw a definite line between anonymous and personal data. A comparison between the potential risks from the use of personal data and the utility of anonymous data must be determined on a case-by-case basis, taking also into consideration the risk of re-identification of anonymous data based on the means available and likely to be used at a given time.

In addition, an alternative to data anonymisation and privacy-preserving techniques has gained ground in the past years: generating synthetic data.[192] Synthetic data are data that are artificially created, typically using mathematical models or algorithms, to resemble real-world data. They share similar statistical properties, patterns, and characteristics with real data but do not (in theory) contain any information about specific individuals or entities.[193] They can be used for various applications, such as developing and testing analytics processes, training and testing machine learning models, and performing simulations, while ensuring data privacy.

The use of synthetic data is motivated by several constraints. First is missing data and data availability in general. Data processing, and in particular machine learning, demands high-quality data in large quantities. Often, synthetic data are used as a tool to complete missing values or to *augment* the available data.[194] When completing missing values, one can use the mean of the sample for particular attributes, regression based on available values, or, in the case of mobility data, interpolation between GPS points.

[192] Farrell, E.; Minghini, M.;Kotsev, A.; Soler-Garrido, J.; Tapsall, B.; Micheli, M.; Posada, M.; Signorelli, S.; Tartaro, A.; Bernal, J.; Vespe, M.; Di Leo, M.; Carballa-Smichowski, B.; Smith, R.; Schade, S.; Pogorzelska K.; Gabrielli, L.; De Marchi, D., *European Data Spaces: Scientific insights into data sharing and utilisation at scale*, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/400188, JRC129900.

[193] Rubin, D.B., 1993. Statistical disclosure limitation. Journal of official Statistics, 9(2), pp.461-468.

[194] Reiter, J.P., 2004. Simultaneous use of multiple imputation for missing data and disclosure limitation. Survey Methodology, 30(2), pp.235-242. Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J. and Greenspan, H., 2018, April. Synthetic data augmentation using GAN for improved liver lesion classification. In 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) (pp. 289-293). IEEE.

Funded by the European Union

For data *augmentation*, i.e., increasing the amount of available data, transformation of existing data (for example, scaling and rotations in image data) and generation of new data (for example, with generative ML models) techniques are often used.

For years, synthetic data has been used to protect the privacy of respondents[195] and, in recent years, improvements in generation mechanisms based on machine learning and growing privacy concerns and related regulations have increased the interest of synthetic data generation. The idea is that randomly generated data, even if their structure or statistical measures are preserved, do not belong to any individual. Thus, re-identification or any other privacy violations are avoided.

This protection is rather nuanced since synthetic data, especially high-quality synthetic data, may unintentionally contain data that matches that of some real individual. It has also been argued that even synthetic data imply a certain risk of re-identification, which can, to a large extent, be controlled by design.[196] From a legal perspective, there is no consensus on the legal nature of synthetic data. They can be considered personal or anonymous data based on whether the identifiability threshold set by the GDPR is met. The considerations of “zero-risk” vs “acceptable risk” also come into play in this context. Supporters of synthetic data as anonymous data rather support an “acceptable-risk” approach, whereas those that oppose them rather identify themselves with the “zero-risk” approach.[197]

3. Determine GDPR responsibilities based on each scenario

It is evident from the above analysis that there are many determinants in identifying the exact role and relationships in a data sharing mobility scenario. As such scenarios vary depending on the envisaged aim, it is impossible to *a priori* set a roadmap that will determine who is responsible for GDPR compliance. Nevertheless, companies should use the current CJEU case law as guidance to determine whether they fit the tests set by the EU Courts.

4. Ensure accountability: documenting the entire data lifecycle

In line with the accountability principle underlying data protection law, a data controller will need to be able to justify and provide documentation on the processing of personal data and compliance with data protection legislation.[198] This also includes the very process of anonymisation itself, as this needs to be considered a processing of personal data. It is more importantly a processing that aims to create new data that is no longer personal, thereby placing this data outside of the scope of data protection law. Consequently, the data controllers should provide clear documentation, containing a full assessment of the anonymisation process, that supports the claim that the result of the anonymisation process indeed creates data that is no longer considered to be personal data under data protection law.

[195] Rubin, D.B., 1993. Statistical disclosure limitation. Journal of official Statistics, 9(2), pp.461-468.

[196] Farrell, E.; Minghini, M.;Kotsev, A.; Soler-Garrido, J.; Tapsall, B.; Micheli, M.; Posada, M.; Signorelli, S.; Tartaro, A.; Bernal, J.; Vespe, M.; Di Leo, M.; Carballa-Smichowski, B.; Smith, R.; Schade, S.; Pogorzelska K.; Gabrielli, L.; De Marchi, D., *European Data Spaces: Scientific insights into data sharing and utilisation at scale*, Publications Office of the European Union, Luxembourg, 2023, doi:10.2760/400188, JRC129900.

[197] On the legal nature of synthetic data’, César Augusto Fontanillo López, Abdullah Elbi, p.7 and the sources mentioned therein. https://openreview.net/pdf?id=M0KMbGL2yr.

[198] GDPR, Article 5(2).

Funded by the European Union

4.2. Competition law

As the debate is not mature yet, the recommendation is to continue monitoring for European Commission competition law cases that may develop such market analysis and anything relevant analysed in the literature.

4.3. MaaS

1. MaaS definition

Looking at EU-level regulation, it should be examined and clarified whether the EC's upcoming Regulation on "Multimodal Digital Mobility Services" seeks to regulate MaaS as described above, or additional platforms that do not qualify as MaaS providers. If the former, the aim should be to avoid inconsistency between the terms used to avoid confusion on the applicable regime. This practice was evident in the case of platform regulation, where, due to the many different legislation enacted, a certain platform could fall under many different definitions at the same time, for example, "intermediary", "video sharing platform" or "online intermediation service".

Regulators should also consider the following:

- How the service is defined and the characteristics entailed thereof can have consequences on the legal status of the MaaS provider and its contractual accountability towards users. Qualifying this service correctly from a legal perspective is essential since it determines which regulation applies and therefore which obligations, rights and procedures must be respected.
- Different modes of transport are regulated differently and any horizontal legislation should align with mode-specific legislation.

2. Definition of MaaS roles and responsibilities

The analysis has demonstrated the variety of MaaS business models and the large number of actors involved in the design of a MaaS system amongst which public transport companies, transport authorities, private on-demand players, data providers, technical service and IT providers, ICT infrastructure providers and customers. Coordination and alignment among actors is crucial for the success of MaaS implementation and this is one of the biggest challenges. This suggests that it is not possible to determine *a priori* roles and responsibilities but to decide those based on the guidelines of the given legislation (e.g., GDPR) and on a case-by-case basis.

Funded by the European Union

3. Access to data under competition law

In the absence of specific case law, the application of the essential facilities doctrine for MaaS operators to access data requires further refinement and case-by-case analysis, e.g. how a market will be defined (i.e. whether "MaaS" constitutes a separate downstream market or a complementary market). Further research could also evaluate the potential benefit of invoking the GDPR portability provisions.

4. Access to fare data and integrating ticketing systems

The above analysis demonstrated that a full MaaS offering cannot be realised without integrated ticket services and access to the relevant data. However, there are inherent risks in sharing potentially competitive sensitive information such as fares and those need to be examined. As there is no proposal for legislation that envisages similar provisions at the EU level, efforts should focus on implementing obligations under current legislation (i.e. sharing of data through NAPs under the ITS Directive).

Funded by the European Union

MobiDataLab consortium

The consortium of MobiDataLab consists of 10 partners with multidisciplinary and complementary competencies. This includes leading universities, networks and industry sector specialists.

@MobiDataLab
#MobiDataLab

https://www.linkedin.com/company/mobidatalab

For further information please visit **www.mobidatalab.eu**

MobiDataLab is co-funded by the EU under the H2020 Research and Innovation Programme (grant agreement No 101006879).

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 MobiDataLab consortium members shall have no liability for damages of any kind that may result from the use of these materials.