Interreg
2 Seas Mers Zeeën
MOBI-MIX
European Regional Development Fund

MOBI-MIX insight report

Shared mobility data for policy making

BAX & COMPANY
VALUE FROM SCIENCE AND TECHNOLOGY

POLIS
CITIES AND REGIONS FOR TRANSPORT INNOVATION

Acknowledgements

Written by Joan Estrada, Samuel Rouquette (Bax & Company) and Laura Babío (POLIS)

Reviewers and contributors: Antonia Roberts (CoMo UK), Wassim Hached (UGE)

Designed by Katie Basham (Bax & Company)

This insight report relies on secondary information obtained through desk research. This was complemented by primary information collected by the Bax & Company team through interviews. The following people have kindly shared their time and knowledge with us:

- Mélanie Gidel (city of Paris)
- Lars Petter Klem & Bård Dagestad Carlsen (city of Bergen)
- Alex Pazuchanics (Vianova)
- Stijn Vernaillen & Jelle de Keyser (city of Antwerp)
- Sabrina Schimmel (city of Berlin)

1. Setting the context

Mobility has experienced many disruptions in the last years that will continue growing over the following decades. One of the main disruptions is shared mobility. Currently, there's quite a broad understanding and agreement that shared mobility is already here to stay. The question just lies in the how.

At the same time, shared mobility has become a rich source of data and recently cities have realised of the untapped potential it holds to analyse and evaluate transportation and mobility goals. In this document, we want to highlight the value this data holds, how to obtain it, and what use cases can cities make out of it, following some common approaches and best practices around EU. In the end, we also showcase some challenges brought by this data as well as future recommendations to ensure that the data ecosystem lives up to its promises.

What is (shared) mobility data?

The term "data" has experienced an increased growth in interest over the last years, and data is being considered by some as the new oil. New data streams are being made available by a lot of different sources. However, to set the scope of the document, we're going to focus on data stemming from mobility related sources. Let us describe first what is meant by the term "mobility data".

Some simplistic definitions define mobility data as "data describing people's movements from one location to another and the mode of transport" (The ODI, 2020). However, with the new mobility revolution (Deloitte, 2016), the definition can be significantly expanded, as new data streams are able to provide more information than just origin/destination matrixes per mode.

Other authors define mobility data as information generated by activity, events or transactions using digitally-enabled mobility devices or services. This data is frequently recorded as a series of points with geographic coordinates collected at regular time intervals by connected devices such as smartphones, shared micromobility vehicles, on-board vehicle computers, or app-based navigation systems (e.g. Waze, GoogleMaps, etc.). Mobility data often provides a temporal element, assigning time as well as location to each point. Depending on the device used to capture the data, other characteristics, such as the speed of travel, or who is making the trip, can be connected to each individual latitude/longitude point (NACTO, 2019).

As the definition already suggests, mobility data can be collected from several sources, ranging from the more classical ones to the modern ones. Classical methods include agents standing on corners and keeping tallies or the usual travel surveys that cities or entities do on their citizens.

Among the modern ones, most of them are enabled by the digital revolution. They include GPS tracks (from shared mobility records, mapping apps, phone operators, etc.), origin-destination matrixes coming from journey planner services or smart ticketing data, tagging locations on social media, traffic cameras, etc.

Since MOBI-MIX is a project focused on shared mobility (and related services such as MaaS and mobility hubs), we'll put a special focus on the data generated by these services and how can cities use it to achieve their policy goals and objectives.

Data sharing vs. data reporting

Making data accessible to external parties, either public or private, can be a great opportunity for all within the mobility ecosystem to make it dynamic, accessible and responsive. However, it also poses some challenges for providers regarding privacy and proprietary interests. While some cities and countries are more advanced in settling agreements with private operators, others are still confused about which data to share or to ask for from partners, or how far they should push them to share data, and in which format. This confusion is mainly often due to privacy issues, lack of capacity, or missing a clear purpose, among others. As stated by D'agostino et al. (2019), oversharing and undersharing are both problematic.Overly demanding data-sharing requirements can deter companies from investing in a city. In addition, the associated costs with meeting these requirements could be passed onto consumers. The middle-ground, allowing both maximizing the benefits while minimizing the risks, is still under research.

In this report, we will try to shed more light on the purpose behind data sharing together with best practices and lessons learnt. An opportunity-cost approach would suggest that we ask ourselves: what would happen if shared mobility operators did not share data? To answer this question, we need to first distinguish between data sharing and data reporting, following the definitions proposed by ITF (2021).

According to the study, "mobility data sharing refers to data shared among market actors and other stakeholders that enables the delivery of mobility and other services and supports the functioning of transport markets". In other terms, it takes into account the data exchanges that need to occur, so the transport system can work properly and efficiently. Some consider this as the "public-facing data", which is intended to facilitate trip planning and booking (Populus, 2021).

On the other hand, "data reporting refers to data provided by stakeholders and market actors to public authorities that enables the latter to monitor, guide and intervene to enact public policy". It is also defined by Hausigke and Kruse (2021) as a transport planning instrument working towards more people-oriented mobility planning, that is to be seen as a supplement to conventional transport and infrastructure planning. In line with what we mentioned in the previous paragraph, Populus (2021) considers reported data as "non-public-facing data", implemented by public authorities to manage the use of public space, inform planning and infrastructure investments, and to provide incentives and nudges to riders based on the characteristics and impacts of their trips.

Among the modern ones, most of them are enabled by the digital revolution. They include GPS tracks (from shared mobility records, mapping apps, phone operators, etc.), origin-destination matrixes coming from journey planner services or smart ticketing data, tagging locations on social media, traffic cameras, etc.

Since MOBI-MIX is a project focused on shared mobility (and related services such as MaaS and mobility hubs), we'll put a special focus on the data generated by these services and how can cities use it to achieve their policy goals and objectives.

Data sharing vs. data reporting

Making data accessible to external parties, either public or private, can be a great opportunity for all within the mobility ecosystem to make it dynamic, accessible and responsive. However, it also poses some challenges for providers regarding privacy and proprietary interests. While some cities and countries are more advanced in settling agreements with private operators, others are still confused about which data to share or to ask for from partners, or how far they should push them to share data, and in which format. This confusion is mainly often due to privacy issues, lack of capacity, or missing a clear purpose, among others. As stated by D'agostino et al. (2019), oversharing and undersharing are both problematic.Overly demanding data-sharing requirements can deter companies from investing in a city. In addition, the associated costs with meeting these requirements could be passed onto consumers. The middle-ground, allowing both maximizing the benefits while minimizing the risks, is still under research.

In this report, we will try to shed more light on the purpose behind data sharing together with best practices and lessons learnt. An opportunity-cost approach would suggest that we ask ourselves: what would happen if shared mobility operators did not share data? To answer this question, we need to first distinguish between data sharing and data reporting, following the definitions proposed by ITF (2021).

According to the study, "mobility data sharing refers to data shared among market actors and other stakeholders that enables the delivery of mobility and other services and supports the functioning of transport markets". In other terms, it takes into account the data exchanges that need to occur, so the transport system can work properly and efficiently. Some consider this as the "public-facing data", which is intended to facilitate trip planning and booking (Populus, 2021).

On the other hand, "data reporting refers to data provided by stakeholders and market actors to public authorities that enables the latter to monitor, guide and intervene to enact public policy". It is also defined by Hausigke and Kruse (2021) as a transport planning instrument working towards more people-oriented mobility planning, that is to be seen as a supplement to conventional transport and infrastructure planning. In line with what we mentioned in the previous paragraph, Populus (2021) considers reported data as "non-public-facing data", implemented by public authorities to manage the use of public space, inform planning and infrastructure investments, and to provide incentives and nudges to riders based on the characteristics and impacts of their trips.

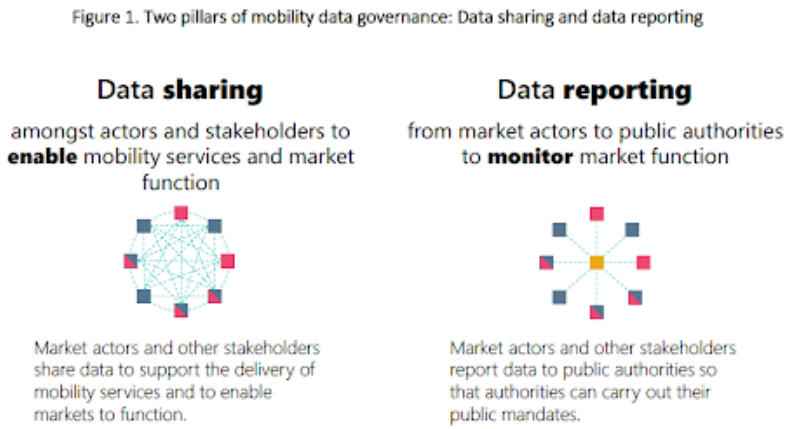

Figure 1: Distinction between data sharing and data reporting (Source: ITF, 2021)

Now, let us move back to our question of what would happen if (shared) mobility data would not be shared or reported. In the case of data sharing, the answer is pretty obvious as the transport system would simply not work: vehicle availability wouldn't be visible to users, stations' location couldn't be placed on a map, vehicle's battery status would be unknown, interoperability of transport modes would not be possible, services could not be bundled in a MaaS app, integrated payment options could not be implemented, etc. In other words, data sharing is an opportunity for providers and users as it increases the visibility of the vehicles, their intermodality, and allows for more selling channels which in turn could result in an expanded user base. However, when it comes to data reporting, the consequences of omitting it are not so straightforward. A key question would be: what is the data used for? The lack of common practices, the variety in public-private agreements, the availability of different data formats and many other factors make the answer to this question very extensive.

According to ITF (2021), there are three broad use cases for such data: data for planning, data for the operational management of traffic and public spaces, and data for enforcement actions. In summary, data reporting is mainly useful for cities, to ensure that shared mobility is complying with the rules and is aligned with higher policy goals.

In this report, we will focus on the first two use cases of data reporting (for planning and for operational management of traffic and public spaces), taking a deeper look into the main standards, datasets and KPIs available, stakeholders involved and their roles, together with the impact of policies and collaborations agreement. This information is used to answer the question of which data is used for which purpose, identifying best practices and main barriers.

2. Current status of data reporting

Data acquisition models

When it comes to shared mobility, in most cases, data are just not open and accessible to everybody. On the contrary, they are held by private firms and must be acquired or accessed by the public authorities to carry out their governance functions.

Following the European Investment Bank's identification of data acquisition methods for public authorities together with other reviews, ITF defines four data acquisition models for public authorities to gather data that is not generated by themselves: compulsion, conditionality, co-operation, and commercial terms.

Compulsion

Simply requiring or compelling an individual or a firm to share data with a public authority. This is the strongest and most constraining method, and it is generally deployed when the social benefits of obtaining such data outweigh individual privacy or commercial sensitivities related to that data remaining out of government hands. Public authorities must still ensure that personal or sensitive data are not released or used in ways not aligned with public policy objectives, including protecting individual privacy and commercial competitiveness.

One example of this model can be found in Paris, where since 2019 the city adopted regulations regarding the operation of free-floating vehicles in the public space. These regulations also include the provision of operational data (mainly vehicle ID, position, type and activity). The data is fed into an internal dashboard that allows the city to monitor the fleet, analyse parking and verify if the annual fee has been paid (EIB, 2021).

Conditionality

Conditionality includes requiring data in order to access a service, a set of rights, or to be granted a licence to operate. It is closely related to compulsion as some services or licences are so ubiquitous or necessary to function that their acquisition is almost compulsory (e.g. parking permits).

An example within MOBI-MIX can be found in the city of Antwerp, which requires data from mobility providers to operate in the city. In this case, instead of imposing the framework, the city engaged in bilateral conversations with service providers, resulting in two agreements. One agreement concerns a generic data licence that determines what data must be provided and for which purposes. The second agreement, on service level, determines the minimum availability and quality of the data, what support is expected and how long the data must remain available (EIB, 2021).

Co-operation

Regarding co-operation, firms or individuals voluntarily transfer data to public authorities to obtain mutually beneficial outcomes. Firms may also want to share data with public authorities via voluntary agreements as a way of building trust, avoiding more constraining data-reporting requirements, or improving reputation, media attention or access to customers.

According to EIB (2021), these partnerships can be a good option to get data when there are no hard guarantees needed on data quality or availability and the data is not critical for the city. When the usefulness of data is not clear for certain outcomes, such low-cost data can be used for a proof of concept.

A good example is the Waze for Cities programme, where participating cities have access to the data collected by the Waze app users: travel times, road works, road closures and notifications from Waze users like accidents, potholes, missing signs and dangerous obstacles on the road. Cities have access to the data and to a dashboard that enables them to analyse it. Additionally, the city provides information on roadworks and road closures. This gives the app insight into future road conditions, information that is not always easy to get without the city's involvement (EIB, 2021).

Commercial terms

Governments may require data from data aggregators and processors by purchasing it. The data are aggregated, anonymised and processed to meet a particular public authority's (or commercial) need for which governments (or firms) are willing to pay. Examples include processed location-based data built on data from mobile network operators or routing engines. In this line, Paris awarded a contract in 2016 to a consulting company to collect raw traffic flow data, specifically floating vehicle data and positioning data generated by smartphones, from public and private sector partners for the French road network. The tender was specifically asking for traffic speed data on the main Parisian roads for a specific period (EIB, 2021).

However, as investigated by POLIS in their report about micromobility data, more than 75% of the cities surveyed indicated that they receive the digital datasets directly, so the intervention of a third party is currently an uncommon practice (Homem de Gouveia and Babio, 2021).

Comparison across models

Every data acquisition model has its pros and cons, and although some models seem to be more widely used, there doesn't seem to be a clear winner yet. Perhaps there will never be a single winner, as cities might use one tailored to the situation at the time, the costs, the hurdles, and the benefits. To shed a bit more light on which model to choose, in the table below we provide a qualitative comparison among the exposed acquisition models based on some parameters (from EIB and other sources):

- **Commonality:** how commonly used is this acquisition models nowadays.
- **Trust:** to what extent an elevated level of trust is needed and generated between the provider and the city and how relationships between both impact the model.

Cost: budget needed on the city's side to get the necessary data.
Control: indicating how much control does the city have over data specifications like quality, availability, formats, etc.
Flexibility: how much can the specifications, perspectives or outcomes of the data be changed throughout the process.

Organisational impact: it refers to the changes and workload that will be imposed to the mobility departments.
Technical capacity: indicating what is the technical expertise needed to have when using this model.
I**mplementation speed**: speed to achieve results.

Data acquisition model	Commonality	Trust	Cost	Control	Flexibility	Organisational impact	Technical capacity	Implementation speed
Compulsion	++	+	+	+++	+	++	+++	++
Conditionality	+++	++	++	+++	+	++	+++	++
Co-operation	+	+++	+	+	+	+	+++	++
Commercial terms	++	++	+++	+++	+	+	++	++

The relative cost of different data acquisition models (both for governments and concerning the burdens imposed on individuals or firms) are important, as are the proportionality and balance of these costs versus the expected benefits to be derived from the data collection. These benefits of each model are often still uncertain. Behind direct or conditional requirements to report data to public authorities is their need to use these data to deliver actionable and policy-relevant insights. As we can see, making data reporting compulsory enables the city to have higher control over the data shared with a relatively low cost. However, some providers might not have an aligned mobility vision with the city, and this imposition of data reporting might turn in low levels of trust, which could be worsened if the city doesn't make good use of the data it mandates to obtain.

For this reason, cities must have enough technical capacity and organisation to get the best out of the reported data, otherwise its credibility to mandate data reporting will be hindered, threatening the shared mobility landscape's stability.

Similar to the compulsion model, data reporting subject to certain conditions enables high control levels over data. The downside of this model comes with a bit higher cost for the city compared to the compulsory one, not necessarily a monetary cost, but the city will need to cede in certain aspects that are part of the conditions (i.e. parking permits, operating licences, financial benefits, etc.). The good part is that the level of trust enabled by this model is higher than the compulsory one, and it will increase if both the city and the shared mobility provider are getting valuable insights and are comfortable with the conditions, respectively.

Regarding trust, the co-operation model is the one that both requires stronger relationships between the public and private mobility stakeholders and, in the same turn, contributes more to strengthen such relationships. Also, this trust level and aligned interests have a positive impact on cost, with the data being more affordable by the city. On the other hand, this model comes with a lower control level over the data, although the better the relationship, the better the control. Like the conditional model, cities need to have sufficient technical capacity to get valuable insights out of the shared data. This will help to justify the established relationship and keep improving it. Finally, the model based on commercial terms, enables high control over the data because all the necessary conditions will be established on a contract basis. Obviously, this comes at a higher cost than the other models, normally monetary. The advantages of this model are that it might require a bit less technical capacity and has less organisational impact on the city's side, because data might be processed and prepared by the provider or a third party, depending on the contract's conditions. It is a kind of model that seems appropriate for cities that are starting to consider whether to acquire data and what to do with it. In summary, public authorities' technical, budgetary and human resource capacity to materially collect, process and manage data are essential elements to consider when selecting a data acquisition model. Public authorities should increase their technical and organisational capacity to process and manage data so that they may carry out these tasks directly or become more proficient at managing contracts with the third parties they delegate to carry out these functions.

For this reason, public authorities should carefully consider which data should be reported to them, irrespective of whether data processing is carried out by them or by a third party.

Mobility data standards

For years, data generated from shared mobility providers has been scattered and quite inconsistent. Data sharing between private and public stakeholders did not always appear as a matter of course, making the design of data-driven policies almost impossible. The lack of data standardization has been identified as "one of the most important issues surrounding mobility data sharing" (D'Agostino et al., 2019). In this line, a data standard homogenises "how data is created, defined, and organized in order to ensure consistency and interoperability" (Oregon DoT, 2020). Having standardised data allows third-party apps or cities access to information from hundreds of providers because their data follows a shared format. FreeNow showed in the POLIS conference in 2021 the usefulness of standards when it comes to implement different APIs for their multi-mobility offering. Having the same standard makes API implementations faster. As a matter of fact, they stated that the time required for the same API standards amounts only to a fourth of the time required for different APIs. From the development of standards to their adoption, policies have a major role to play to ensure confidentiality, access, and exploitability of the data generated by mobility users and citizens in general. The need for standardization stems from it, as it is essential to ensure interoperability of data.

Adopting again an opportunity-cost approach, the absence of standardization generally leads to inconsistent data collection and sharing practices, thus impeding public actors to build insights for monitoring, planning and regulatory purposes.

Public authorities and mobility operators have understood this urgent need and as a result, numbers of standards have been created, most of them emerging in the United States. These standards can be divided in two main categories: those that are public-facing and intended to facilitate trip planning and booking, and those that are non-public and regulatory, implemented by public agencies to manage the use of public space by shared mobility services. These latter standards also inform planning and infrastructure investments, in order to provide adequate incentives and nudges to users based on the characteristics and impacts of their trips. This section focuses on briefly describing the most relevant/common shared mobility data standards and their use.

GTFS

Among the public-facing standards, the earliest one settled was the General Transit Feed Specification (GTFS), created in 2005 when TriMet (Oregon's public transport operator) approached Google to deal with the problem of online transit trip planners through the use of open datasets. It provides static information on public transport (e.g. routes, stops, schedules, arrival and departure times, etc.). Its early success resulted in widespread adoption, with GTFS being now used by more than 1,000 transport operators in the world. Also, since then, it has been extended to include real-time data.

GBFS

Sharing the same objective of powering trip planning service, the General Bikeshare Feed Specification (GBFS) came to life in 2015, as an open-data standard during the rise of bike sharing. Initially designed for docked bikes, its application area was then extended to other shared-mobility services. It follows the same principle of standardization as GTFS, but also including real-time micromobility data. For all this, it is specifically designed to allow open and public access to data regarding shared mobility services. The specification consists of thirteen files that contain different types of mobility data. Some of these files and their associated fields are required to comply with the standard, while some others are optional. Municipalities may require some optional files or fields in their regulations. Comprehensive information about GBFS can be found in this article from Mobility Data.

MDS

Following the success of the first standards and the swift development of micro-mobilities, the MDS (Mobility Data Specification) is a regulatory standard that aims, through its bilateral functioning/ons, to facilitate collaborations between mobility providers and public authorities. Developed by the Los Angeles Department of Transportation, it is now managed by a not-for-profit foundation (OMF – Open Mobility Foundation) and widely adopted in the US. This standard presents all the functionalities of GBFS while bringing more details on the vehicle state and its historical trips. However, this standard has raised some concerns regarding users' privacy and GDPR compliance.

CDS-M

Eventually, these concerns brought a consortium of Dutch cities to develop a new standard, the CDS-M (City Data Standard for Mobility). It aimed to bring a robust compliance within European GDPR while offering a similar set of functionalities as MDS, spreading it afar from micromobility. It covers three main public authorities' requirements in terms of planning, policy and enforcement. Although originally it was conceived as a standard, it has evolved into a set of use cases and protocols to ensure secure data exchange. It focuses on specific questions or problems that a city might want to solve and why, what data is needed for this purpose, which data formats are best suited, and what legal requirements and privacy assessments need to be considered. Its compliance with GDPR regulations, and its replicability across different cities and providers could potentially allow its widespread adoption within the EU. However, as of today, only 5 Dutch cities have adopted it.

TOMP-API

In response to the urgent need of data standardization between MaaS providers and transport operators, an open-source working group constituted by both public and private stakeholders started developing the TOMP-API standard in 2019 (Transport Operators to MaaS providers – Application Programming Interface). Seeking the implementation of a standardized language for data exchange between these specific stakeholders to improve interoperability and monitoring, the TOMP-API now provides information on an entire MaaS journey. This includes operators' information, registration, planning, booking, trip execution, payment and support, which interactions can be found in the figure below.

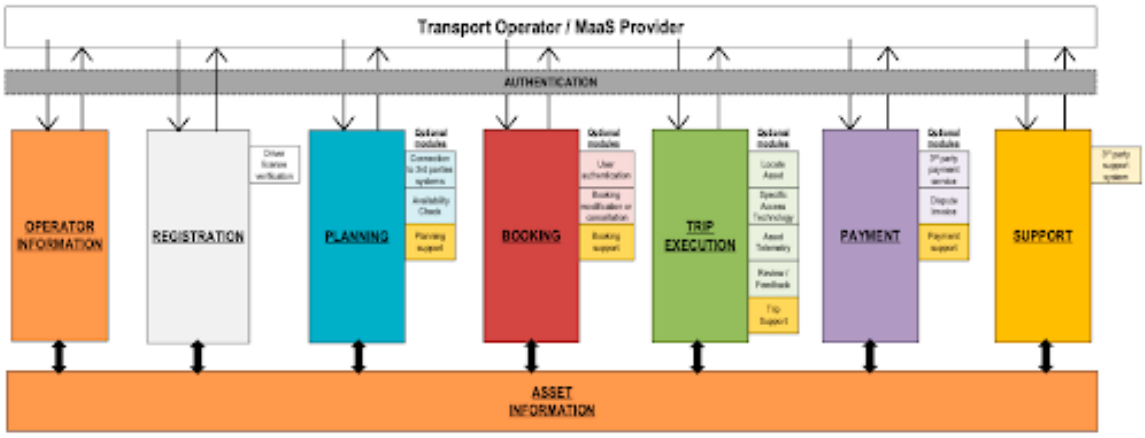

Fig. 1: MaaS Provider to/from Transport Operator (TO-MP) API – Functional Blocks visualization

Figure 2: Functional blocks of TOMP-API (Source: Dutch Ministry of Infrastructure and Water Management, 2019).

1 Facilitated by the initiative of the Dutch Ministry of Infrastructure and Water Management

2 Static data on its regions, stations, opening hours, pricing, terms and conditions

When considering the roadmap proposed by the TOMP-API initiative and the numerous implementations already in place and planned, it seems that the standard is constantly under development to ensure a maximal coverage of use cases, all while keeping an open-source functioning. The challenge now lies in making this standardized framework widely accepted and used.

To do so, there is first a need for clarity for cities to understand how useful data can be for them to leverage their policies and deal with regulatory, social, and environmental aspects that stems from the use of micromobility services (among others).
As we will illustrate below, data standards are crucial when building indicators that help cities monitor and evaluate the impact of such services within their boundaries. However, their fragmentation makes it complex to choose between standards, besides the fact that cities do not always possess the human resources to reflect, plan and decide on such matters- often due to higher labour costs needed to acquire such skills.

Moreover, even though European cities are undertaking strong efforts to reach a widespread standardization, shared mobility providers and public authorities have not agreed yet on one open standard for historical, aggregated, and anonymized mobility data. Eventually, this could limit the former's expansion across borders at both EU and international level. Indeed, for a mobility provider, adopting and conforming mobility data to different standards in a row may require high investments, which obviously imply a negative impact on competitiveness. This could discourage/demote some companies from expanding into new geographical areas, where they would find it too costly to comply with the existing standards of the targeted expansion area.

Also, consulting with operators on the data format will allow them to use their expertise to feed into the data reporting plan and suggest a format common with other cities. Using the same template will reduce costs for the operator and could allow for comparisons with other cities.

The case of Paris, France

"Mutual interest both from the city and private providers resulted in the development of the SIVU standard"

Interviewee	City of Paris
Motivation	Initially the city had sensors mounted to monitor mobility. These used to work for cars, but when micromobility appeared, they had nothing to monitor them. After an open innovation program and meetings with providers, there was a mutual interest to share information to solve parking issues. This is when the SIVU "standard" was born.
Data regulations	After the open innovation program finished, the city made it mandatory to share data by municipal by-laws, regulations, and tenders.
Data requirements	- Provide detailed data via 3 APIs - Monthly statistical reports
Standards collected	GBFS – The city also stores GBFS data, so evolution and trends can be observed. MDS – Not stored since it's more complicated than GBFS. Used for occasional ad-hoc analyses. The advantage for the city is that they can request the historical data, so it doesn't need to be stored. SIVU – Simple and pragmatic format designed in Paris both by the city and providers. Its simplicity allows the city to not have external data processing. New operators are not required to apply to it because they already abide by the other ones and it would require them additional specific developments.
Data acquisition model	Conditional – thanks to a national framework, cities in France can request providers to share data in order to operate in the city
Value creation – data use	For now, GBFS and SIVU data are used in their own dashboard developed in-house to monitor fleet sizes and parking rules. Not still able to explore the full potential of MDS because of a lack of time and resources.
Challenges	It is more complicated than originally thought to integrate MDS into the city dashboard. Due to capacity limitations and to avoid the costs of externalising its analysis and integration, the city still has little use of MDS. SIVU and GBFS cover most of the city's current use cases. Some operators are still struggling to share MDS data, because they are old companies and are slow adapting to the new standards. There are also difficulties to evaluate parking compliance because of vehicle's geoposition accuracy
Other interesting points	Shared mobility providers need to prove that their environmental footprint is low, and they are actually compared and scored on this topic when they are being selected. If the LCA of the vehicles is evaluated by an external company, the score is higher.

2.3 Regulating data exchange

At the EU level, there are several legal and regulatory frameworks that tackle data sharing and reuse in the transport sector. The EU project MobiDataLab has prepared a gap analysis of the legal and regulatory data sharing frameworks in the EU, including horizontal and sector-specific legislation, that provides a good overview of the data legal ecosystem in the EU.

Horizontal legislation

- Privacy and Data Protection:

The EU General Data Protection Regulation (GDPR), which has a broad scope and strict criteria, has swiftly become a global privacy standard for all sectors dealing with personal data. To put it very simply: whenever data sharing activities require the processing of personal data, GDPR will be applicable (e.g. for ticketing validation or enforcement). Personal data is very broadly defined as any type of information that relates to an identified or identifiable natural person, which includes not only private or sensitive data, but encompasses all types of information related to the person. Considering the multitude of actors active in a data-sharing ecosystem in the transport sector, it can be challenging to determine the obligations of each actor to comply with GDPR. The e-Privacy Directive aims to "particularise and complement" the provisions of the GDPR, concerning the processing of personal data in the electronic communication sector.

- Competition law:

Competition law is a set of regulations designed to safeguard the competitive process and maximize consumer welfare. Although it has wide applications in the sector, it is particularly interesting in the MaaS case, since organisations that hold significant market power in the mobility sector can be subject to stricter requirements under competition law. However, by barring those market leaders from limiting access to new players, competition law can also greatly help in the deployment of MaaS.

- The Public Sector Information (PSI) Directives:

The PSI Directive (and its revision in 2019) provide a regulatory framework for the re-use of public sector information, including the minimum rules for public authorities to make their data available for commercial or non-commercial purposes. The Directive introduced provisions on non-discrimination, charging for re-use, exclusive arrangements, transparency, licensing, and practical tools facilitating the re-use of public sector documents.
The Open Data Directive represents the latest major upgrade of the PSI legislation in the EU. It amends previous PSI directives establishing a stronger requirement on the principle of reuse of data by default and the principle that research data resulting from publicly funded research must be Open Access by default.

- The Regulation on the free flow of non-personal data:

The EU regulation on the Free flow of non-personal data is the first legislation dealing with non-personal data and is perceived to be a counterpart to GDPR. This regulation sets rules concerning data localisation requirements, the availability of data to competent authorities and the transfer or

'porting' of data for professional users. The broad definition of personal data under GDPR generates uncertainty as to what information falls within the scope of each regulation.

- Legislation concerning digital platforms and/or intermediaries:

The Commerce Directive 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.
The Platform to Business Regulation (often known as the "P2B Regulation") aims to rebalance the relationship between online platforms and the businesses that utilize them. Its goal is to promote fairness and transparency for business users of online intermediation services. The Regulation essentially contains clauses similar to those found in consumer protection laws (such as transparency, notification of Terms and Conditions changes, class action lawsuits, or other forms of recourse). In 2020, the EC proposed two new pieces of legislation: the Digital Services Act and the Digital Markets Act. Once adopted, the proposed regulations will affect the legal obligations and responsibilities of online users, customers, and corporate users in the EU. These initiatives have two main goals:

1. to create a safer digital space in which the fundamental rights of all users of digital services are protected;
2. to establish a level playing field to foster innovation, growth, and competitiveness, both in the European Single Market and globally.

- The proposal for a Data Governance Act (DGA):

By fostering trust in data sharing and data intermediaries, the DGA proposal seeks to create the conditions for accelerating the development of the common European data spaces. In this regard, the DGA proposal establishes a broad framework with measures that are applicable to all common European data spaces while allowing for the application of sector-specific regulations.

Through the creation of data intermediaries serving as neutral market facilitators, the Commission hopes to develop a model for data exchange. The obligation to maintain neutrality relates to the data shared and ensuring that all users have fair, transparent, and non-discriminatory access to their services.

Sector-specific regulation

The Intelligent Transport Systems (ITS) directive was published in 2010 with the aim to accelerate and coordinate the deployment and use of these systems. ITS are defined as "systems in which information and communication technologies are applied in the field of road transport, including infrastructure, vehicles, and users, and in traffic management and mobility management, as well as for interfaces with other modes of transport". In 2021, the EC adopted a revision of the ITS directive to tackle some shortcomings and future-proof the regulation in view of emerging solutions such as CCAM and MaaS. At this time, the revision is at the proposal stage but it includes far-reaching scope changes such as the mandatory creation of data by public authorities or the mandatory deployment of ITS services (currently these had only applied to safety-related information services).

Based on the ITS Directive, the Commission has introduced legally binding specifications for interoperability and continuity through delegated acts and the creation of certain necessary standards. These initiatives follow the six priority actions established in the mother directive:

1. the provision of EU-wide multimodal travel information services,
2. the provision of EU-wide real-time traffic information services,
3. data and procedure for the provision, where possible, of road safety related minimum universal traffic information free of charge for users,
4. the harmonised provision for an interoperable EU-wide eCall,
5. the provision of information services for safe and secure parking places for trucks and commercial vehicles,
6. the provision of reservation services for safe and secure parking places for trucks and commercial vehicles.

The most relevant initiatives for new (digital) mobility services are the RTTI, MMTIS and, in the future, a new iniative on MDMS. While MMTIS concerns the provision of EU-wide multimodal travel information services, RTTI covers the provision of real-time traffic information. The objective is for MDMS to cover data-sharing between public authorities, transport operators and digital service providers to enable multimodality at EU level – the initiative is currently under development.

According to the delegated regulations of MMTIS and RTTI, every Member State must establish a National Access Point (NAP) where the data required by the regulations can be made available. The goal of the NAP is to make it simpler for data users to discover the variety of data available. However, the regulations do not establish which form should the NAP take and different approaches are appearing throughout the EU at different paces. To work on a better alignment of the NAPs, the European Commission is co-financing the NAPCORE project to work as a coordination mechanism. NAPCORE ambitions go far beyond the NAPs as it aims to improve the interoperability of mobility data in Europe through standard harmonization and alignment, and by enabling the ITS directive and delegated regulations.

Regulation on data does not happen exclusively at EU level, since several regulations have been approved at national level and more are expected to appear in coming years. France has recently passed two relevant pieces of legislation that will set precedent for the future: the Loi d'orientation des Mobilités (LOM) and the Climate Law. The LOM broadens the scope of data for transport and mobility services, which must be open freely and free of charge – this piece aims at accelerating the deployment of digital mobility services. The Climate Law regulates digital service providers (both unimodal and multimodal) that have a route planning or a location tracking component, and it determines a series of policy measures that digital service providers must implement (in different stages) to ultimately achieve higher sustainability goals. Other regulations exist that govern the way data sharing should happen from public authorities to digital mobility services providers, such as the Finnish Act on Transport Services, but the data flow towards public authorities is a less regulated subject.

This issue leaves many local public authorities in a limbo, without EU or national regulation supporting their data reporting demands and limiting their capacity to understand and manage the local transport system. Whenever possible, cities are working with licence agreements or contracts that enable them to establish a data governance framework with new mobility services, because in the absence of national/EU-wide regulatory mechanisms they have very little leverage to demand data.

As seen in this section, the legal framework governing data, and thus, data sharing and reporting, is complex and fragmented. This is mostly due to the absence of a unified legal definition of data and a generally recognized legal status for data (as intangible assets).Currently, the absence of a standard status or access rule for data severely restricts data exchange and furthers the power imbalance in favour of those who create or run the systems, and hence control the data. Despite the Commission's efforts to advocate for data-specific legislation, there seems to be a number of legal gaps that impose restrictions on data exchange. The MobiDataLab analysis indicates that the different legislative provisions are not optimized to maintain uniformity, which seems to be particularly apparent when looking at the intersection of two separate pieces of legislation (referencing for instance the ITS and the PSI directives). The issue becomes even more concerning when looking at the pipeline of legislative interventions for the coming years in the EU (revision of the ITS directive, MDMS delegated regulation, additional delegated regulations...) and the existing and incoming regulations at the national level (Finnish Transport Act, LOM, French Climate Act). Considerable efforts need to be made to harmonize existing and future initiatives, and high expectations are being set in the NAPCORE project. Although the NAPCORE project should facilitate harmonization of the sector specific regulations, the EC cannot forget to work on the horizontal harmonization of data to minimize the existing legal gaps.

3. The data reporting value chain

Cities and public authorities in general need to ask for data to to address specific concerns, but before that they need to reflect on if data will be a solution. Data is a tool, not a solution. As mentioned by Stijn Vernaillen from the city of Antwerp: "Sometimes one well-placed block of concrete can do much more than a handful of programmers" (EIB, 2021). A vision, policies and good infrastructure are always the basis. Data is a layer that comes on top of that. Conversations with experts show that it is very necessary to have a data plan. Data can be used to support the SUMP. Just setting a data infrastructure, stuff it with some datasets and wait for the system to produce meaningful results is not enough. In this line, in the past, many cities have collected data without sense. Some even use the term "data hoarding": give me all the data because I think it will be useful in the future. This approach has proven to be expensive and provides no guarantee of results, plus it threatens trust and relationships between public authorities and private mobility providers, let alone the increased workload.

In this chapter, we will delineate about the purpose of collecting data, coupled with some relevant developed frameworks to inspire cities when getting ideas for use cases. In addition, we will also mention which data sources are generally available to evaluate such use cases.

3.1 Data with a purpose

As shared mobility services started permeating cities' streets, they realized the potential that the data generated by these services could have to achieve their long-desired goals. In this line, in order to guarantee order and safety on the streets, cities started requesting data from providers, but without a clear idea of what to do with it.

On top of that, some cities and private shared mobility providers started questioning if the city should intervene at all. Normally, the public sector intervenes when there is a market failure. When it comes to data reporting, there is a failure, in that the incentive to share data (provide accessible, efficient, and safe mobility for all) is not as great as the disincentive to share it. Note here that the term "market failure" should not only be understood as bankruptcy but as the dissociation between needs and wants (demand – the city – wants one thing and the offer – mobility providers – want a different one). Put it differently, if there was no "failure", all providers would share/report data straightforward when they arrive to a city, however, this is not the common case. At the same time, private providers began to see this practice both as a hurdle and as a threat. A hurdle because now they needed to curate their in-house data in order to provide it to the city in their desired way and also curate it in a way that it wouldn't expose sensitive commercial and personal information. And a threat because of what we just mentioned, reporting data always exposes the business in some sense. As an example, some providers say that with the appropriate algorithms, financial information from a shared mobility company could be obtained using the open data.

Among the main reasons for the lack of trust between public and private stakeholders are:

- Lack of capacity from the public side to digest and manage the amounts of data and its commercial sensitivity and personal information.
- A reaction to the very first way data reporting started, as a "give it all and we'll see what we do with it".
- Cities asking for data in different formats increasing workload and costs for providers.
- Lack of interest from private shared mobility providers to integrate with public policies.
- Private providers not operating correctly: parking requirements, speed regulations, geofencing, number of vehicles per area, etc.
- (Unknown) reasons behind providers not sharing data voluntarily.

To overcome this, cities should define a clear and use case-based framework for mobility data reporting, identifying data types and groups that support precise policy goals and targets.

In this line, the New Urban Mobility Alliance (NUMO) already started the work by compiling a set of use cases linked to policy goals. In this line, ITF also recommends aligning data-reporting mandates to targeted outcomes, providing a "why" to data reporting (ITF, 2021).

In them, the various **high-level** outcomes settled by cities are addressed through four groups, being:

- **Equity:** access to necessities for every community resident.
- **Safety:** ensuring streets are safe for residents, riders and pedestrians.
- **Environment:** reviewing the environmental impact of vehicles and services.
- **Usage:** understand how, where, and by whom are shared mobility services being used.

Each outcome is composed of different **use cases** that in turn are evaluated by a series of **key questions**, such as "How far do users have to travel to find a vehicle?". Each question is then evaluated by several metrics to monitor such policy outcomes. The metrics are classified in 4 groups:

- Evaluation Metrics
- Policy Metrics
- Equity Metrics
- Compliance Metrics

These metrics can be composed of various indicators, like the “average distance to an available vehicle” to “number of educational facilities accessible with micro-mobility”, and they address a broad array of aspects/issues that cities aim to encompass, sometimes reaching beyond mobility matters.

Finally, to measure each metric, relevant data sources required are included. We provide below an example, including the high-level outcome they target, the use case, the required data derived from it, as well as its sources, which is where most of the challenges lie regarding the variety of standards needed to build such indicators.

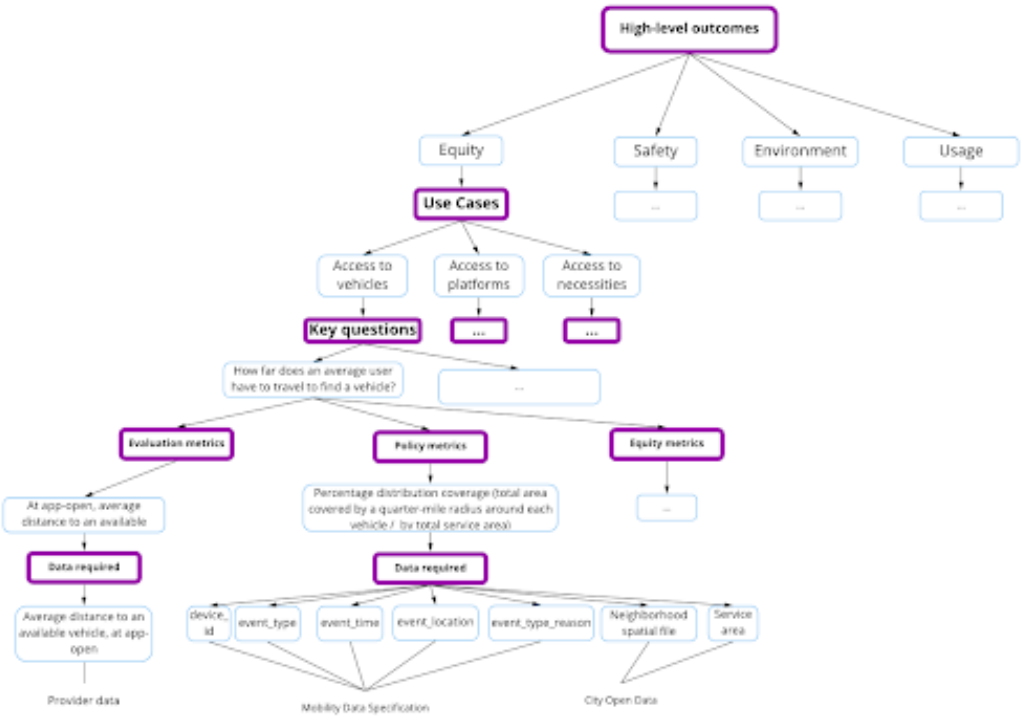

Figure 3: Breakdown example of a NUMO equity use case -
Access to vehicles (Source: Own generation based on NUMO use cases).

With a similar purpose, the CDS-M organisation started setting up a library of use cases to focus on specific questions or problems a city might have. In each use case, a research question is established, followed by a list of metrics and data that will contribute to its answer. In addition to the NUMO use cases, it specifies which standards (GBFS, MDS, TOMP, etc.) are needed for each use case, scoring them in terms of privacy and specifying if they contain personal data or not.

Such guiding frameworks can be of great help for cities, as stated by Vianova during the MOBI-MIX data session, cities often ask for data, but they don’t know exactly how it is going to answer their questions. This can also discourage operators from sharing. Companies like Vianova, a third party data aggregator and manager, prefer that cities reach out to them with the goal in hand rather than a “wish list” of data needs, so they can better tailor the data needs with the policy goals themselves.

The case of Antwerp, Belgium

"The uses cases developed from CDS-M can be interesting for smaller cities that don't have enough workforce working on data. They don't need to start from scratch"

Interviewee	City of Antwerp
Data regulations	The city requests data from providers. With the contract established with the providers, the city becomes the owner of data, under certain conditions. Data collected by the city can be shared with non-commercial parties, like research or governmental entities.
Standards collected	In the contract with providers doesn't specify a concrete standard, but the contract includes requirements that reflect GBFS and MDS. MDS is shared by almost all micromobility providers, and two car-sharing providers, with a third one following soon. GBFS is shared by all. Some providers also comply with TOMP-API, useful for the transactions component.
Data acquisition model	Conditionality – operators need to report and share data if they want to operate in the city
Value creation – data use	Use cases are not defined in a structured way yet. However, it is part of the city's roadmap, although it has a low priority at the moment. The city follows the use cases developed by NUMO and CDS-M. Part of the regulations is the maximum nº of vehicles. Now with the data, the city can monitor compliance with it. They can also ensure proper distribution of the vehicles in every neighbourhood/area of the city, as well as encouraging trips to more underserved parts. In addition, data has allowed us to have a more frequent dialog with providers, for example when they are not abiding by some regulations. At the same time, though, the fact of having a standardised way of working with providers helps to avoid unnecessary discussions.
Challenges	Most providers have concerns with MDS standards because of all the business intelligence and level of detail included (privacy concerns). If use cases for the data are not defined, many providers don't want to share/report it as it is not clear what is going to be done with it.
Lessons learnt	Three tips for newcomers: 1. Define the maximum number of vehicles 2. Create a policy for data requirements so there's an element to stick to. Beware of regulations not flexible enough to keep up with changes 3. Don't do pilots. First because there's no real impact from them and second because it's not interesting for operators, who need to be sure they operate for minimum 5 years with a big enough fleet to make some profit.

3.2 Indicators to monitor mobility

Using indicators can help make decisions on policy and planning of sustainable transport systems, and it is in fact common practice and part of any SUMP development process. Indicators measuring the strengths and shortcomings of any transport system try to consider all three dimensions of sustainability: environmental, economic, and social. The SUMP guidelines set four clear aims suitable for any monitoring process:

- Identify problems, bottlenecks and other challenges for on-time implementation.
- Keep track of progress towards achieving the targets.
- Adapt to new technological, legal, funding or political developments.
- Adapt and optimise the implementation process.

To standardize the evaluation of urban mobility systems, the European Commission has developed a comprehensive set of practical and reliable indicators to support cities when monitoring and evaluating sustainable transport measures. The Sustainable Urban Mobility Indicators (SUMI) can help cities to evaluate the performance of their policies, as well as compare it to that of other cities using a benchmarking tool. The indicator set is split into core and non-core indicators, it includes a wide range of mobility-related aspects and has also been designed to enable cities to measure improvements stemming from novel mobility policies or practices.

Core Indicators
Affordability of public transport for the poorest group
Accessibility of public transport for mobility-impaired groups
Air pollutant emissions
Noise hindrance
Road deaths
Access to mobility services
Greenhouse gas emissions
Congestion and delays
Energy efficiency
Opportunity for Active Mobility
Multimodal integration
Satisfaction with public transport
Traffic safety active modes

Non-core Indicators
Quality of public spaces
Urban functional diversity
Commuting travel time
Mobility space usage
Security

After an initial pilot with 50 cities within the SUMI1 project, feedback regarding the calculation of the indicators and the availability of data is now being implemented, and a revisited and simplified version will come out under the SUMI2 project.

Lack of good quality data is often one of the biggest barriers to frequent monitoring and evaluation of sustainable mobility indicators. Regularly collected data can help demonstrate that goals are met, that what is meant to work actually does, and can demonstrate to policymakers and residents that their city is improving. However, collecting data on a regular basis for all indicators can be a challenging task due to the various resource constraints that public authorities face in this domain. Considering data collection (and generation) is inherent to the operation of digital mobility services, their deployment and up-take could greatly help cities overcome this barrier, that is if appropriate data governance frameworks are in place. With innovations such as shared mobility and MaaS, cities all over the world have had these services operating well before their impacts are known. Given there are no defined KPIs to measure the impact of MaaS and digital mobility services (yet), they can be difficult to model, which makes data from service usage a crucial resource in order to really understand the impact of these services and their true value in the sustainable transport system. Cities and mobility service providers should agree on an appropriate set of indicators that can help evaluate the impact of these new (digital) mobility services in the system.

Having a set of defined indicators would also allow public authorities to monitor and manage the alignment of new mobility services with high-level sustainability goals, through the creation of incentives or advantages when sustainability targets are met or penalties in case of disproportionate negative effects.

3.3 An overview of sources and the information they contain

Main sources

Most of the data stemming from shared mobility is considered “floating vehicle data” (also known as floating car data, FCD): vehicles that can transfer their real-time location. FCD can be used for speed information, origin-destination, historical trends, travel times, etc. Motorised vehicles can, by their very nature, provide more data due to the use of navigation and on-board units. However, in recent years more data has become available for public transport and bicycles as well, thanks to the use of smartphone applications.

Among the types of data that can be collected through shared mobility, we can find:

- Speeds and travel times
- Origin-destination
- Individual trip information (place where it started/ended, date and time when it started/ended, distance, duration, transport mode, cost)

- Human behaviour (trip purpose, trip schedule, chosen route)
- Service provisioning (available vehicles, available distance on remaining charge, unique monthly users, average rides per user, unfulfilled trips, average distance to a trip at app open, monthly vehicles' removal/maintenance
- Vehicle characteristics (propulsion type, ID, battery status, life span)

Some cities in Sweden and in the Netherlands are substituting roadside hardware (like cameras or inductive loops) by acquiring FCD from third-party provider, having lots of benefits. For example, compared to roadside equipment, it is faster to set up and can cover bigger areas. Other cities, like Lisbon, won't substitute it. In any case, it's important to keep in mind that the quality depends on the number of individual vehicles that contribute to the dataset.

Complementary sources

Data stemming from shared mobility services also has its limitations. However, other data sources can add very valuable complementary information. In addition, cities are often already in possession of several data sources that could very well complement shared mobility data. Although sometimes urban authorities might not be aware of the whole set of data they own.

Mobile phone application location data

Many mobile phone apps collect location data of their users. Some emerging advanced algorithms are able to identify travel patterns together with the transport mode used for a specific journey (although this requires some further and more advanced data analysis, as companies like Nommon do). Among the information that can be derived from mobile apps' location data we can find:

- Speeds and travel times between two different locations
- Origins and destinations
- Individual trip information, such as route and/or mode taken for a trip
- Human behaviour, providing trip purposes
- System and infrastructure data
- Service provisioning

However, some concerns arise around these datasets due to the high possibilities they offer coupled with a high risk of privacy violations, as delineated in this article by the New York Times.

Bluetooth (and Wi-Fi) tracking

Sensors mounted in the city can detect Bluetooth enabled devices (like smartphones or vehicles). Suitable for detecting travel times or origin-destination.

- Speeds / travel times
- Origin – destination
- Individual trip information
- Service provisioning

Survey data

When none of the sources cited above are available, some approximations retrieved from surveys could still help to assess mobility patterns. As an example, census surveys often and increasingly include questions on vehicle ownership, mobility habits, workplace and home location, all being very exhaustive and openly available in high scales when run nationally.

Or even if data is available, some key questions cities have cannot be answered by trip or vehicle data alone. Surveys can bring estimations on modal splits and mode choice, familiarity with different modes, vehicle ownership decisions, users' demographics, etc.

As an example, the MOBI-MIX impact assessment survey was asking respondents about socio-economic characteristics, travel habits, familiarity with different mobility options, or intended change in response to a new shared mobility pilot available.

In addition, surveys can help respond more subjective questions cities might have, such as:

- Users' safety feeling when riding the vehicles.
- Travel mode that would have been used had the shared mobility service not been present.
- Reasons for using/not using the shared mobility service
- Users' vehicle ownership choices

The case of Vianova

"Operators are able to provide data. Cities, understand that the more data you have the more you can be aware of what's going on."

Interviewee	Vianova (data agregator)
Motivation	When shared mobility vehicles arrive in a city, most of the cities start wondering about basic operational monitoring such as the number of vehicles, their location, etc. Gradually, over time and exposure to data and its benefits, they start asking themselves more sophisticated questions, like parking regulations' compliance, impact of shared mobility on public transport or on carbon emissions, etc.
Standards collected	GBFS, MDS
Data provision model	Generally, the city will require data as part of an agreement/tender/contract. In other cases (more in the US but now getting common in the EU) cities ask for very specific data. Vianova has the advantage to get these kind of data avoiding cities the hurdles of having to deal with very specific or precise data.
Value creation – data use	Apart from the NUMO use cases, CDS-M has also started setting up a library of use cases, which is really useful for inspiration. Sharing use cases is incredibly useful for cities to not do more work than they need to do, and to know exactly what and where to get what they need.
Challenges	Data privacy is a real issue. Standards need to be compliant with GDPR. CDS-M is taking the DPIA process from GDPR. A different question is that commercially sensitive information requires a trade-off between cities and operators. There's a need to define and agree on what information can be made public and how confidential data will be managed. MDS is aware of these challenges and the goal is to make sure there is a conversation between cities and operators when designing contracts/tenders.
Lessons learnt	For providers: data sharing doesn't need to be punitive. The more data cities have access to, the more benefits providers can get out of it. For cities: understand the power of data, how to leverage it, and be explicit in the conditions of the permits. This is important to remember when setting up a use case and to know what to use the data for.
Gaps	In general, the more evolved and time cities have spent dealing with data, the more staffed they are. There's also a distinction between data informed and data driven decisions. Cities do not always retrieve insights from data to build policies, for example, adding a new cycling path. There is also a need to complement current data with data about gender, age, mobility behaviour, etc. Part of operators' work can be to provide more qualitative data about information pure data (standards) are not able to provide (socio economic characteristics, trip purposes). But this needs to be anonymised.

4. Policy actions and recommendations

In this final chapter, we will delineate which are currently the main challenges regarding mobility data sharing and reporting, and a final summary of recommendations both at EU level and at regional/local level.

Data sharing & reporting challenges

As it has been mentioned along the report, several challenges still remain on the path of successful data sharing and reporting. These challenges affect both the data providers (shared mobility companies) and the data receivers (public authorities). Some of these challenges are listed below.

Share Personally Identifiable Information (PII)

PII is a term that refers to data that can identify a person uniquely, or that combined with other data sources can successfully recognize an individual. An alarming example of this was shown by a researcher in 2014 after the release of New York taxi data containing details of every taxi ride in 2013. By combining this data with other publicly available sources he was able to track the home addresses of people hailing taxis in front of a night club, together with the property value, person's ethnicity, relationship status, court records and even their profile picture from social media.

Safeguarding proprietary and commercially sensitive information

In the same line as PII, openly sharing data on trips can pose the challenge of opening commercial information of a shared mobility company. Confidential data are sensitive, but not all sensitive data are confidential. Therefore, some commercially sensitive data may be shared or reported in certain circumstances and with adequate safeguards. But if this is not well established, risks of exposing a business may emerge. For example, the profitability of a certain mobility company, can be obtained using service ridership and by estimating costs (ITF, 2021).

Difficulty of anonymising data

Following the same example of the NYC taxi trip data, a possible solution goes through make data "more" anonymous. Aggregation of trips could be a solution, but it can also hinder data utility. In addition, aggregation is not always the solution. For example, if we know the total income of a neighbourhood, and we also know that the next month a certain person is leaving the neighbourhood, that person's income can be easily deducted. Some authors propose what is called "differential privacy" (Tockar, 2014), but the challenge demands that other alternatives are also explored, and that regulations ensure it.

Cost of data collection and storage

Shared mobility providers need to invest money, time, and personnel to meet data reporting requirements or agreements. If those are inconsistent and/or excessive, costs could escalate too highly and menace their business stability. The challenge lies in compensating these companies (not necessarily monetarily) enough to guarantee data reporting practices that are sustainable in time.

At the same time, cities also need to invest in digesting these amounts of data to give them a purpose, and this implies training or staffing their teams, let alone financially compensating providers for the data.

Lack of alignment on data standards

The variety of data sharing standards (GBFS, MDS, CDS-M, TOMP-API, etc.) poses a challenge for providers when different cities require different standards. Many of the existing companies operate in different cities, but sometimes their teams are not prepared to adapt to a new standard, and this supposes a challenge for their upscaling. This is particularly the case for small or "old" providers, as they find that they need to invest considerable amounts of money, time, and human capacity in adapting to new cities.

On the public authorities' side, many cities might feel lost at the time of choosing which standard suits their objectives best. Sometimes they might end up requesting compliance with several standards at the same time for different purposes, which increases the workload needed to process and use the data.

Expertise to analyse and visualise data

As the new data economy flourishes, many cities are finding themselves unprepared to face it. As stated in this report from POLIS, generally cities are not well-prepared in terms of capacity to clean, manage and analyse the data in general. On the other hand, the private sector is generally more prepared as mentioned in the report.

A reason behind this might be because recruiting and retaining data specialists is not easy for urban governments, due to rigid recruitment and career development policies and an inability to generally compete salary-wise with the private sector (EIB, 2021). A potential solution is to train current employees or to partner with universities to learn from data specialists and work with data tools, like in Lisbon.

The case of Berlin, Germany

"Bring all stakeholders together at some point [...] And start the process as early as possible!"

Interviewee	City of Berlin
Motivation	Know what is going on inside the city regarding shared mobility, particularly monitoring the number of vehicles.
Data regulations	The city has developed a permit to which providers need to apply. The approach follows the same permit that restaurant terraces or building charging infrastructure need to have – if any player wants to use public space, they need to get this permit. For shared mobility, a new part was added in this regulation, although only for the city of Berlin. The permit comes with a set of additional rules developed together with the operators during workshops. Regulations do not cap the maximum number of vehicles, because the city first wanted to get the data, without imposing any hurdles. In addition, the city still does not know the maximum number of vehicles it would need, because they didn't know how many vehicles were there, prior to this permit. Evaluation will follow in the next 2 years, if providers prove to be beneficial.
Standards collected	MDS, included in the permit for e-scooters. For bikes, MDS as well. Most bike sharing companies are also e-scooter providers, so they can provide MDS easily. Except for the city company (Nextbike). Deutsche-Bahn should also be able to provide this standard. Some providers have expressed complaints about it. The city gives them 3 months to provide it, if not, they have to leave. For car sharing, since the city is in the middle of a legal battle with some providers, it is not defined yet. Mopeds could apply to MDS because most providers are the same as e-scooter ones, although a workshop will take place in January to decide with them.
Data acquisition model	Conditional – providers need a permit to operate in the city, and the permit requires data sharing and reporting
Value creation – data use	Right now, data is provided in excel, in an aggregated manner, including: • N° of vehicles • Trip distance statistics • Trip duration statistics Currently the city has no time or resources to analyse it, although the data is not telling much. No specific use case defined yet. The city has launched a tender to get a data platform/aggregator and soon it will start gathering data using MDS. With it, they intend to analyse origin and destination points, mobility patterns, where to reinforce public transport (e.g. where a lof of car sharing trips are being done for out-of-city trips), etc.
Challenges	Some car sharing companies are not used to share/report data, as they started operating 20 years ago and this was not a requirement. Further, sometimes they do not know how MDS works and they need additional IT support to adapt. This also complicates the discussions with the city, if the providers are not aware of the technical details. Also, vehicles are not allowed to park in certain areas (parks, narrow streets, in front of museums or delicate places like holocaust museums, etc.). The city asked for this data to the different districts, but they sent couloured maps marked with pencil, when the city wanted to work with GIS (.shp format), and they needed to "translate" it all. In summary, the city lacks staff when mapping, and there's no special people to analyse the data, which is right now a gap.
Lessons learnt	Bring all stakeholders together at some point, also possible users and non-users that are affected as well (like people with disabilities, since these vehicles are affecting their lives). Also, help the operators to develop. And start the process as early as possible! Secondly, digitalisation. Don't get stuck in the past, invest in new technologies and in people that know how it works, with specific job descriptions for this (sometimes a city worker might omit working on it by saying it's not in their work description, to work with data).

Recommendations

EU level

Define and share concrete use cases

As it has been stated throughout the report, any obligations to generate and share data (both static, dynamic and historic data) should be based on concrete use cases. EU stakeholders, both public and private, should collaborate in the sharing of best practices and use case definition.

Establish an EU-wide legal framework

Establishing a global legal framework (at the EU level) would be helpful for allowing justified use cases and enforcing limitations on overreach. While any legal provision should be sufficiently general to accommodate various national contexts, to allow for the emergence of national and local approaches, and to respect subsidiarity, it should also have a certain level of specification to avoid drawn-out discussions about data protection and GDPR compliance on shared mobility and MaaS.

The definition of a high-level framework for data sharing at the EU level -including data types, formats, and granularity- could make the creation of data sharing agreements or contracts between public authorities and private mobility providers more straightforward. This can be especially relevant for smaller public authorities or private mobility providers by helping to leverage the mobility playing field.

Harmonise and clarify privacy legislation

Different Member States seem to have far stricter interpretations of GDPR than others. The EU needs to spend wider efforts harmonising and clarifying its privacy legislation, facilitating the implementation of GDPR in shared mobility policies and providing legal clarity for the secondary use of private mobility data. Horizontal harmonization across data regulations is necessary to lift the barriers and encourage data sharing and data reporting.

Accommodate national/regional/local specificities

As new transport-specific policies are planned, it remains crucial that any new legislation or policy initiative should acknowledge the important context specificity of mobility and the decisive role of local public authorities—given their expertise, central role as integrators, local knowledge, and legitimacy as elected officials.

Set up digital governance structures

Effective governance of digital environments will need the emergence of new and adapted governance structures that allow public-value data reporting by default. In this, line, the EC is committed to develop a European Mobility Data Space, which aims to be a sort of marketplace for data sharing, subject to rules and where every data owner will be able to decide who can have access to their data and under which conditions. Establishing such governance structures at the EU level can speed up the digitalisation and decarbonisation of cities, while increasing the capacities of public authorities to manage and govern data.

Establish a common, safe and useful standard

Adopting privacy-focused standards and requirements would benefit both the public and the private sector. All relevant EU stakeholders should cooperate and aim for a deeper involvement in standards' communities to ensure data specifications can address desired policy outcomes.

Local/regional level

Plan

Start with the urban strategy. Mobility objectives or the SUMP, and how to determine if data is needed to plan, monitor, evaluate or operate mobility. Clarify the ultimate goals of deploying shared mobility in your city (CO_2 emissions reduction, modal shift, etc.). Make sure to align and leverage European, national and regional policies for data reporting and sharing.

Decide what data do you need

As delineated before, data is not a solution if the problem is not well-defined. Determine how data can benefit the evaluation and monitoring of such goals. Determine which data type is needed to contribute to the objectives set. Section 1.5 of the EIB document on data sharing contains information on some data types and examples of how they can be applied. Ensure some redundancy, not relying on a single data source.

Investigate what you have

Investigate which data sources are available to provide the data types needed. In the same document mentioned above, there is a table linking data types with the sources where they are available.

At the same time, it is important to know which data is already owned within the city. It is also worthwhile checking other departments and regional and national authorities, as they might already have such data.

Look at your peers

Look at what other cities, networks or organisations have done in terms of strategies, measures, or projects and enquire about lessons learned. Even look at what data practices do other city departments or organisations have in place (e.g., public health, public works, or criminal justice departments).

Define the scope

Define the scope of the project, in terms of scale, resources and objectives. Once it has been achieved, it will be possible to expand it.

Set up data acquisition

When the data needs have been defined and the sources identified, a process to acquire it needs to be defined. Chapter 2.1 presents some data acquisition models and their pros and cons. Investigate for parties to work with and acquire data from. The outcome should be a list of data requirements to be shared/reported by mobility providers. Define also the format/standard in which you want to receive the data. You also need to define how often is data needed to be reported/shared (i.e. frequency) and for how long is it going to be stored, as well as who will have access to it.

Set up capacity

Look for what in-house expertise is there in the city, both in urban mobility and data processing and interpretation. If required, consider acquiring external expertise or staffing your team to ensure capacity is ready to define the specifications and evaluate the work done by third parties.

Establish the legal framework

Define what contracts and agreements need to be in place, while looking into what regulation is already in place and what needs to be modified. Pay special attention to data ownership, data acquisition, data usage and privacy.

Learn about the challenges between shared mobility data and PII. Some recommend engaging with the community of potential shared mobility users to understand concerns about shared mobility data.

Evaluate

Once the process has started, check that the data acquired meets the objectives that were set initially. Check if results make sense. If the outcome is satisfactory and useful, adapt and scale it up.

The case of Bergen, Norway

"Micro-mobility can be micro-regulated: it's important to go case by case and understand the provider's needs and motivations"

Interviewee	City of Bergen
Motivation	Investigate geofencing with Nivel
Data regulations	E-scooter providers are required to share data when they want to operate in the city, linked to compliance with national laws. They need to prove that they follow a standard on their application since day 1. Car sharing operators have to agree to share data with the city to be allowed to apply for their reserved parking slots. Further, when they apply for reserved parking slots, they have to bring the city data on the usage of shared cars (number of reservations per car the last 6 months) they have in the vicinity. However, there's no digital regulation as with the e-scooters, meaning the city gets this information per application.
Standards collected	MDS (e-scooters)
Data acquisition model	Commercial terms (through Nivel), as the city doesn't have resources and knowledge to deal with the data standards. It is also easier for operators, because they already have data agreements in place with Nivel.
Value creation – data use	• Enforcing parking • Regulating riding rules (e.g. maximum speed) • Monitoring the number of vehicles The first two points are related to national laws. MDS allows a two-way sharing. The city established the geofenced areas and they can get real-time data to see that providers are complying with the zones. In addition, the city tried to carry out a pilot intended to test differentiated pricing depending on the zone where e-scooters were parked, but this was never implemented and thus they couldn't see an effect on where operators had their e-scooters. The reason this pricing model wasn't tested was because one operator in the city didn't join the pilot project. To charge the participants in the pilot would therefore be an extra cost (a "penalty") for the operators that cooperated with the city. Once the regulations came in, and it became mandatory to have permission from the city to operate, the pilot project ended. In this situation the national laws don't allow this pricing model. It was possible in the pilot because that was a private, and voluntary, agreement between the city and the operators. For now, data is not being used to gain additional insights, like improving the transport system or deciding where to invest in cycling infrastructure.
Challenges	Regulations – to be able to require data on a tender, the local regulation needs to accommodate it, and it is based on national laws. But national laws regulate how cities should operate e-scooters (e.g. not riding at night), not which data they should request. National law acts as a "framework", but then cities need to create their own regulations based on national law, before they can make demands to operators. The cities/municipalities can choose how to regulate most aspects (e.g., riding at night, maximum number of vehicles, or parking rules). However, some aspects are subject to national laws and cannot be freely chosen (e.g., maximum size or speed (20 kph) of e-scooters, employer rights, fire safety).
Lessons learnt	Micromobility can be "micro-regulated": don't go big at the beginning, regulate case by case and understand the context and provider's needs and motivations.

5. Conclusion

In conclusion, data sharing and reporting is a practice that can bring multiple advantages and can be a tool in making cities more liveable. Data has the potential to ensure shared mobility brings benefits to cities in reducing the need for parking space and the number of cars and as a consequence, CO2 emissions.

Although different data acquisition models and data sources exist, the factors ensuring success are similar across cities. On the one hand, local authorities can plan and set up concrete use cases (parking regulation, speed limitations, underserved areas, etc.) with the goal of defining the data needs. On the other hand, the city should bring all stakeholders together to define collaboration agreements and secure an ecosystem where data sharing can flourish while creating the right circumstances for shared mobility providers.

Even though the impact of data sharing and reporting is visible both in the short and in the long term, the number of use cases is still limited and the potential to exploit it is high. The “use case libraries” that are being developed can bridge this gap, inspiring other cities and showing the benefits of data.

Nevertheless, as this insight report has shown, there are already numerous good practices and valuable lessons to learn from. Cities wanting to implement a data sharing and reporting ecosystem have the necessary tools to avoid potential negative issues. Hopefully, this insight report supports local authorities in this sense, providing them with insights on how (shared) mobility data can contribute to reaching wider liveability and decarbonisation goals.

Reference list

D'Agostino, M. et al. (2019). Mobility Data Sharing: Challenges and Policy Recommendations Contributors.

Deloitte (2016). Gearing for change - preparing for transformation in the automotive ecosystem. Available at: https://www2.deloitte.com/content/dam/insights/us/articles/3474_Future-of-mobility-gearing-for-change/DUP_Future-of-mobility-gearing-for-change.pdf

Dutch Ministry of Infrastructure and Water Management (2019). Blueprint for an Application Programming Interface (API) from Transport Operator to MaaS Provider. Available at: https://dutchmobilityinnovations.com/attachment?file=FN%2F9HXXdbnjwbPscgpt1LA%3D%3D

EIB (2021). EIB Technical note on data sharing in transport. Available at: https://www.eib.org/attachments/publications/technical_note_on_data_sharing_in_transport_en.pdf

Hausigke, S. and Kruse, C. (2021). 'Öffentliche Mobilität gestalten – Die Mobilitätsberichterstattung', in Öffentliche Mobilität. Springer Fachmedien Wiesbaden, pp. 269–300. Available at: https://doi.org/10.1007/978-3-658-32106-2_11.

Homem de Gouveia, P. and Babio, L. (2021). Sharing Data from Shared Micromobility. Available at: https://www.polisnetwork.eu/wp-content/uploads/2021/01/SHARING-DATA-FROM-SHARED-MICROMOBILITY_FINAL.pdf

ITF (2021). Reporting Mobility Data: Good Governance Principles and Practices. International Transport Forum Policy Papers, No. 101, OECD Publishing, Paris. Available at: https://www.itf-oecd.org/sites/default/files/docs/reporting-mobility-data-governance-principles-practice.pdf

NACTO, I. (2019). 'Managing Mobility Data'. Available at: https://nacto.org/wp-content/uploads/2019/05/NACTO_IMLA_Managing-Mobility-Data.pdf

Oregon DoT (2020). 'SHARED MOBILITY DATA A PRIMER FOR OREGON COMMUNITIES'. Available at: https://www.oregon.gov/odot/RPTD/RPTD%20Document%20Library/Shared-Use-Mobility-Data-Primer.pdf

Populus (2021). Mobility Data Standards: what they are & how to implement them. Available at: https://www.populus.ai/white-papers/mobility-data-standards

The ODI (2020). What mobility data has been collected and published during Covid-19? Available at: https://theodi.org/article/what-mobility-data-has-been-collected-and-published-during-covid-19/

Tockar, A. (2014). Differential Privacy: The Basics –. Available at: https://agkn.wordpress.com/2014/09/08/differential-privacy-the-basics/

The MOBI-MIX project has been funded by the European Regional Development Fund (ERDF)