EIB Technical Note on Data Sharing in Transport

European Investment Bank
The EU bank

EIB Technical Note on Data Sharing in Transport

European Investment Bank

The EU bank

EIB Technical Note on Data Sharing in Transport

Disclaimer
This report was commissioned by the European Investment Bank. It was prepared by Pieter Morlion, an independent expert on mobility and innovation, and Suzanne Hoadley, an independent advisor on urban mobility policies and intelligent transport systems.

The findings, interpretations and conclusions are those of the authors and do not necessarily reflect the views of the European Investment Bank.

For further information on the EIB's activities, please consult our website, www.eib.org. You can also contact our InfoDesk, info@eib.org.

Published by the European Investment Bank.
Printed on FSC® Paper.

Table of Contents

Acronym

ACPaaS	Antwerp City Platform as a Service
ANPR	Automatic Number Plate Recognition
API	Application Programming Interface
AVL	Automatic Vehicle Location System
B2B	Business to Business
B2G	Business to Government
CDR	Call Detail Records
CEN	European Committee for Standardization
CIRB	Brussels Regional Informatics Centre
CROW	Platform for Transport, Infrastructure and Public Space (NL)
DevOps	Development and Operations
EIB	European Investment Bank
EMT (Madrid)	Municipal Transport Enterprise (Madrid)
FCD	Floating Car Data (see FVD)
FVD	Floating Vehicle Data
GDPR	General Data Protection Regulation
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
IoT	Internet of Things
ITF	International Transport Forum
ITS	Intelligent Transport Systems
ITxPT	Information Technology for Public Transport
LEZ	Low Emission Zone
LoRa	Long Range (communication protocol)
M4	MultiModal Mobility Manager (Antwerp)
MaaS	Mobility-as-a-Service
MDS	Mobility Data Specification
MMTIS	Multimodal Travel Information Services
NACTO	National Association of City Transportation Officials
NAP	National Access Point
NDW	National Road Traffic Data Portal (NL)
NeTEx	Network Timetable Exchange
OD	Origin-Destination

OMF	Open Mobility Foundation
OpenLR	Open Location Referencing
OSM	OpenStreetMap
PCP	Pre-Commercial Procurement
RTTI	Real-time Traffic Information
SaaS	Software-as-a-Service
SLA	Service Level Agreement
SRTI	Safety-Related Traffic Information
STIB/MIVB	Brussels Intercommunal Transport Company
SUMP	Sustainable Urban Mobility Plan
TfL	Transport For London
TMaaS	Traffic Management as a Service
UITP	International Association of Public Transport
UVAR	Urban Vehicle Access Regulations
WBCSD	World Business Council For Sustainable Development

Executive summary

Rationale

This document was commissioned by the European Investment Bank (EIB) to support and advise local authorities on data acquisition in the field of urban mobility. It aims to serve as a starting point for cities and municipalities seeking practical information and to provide a basis for an efficient and effective investment of resources. This document provides information and recommendations on the following topics:

- Strategic, budgetary and capacity requirements in data acquisition initiatives.
- Data types and their relation to commonly used data sources in the field of urban mobility, for use in mobility dashboards and other applications.
- Alternatives to straightforward procurement of data and a division of data acquisition models into seven categories.
- Contracts and licensing models for data acquisition, including data privacy and the General Data Protection Regulation (GDPR).
- EU policy framework in relation to data acquisition by urban authorities.
- Future data trends to consider.
- A concise roadmap for urban mobility data acquisition.

Method

As the field of data and data acquisition is in constant motion, an extensive index was made of current initiatives, organisations and experts on the topic. This cross-section of noteworthy activities in the field (Annex III) was drawn up from an extensive literature review (Annex II) and a survey amongst experts. In consultation with the steering body for this study, a shortlist was made of items and persons to be featured in this document. Seven models for data acquisition were defined to categorise the information collected, and for each model two to four case studies were selected. In total, more than 30 experts from cities, startups and companies, citizen collectives and academics were interviewed to gather best practices and recommendations in their fields of expertise. The best practices and recommendations from the interviews were structured and supplemented with links for clarification and further reading.

Structure

The information obtained from the literature study and the interviews covers many subjects, countries, objectives, transport modes, data sources and applications. The information has been organised as follows:

- An overview of common data types and data sources, including references to mobility dashboards, aggregation, strategies and budgets (chapter 1).
- A data acquisition categorisation consisting of seven models, ranging from straightforward procurement and in-kind partnerships to mandatory data sharing and crowdsourcing. A comparison between the models is described, as well as recommendations and lessons learned from over 30 experts (chapter 2).
- Some guiding principles and recommendations regarding applicable legislation, contracts, data privacy, capacity building, the EU policy framework and future trends (chapter 3).
- A pragmatic roadmap for data acquisition in cities (chapter 4).

Different aspects of a specific case, city, topic or data source are referenced over these chapters. For the sake of coherence, frequent references have been made to other sections of the document. To illustrate this, the box below describes where the various elements of the Traffic Management as a Service (TMaaS) project appear throughout the study.

> Case: The city of Ghent initiated a project to implement the Traffic Management as a Service concept – offering traffic management functions to cities through a digital platform. Two researchers from Ghent University (cf. the list with interviewees on the next page) were interviewed on this topic. This case is mentioned in the intermediaries data acquisition model (cf. section 2.3), as the platform stands between a number of data sources (cf. section 1.6) and urban authorities. The project built a dashboard (discussed in section 1.4 "Mobility dashboards"). The as-a-service platform unburdens cities in terms of contracts (cf. section 3.2.1) and takes care of some legal aspects related to the data (cf. section 3.2). This results in the city being able to focus on capacity and expertise regarding data in urban mobility, which is described in section 3.3.1.

Conclusions

Recommendations and conclusions have been introduced directly in the chapters to accompany the subject to which they are referring. Some general observations are provided below:

- An urban mobility strategy, such as a sustainable urban mobility plan (SUMP), and a data strategy should be the starting point for data projects. Randomly storing large amounts of data without a clear objective might not lead to satisfactory results.
- Many cities already hold useful data in some form in the systems or services they are using – working with these data makes a great starting point.
- Both cities and suppliers see the best results by starting small and scaling up when intermediary outcomes are successful.
- The same data can lead to different conclusions/information when processed differently.
- Mobility dashboards can be a good way to visualise data, but are rarely the end product of a good data strategy.
- Not all data sources are capable of delivering the required data type or information.
- Mobile phone application location data are a data source with great potential, but they are raising substantial privacy concerns.
- There are many alternatives to straightforward procurement for data acquisition, but not all of these acquisition models are applicable in every context.
- Intermediaries like aggregators and integrators are gaining importance in the urban mobility data ecosystem and play a role in many of the cases studied.
- Specific clauses on data (e.g. ownership) should be included in any contract or agreement.
- Technical skills, subject knowledge and procurement expertise are needed in data projects. Some tasks can be outsourced, but cities should have the capacity to understand tender specifications and validate project outcomes.
- At the moment, EU legislation on mobility data is mostly impacting EU Member States – this is expected to shift towards cities.
- The GDPR has a huge impact on conversations and projects on data. Some cities choose to have personal data processed by third parties to avoid privacy and security concerns.

List of interviewees:

Stijn Vernaillen	MaaS expert	City of Antwerp
Ahmet Demirtaş	Business Operations Manager	Parabol
Tuğçe Işık	Growth & Digital Marketing	Parabol
André Ormond	Data & Mobility Consultant	Ormond Consultoria e Treinamento
Tiffany Vlemmings	Project Manager Innovations	National Road Traffic Data Portal
Pedro Barradas	Chief Strategy Officer	Armis
Noam Maital	CEO	Waycare
Mélanie Gidel	Data specialist, Mobility Agency	City of Paris
Olivier Dion	CEO	Onecub
Stephanie Leonard	Head of Innovation & Policy	TomTom
Jeroen Brouwer	Manager Customer Program Management	TomTom
David Thoumas	CTO	Opendatasoft
Peter Mechant	Researcher	Ghent University
Timo Latruwe	Researcher	Ghent University
David Cunha	Smart Cities Advisor	City of Lisbon
Malin Stoldt	Project Manager ITS	City of Gothenburg
Rob Roemers	Head of Data and Analytics	STIB/MIVB
Sergio Fernandez Balaguer	International Projects and Cooperation	EMT Madrid
Jascha Franklin-Hodge	Executive Director	Open Mobility Foundation
Rick Meynen	Data Process Officer	STIB/MIVB
Thibault Castagne	Founder & CEO	Vianova
Hildegunn Grindheim	Product Owner Data Lake	City of Bergen
Edoardo Felici	Policy Officer	European Commission
Bon Bakermans	Advisor Innovation in Mobility	Ministry of Infrastructure and Water Mgmt
Per-Olof Svensk	ITS Advisor & Project Manager	Swedish Transport Administration
Miguel Picornell	Mobility Analytics Director	Nommon Solutions and Technologies
Tomas Straupis	Volunteer & GIS expert	OpenStreetMap
Otto van Boggelen	Coordinator	CROW
Emmanuel de Verdalle	Technical Committee Chair	ITxPT / Convenor of TC 278/WG 3

Kasia Bourée	Standardisation expert	Data4PT
Stephan Corvers	Founder, Owner & CEO	Corvers Procurement Services BV
Tijs de Kler	Technical Information Analyst & Developer	City of Amsterdam
Christoph Pyliser	Program Manager Circulation Plan 2017	City of Ghent

Chapter 1: Data overview

1.1 Introduction

Urban mobility is a specific context that requires other data and insights than, for example, highway management or a city's human resources (HR) department. This first chapter focuses on defining a data strategy that responds to the specific needs of a local mobility department and how to get meaning out of data. Through a study of mobility dashboards, this chapter then describes the types of data that are frequently used, followed by data sources that can provide these types of information. Finally, some insights on budgets for acquiring data are given.

The content of this chapter, like the rest of the study, is based on interviews with more than 30 experts from different parts of the world, working for cities and other public authorities, private companies, academia and citizens' groups. For this chapter specifically, they were asked about their experiences with data acquisition, data sources, dashboards and data strategies. The second important source of information was a literature review based on existing data inventories, dashboards and literature (cf. Annex II). The information from interviews and a literature study was supplemented with the authors' expertise working in the fields of urban mobility and data.

This chapter focuses on the importance of starting from a data strategy. Although the possibilities of big data and artificial intelligence may seem endless, the research has not shown any city achieving a lead that cannot be overtaken. On the contrary, by starting from a clear strategy, learning from peers and engaging in constructive partnerships (cf. chapter 2), cities starting today have good chances of catching up with the front runners.

1.2 A data strategy

Although defining a sustainable urban mobility plan (or SUMP) is outside the scope of this study, any city implementing data projects should clearly link such initiatives to its urban mobility strategy. The SUMP framework offers a number of starting points for data collection (cf. these webinars on SUMP data collection and analysis). Data acquisition, exchange or analysis is one of the tools and measures to support sustainable urban mobility – it is rarely the only solution and is best combined with other measures or intelligent transport systems (ITS).

Pedro Barradas points out that tools exist for SUMP data collection and monitoring, like Way4Smart, a solution from his company Armis: *"By collecting static and dynamic data, W4S characterises the mobility services offered by local operators (new mobility services, road traffic, public transport and others), making it possible to manage and supervise their effectiveness and quality of service, while assessing their suitability for the urban mobility plan and enabling information exchange with other entities, namely the National Access Point."*

Stijn Vernaillen, city of Antwerp:
"Data are mostly the layer of varnish on the entire mobility ecosystem in a city. Data are enablers, but you can't solve real-world issues with fancy data solutions, if all the layers below are not well managed. A vision, policies and good infrastructure are always the basis. Sometimes one well-placed block of concrete can do much more than a handful of programmers."

Given the great progress made in the areas of big data and artificial intelligence, there is sometimes an expectation that it is sufficient to set up an infrastructure, stuff it with random datasets and wait for the system to produce meaningful information. In practice, this is not the case. Case studies, literature and conversations with experts from all over the world have shown it is extremely important to have a clear plan before working with any kind of data. Many cities and projects have collected data haphazardly, without considering what exactly should be done with it. Ahmet Demirtaş from Parabol confirms that some cities practice "data hoarding": collecting all kinds of data, hoping that they will turn out to be useful in the future. This is expensive and offers no guarantee that the data will be useful. Often adjustments are needed afterwards because the data have the wrong format, do not meet quality standards, or there are data missing that cannot be retrieved retroactively.

Noam Maital, Waycare:

"Cities should think about the problem and things they want to solve and THEN ask themselves if data can be a solution. Too many cities start with asking data to see what the problems are."

When defining a data strategy, a good starting point is defining the goal for data collection. This can be, for example, operational mobility management, planning and monitoring urban mobility policies, auditing or following up on contracts or agreements with service providers. Antwerp also collects data to influence behaviour through a large communication campaign. Pedro Barradas from Armis indicates that a lot of guidance is given through delegated regulations by the European Commission (cf. chapter 3) and that *"if your plans fall within the scope of those regulations, any investment is a good investment."* Some cities in the United States and in Europe (like London and Ghent) have published city-wide data plans that can serve as inspiration or as a handhold as well, but beware of some pitfalls. A good way to make the strategies tangible is by defining indicators, like the Sustainable Urban Mobility Indicators proposed by the European Union. These might be cycling volumes, air quality, vehicle speeds or behavioural data (cf. section 1.5 below).

The same data can be used by different departments for different purposes. Therefore, a data strategy can rarely be defined by one department on its own (cf. section 3.3.3 "Integrated approach to data management" and the examples cited above) and is instrumental in promoting synergy between different parts of the city administration (and other stakeholders).

Christoph Pyliser, city of Ghent:

"For the introduction of our 'circulation plan' in 2017, we collected KPI data like travel times, public transport and car usage, number of cyclists and air quality. This allowed us to clearly demonstrate an uptake in walking, cycling and public transport and a decrease in car usage afterwards. In our next urban mobility plans, we would like to collect even more parameters, like the impact of policy measures on local businesses' revenue."

Once the data purpose is defined, data sources or systems (cf. section 1.6 below) can then be looked for that can deliver these types of information, like Bluetooth sensors, floating vehicle data or number plate recognition cameras. On this matter, experts indicated that cities tend to immediately start working with promising third-party data sources or services and sometimes ignore the data potential of systems they already own or operate. This includes information from other city departments. If the data are not yet available to the city, procurement is not the only solution. Chapter 2 of this document elaborates on a number of models for data acquisition.

André Ormond, Ormond Consultoria e Treinamento:

"Cities have amazing data! Instead of immediately acquiring expensive new datasets, they should make sure they gather the knowledge, capacity and vision to work with the data they already own."

Thirdly, before starting to implement, procure or build, it is recommended to consult experienced cities. Reviewing best practices and lessons learned can be insightful and save a lot of time and money. Besides other cities, there are a number of networks and organisations that help disseminate expertise like Ertico, Eurocities, Polis and national ITS organisations. The European Union also actively stimulates and subsidises a number of projects and forums (like Civitas and Eltis) that hold a lot of knowledge.

David Thoumas, Opendatasoft:

"You have to start small. Don't hesitate to start small and quick. Once you start working with – and sharing – a dataset, this will generate insights and feedback, it will feed processes and quality can be improved along the way."

One of the last steps before getting started is budgeting. The city of Bergen emphasises the need to carefully estimate the size of your project and the impact on budgets and infrastructure. Other cities and even the suppliers interviewed also underlined that it is best to start small. Small pilots or proof of concepts allow cities to check their assumptions, get a grip on the actual use of specific data and evaluate if the project is capable of meeting the objectives set. If the project is evaluated positively, it can be fine-tuned and scaled up. This enables efficient spending of funds, energy and resources.

1.3 From data to wisdom

Once datasets have been acquired, they generally do not give immediate insights into urban mobility. In many cases, the data need to be processed, aggregated and visualised. Speed data, for example, can be grouped by day of the week and merged or fused with other datasets, like a list of bank holidays, weather, roadworks or events before it is shown in graphs or a dashboard. In some cases, the data need to be cleaned, meaning that, for example, measurements from defective induction loops and extreme values are removed.

Miguel Picornell, Nommon Solutions and Technologies:

"We have seen cases where other companies work with the exact same data as we do but get completely different results. The way the data are cleaned, processed and the assumptions made have a big impact on the outcome. Not only are the data important, but also the interpretation and knowledge of how they were obtained. It is important to distinguish between data (raw data) and information (processed data)."

In this process, technical expertise on data science and manipulation is logically very important to get the correct information out of the data. Conversions between formats, timestamps and coordinates should be carried out correctly. In the Gothenburg case described in chapter 2, a supplier and the city are using two different maps. Even though OpenLR (a map-agnostic location method) encoding and decoding are used, extensive manual adaptations need to be made by the customer whenever the supplier's map is updated. The city of Lisbon built a horizontal data team in order to ensure access to state-of-the-art technical data expertise for all departments. Because the process from data to information can transcend the city's technical capacities, some urban authorities increasingly choose to have the data processing done by specialised companies (cf. section 2.3 "Intermediaries" and section 3.3.1 "Capacity and expertise").

Data skills aside, domain knowledge on urban mobility is at least as important when manipulating and interpreting data. First of all, familiarity with the city and its policies can help to correctly interpret the data. For example, low-emission zone regulations are different across Europe, meaning enforcement data cannot be simply compared.

André Ormond, Ormond Consultoria e Treinamento:

"To forecast the stress on our mobility network during the 2016 Olympics, we ordered an extensive study. An excellent job was done, but the numbers were completely off. The data for the study dated

from before a huge economic crisis hit Brazil, decreasing road volumes by almost 50% in some areas of the city. The study also did not take into account the huge roadworks preparing for the Olympics that were hindering traffic the years before the games. The study was biased by disregarding economic, political and other contexts at the time of data collection."

Also, insights are needed on how the data are collected. The International Transport Forum warns about information biases: some data only reflect behaviour by social groups that have access to data-generating technologies, like smartphones. Insights into how the data are collected helps estimate the reliability of values. For example, section 1.6 below explains why floating vehicle data might not be reliable in small urban streets and why Bluetooth sensors are better at detecting cars than pedestrians.

Finally, it is important that all involved in collecting, processing and viewing data should maintain the same definitions. For example, STIB/MIVB – Brussels' public transport operator – indicates that Belgian transport operators have different codes for the same station, resulting in difficulties when converting data, for example, converting to the NeTEx standard as is required by EU legislation. This study on the semantic integration of urban mobility data, the SPRINT project and yearly Sem4Tra workshop are working on the subject, as well as the Flemish government, which is creating a common vocabulary for mobility.

Mélanie Gidel, city of Paris:

"In our experience, for certain use cases and types of data (especially GPS data), it is easier for an administration to work with preprocessed or aggregated data."

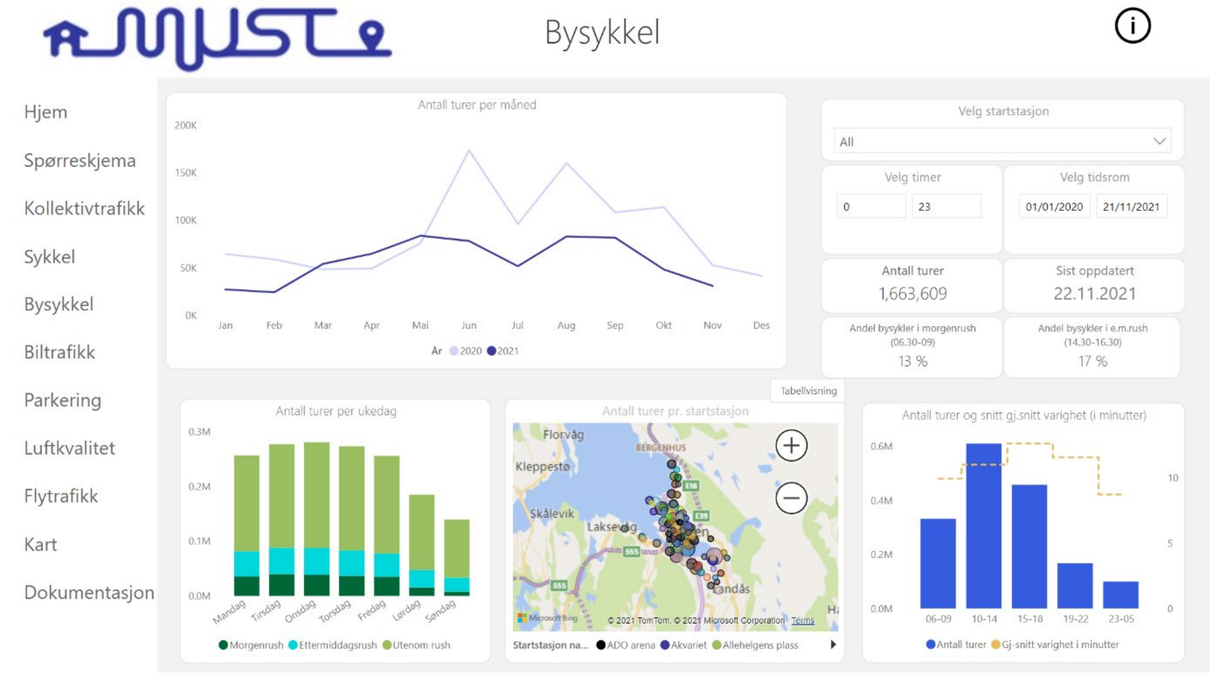
1.4 Mobility dashboards

The rest of this chapter provides an overview of urban mobility data types and data sources commonly used by cities. To maintain some focus – there are hundreds, if not thousands of data types and sources – data that have been used to build publicly available mobility dashboards are described for two reasons. On the one hand, dashboards prove that data are available to cities and it is technically feasible for them to process these data. On the other hand, visualising the data makes it easier for other cities and the public to get a grasp of the possibilities and limitations of these datasets.

Proprietary dashboards (as opposed to publicly available dashboards discussed here) can be useful in the context of:

- Traffic management (like the Metis or Waycare platforms).
- Monitoring service provision (like Vianova, an Arcgis visualisation for MDS, a dashboard to follow up on operators of shared bicycles or public transport concessionaires (cf. case study 5 in the EIB Technical Note on ITS Procurement for Urban Mobility).
- Event management (like Tripservice).

As many of these cases are described in chapter 2, in this section we focus on publicly available mobility dashboards with real-time information like the Izmir City Dashboard, Link (Ghent), MUST Mobilitetsdashboard (Bergen), TfL's road danger reduction dashboard (London) and the COVID-19 mobility dashboards from the Inter-American Development Bank (Waze data), the Spanish government (mobile data), Direct Relief (Facebook data), Cuebiq (smartphone GPS data) and Google Community Mobility Report UK (Google data).

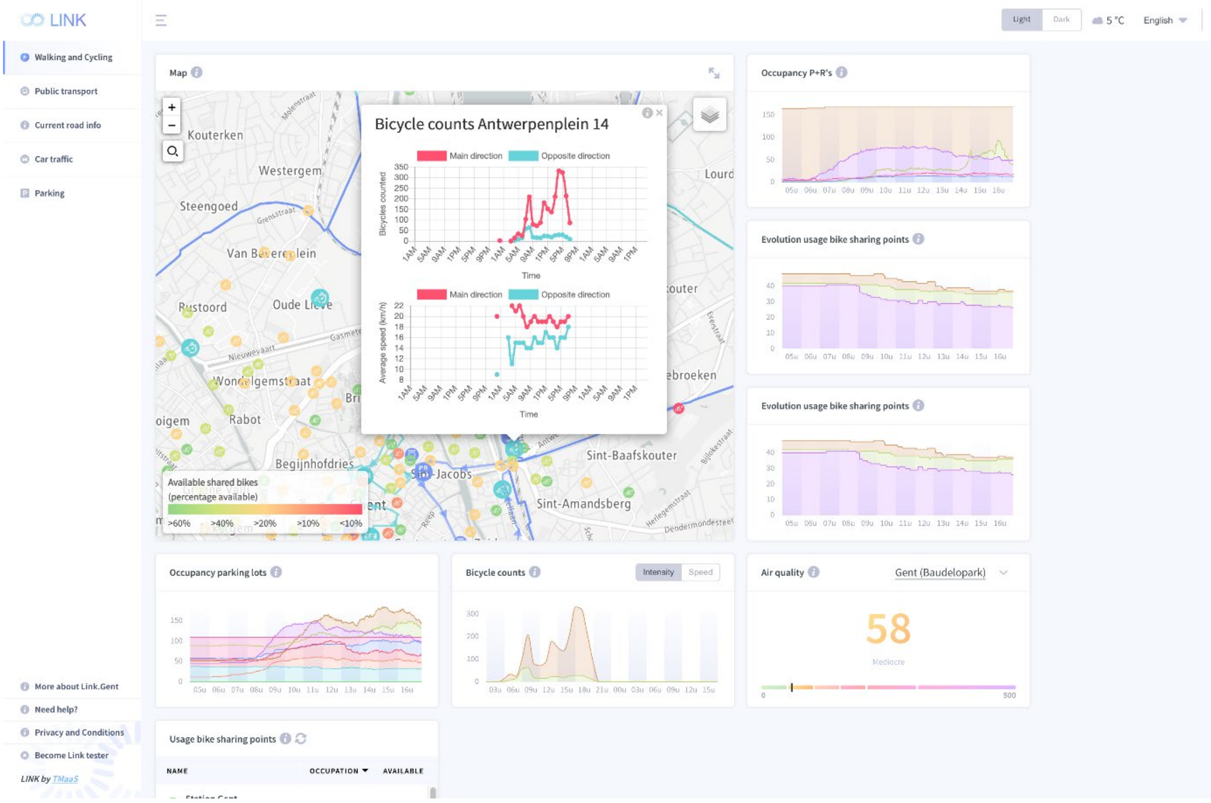


Bergen's mobility dashboard. Source: MUSTlab.no.

David Thoumas, Opendatasoft:

"The experience of the Centre-Val de Loire region releasing a dashboard is very interesting. There was a lot of traction and press attention because it did not communicate on technical matters like data, but on an everyday problem for a lot of people: train delays."

Dashboards can be a good way to make data and efforts visible somewhere along the journey of implementing a data strategy. Collecting data on mobility is a process that can be time-consuming, and its results are mostly invisible to the city council, administrative staff and the public. A dashboard is a good tool for making such efforts visible. It can be the basis for informed conversations about mobility, about what data to collect and how to display them. A public mobility dashboard can, however, rarely be the ultimate goal of a data strategy. This is the case with the Link dashboard, which is the first step in journeys towards big data processing and automatic messaging services.



The Link mobility dashboard. Source: [TMaaS]

Lessons learned from designing a dashboard suggest that it is better to have a clear target group in mind and to ensure support within the city organisation. The absence of these two elements led, for example, to the discontinuation of the Manyways dashboard, even though it was perceived as a versatile dashboard and received a lot of press attention. Most of the companies interviewed also recommend starting small and working iteratively, based on participation and feedback from the public. Cities that are considering building a mobility dashboard can consult these articles about designing a mobility dashboard, the importance of imagination, examples of dashboards and specifications for a tender from the Manyways.be project (in Dutch).

1.5 Data types

While there are hundreds of different data sources on urban mobility, they can be categorised through a number of data types. Before identifying the data sources needed, it is important to assess what data types (like speeds, travel times, number of bicycles, air quality, etc.) are needed to feed or monitor the urban mobility strategy. This chapter gives an insight into what these different types are, their use and what to pay attention to.

	Mobile app data	In-vehicle data	Bluetooth / Wi-Fi	Mobile data	Inductive loops	Floating vehicle data	(ANPR) cameras	Ticketing systems	Automatic vehicle location	Air quality	Weather	Citizen-collected data
Volumes					x		x	x				(x)
Speeds / travel times	x	x	x		(x)	x	x	x	x			
Origin-Destination	x		x	x		x	(x)	x	x			
Individual trip information	x	x	x	x		(x)	x	x				
Parking usage	(x)	(x)					(x)					
Contextual data		x								x	x	x
Human behaviour	x	x		x		x	(x)	x	x			x
System and infrastructure data	x	x										
Service provisioning	x		x	(x)		x	(x)	x	x			x

Volumes

A type of data that is often used by cities is volumes, for example the numbers of cars, pedestrians, cyclists or public transport users. These data are used to monitor how intensively a certain infrastructure or service is used over time (like this Connecticut dashboard) and can give insights into additional traffic due to large events, road closures or policies. The effects of introducing a circulation plan in Ghent were measured using the volumes of cars, cyclists and public transport users in the city.

The data can be collected by placing induction loops in a road or bike path, using a radar or an image recognition camera that counts vehicles, bicycles, buses or pedestrians, like the systems from

Viscando. This can also be a licence plate recognition camera (which would only count vehicles with a licence plate). To measure the number of public transport users, data from smartcard readers, a camera in the vehicle or at stops, or the weight of the vehicle can be used.

It is important that all road users are counted, or enough information is available to safely extrapolate. The use of data sources that map only a relative part of traffic (Bluetooth, floating vehicle data, mobile data) cannot guarantee conclusive results and are therefore not recommended to collect data on volumes.

Speeds and travel times

Speed data can be used to get an idea of traffic flows (and congestion) and speeding offences, or to ascertain the need for speed reduction measures or enforcement (although this sometimes leads to protest: "TomTom sorry for selling driver data to police"). Speed data are also used to monitor service levels of public transport, known as commercial speed. An overview of the average commercial speed in 15 cities around the world can be found in UITP (2009) Buses Today and Tomorrow report. An interesting use is creating accessibility heatmaps: for example, how far can one travel by foot, bicycle, public transport or car in 30 minutes in your city?

Speed data are often collected by double induction loops in the road, ANPR cameras for average speed enforcement, Bluetooth sensors, radars and floating vehicle data. The speed of a vehicle can also be mapped by tracing the smartphone of a driver or passenger, as Waze does for vehicles. Moovit and other companies sometimes use the same approach for public transport.

When speeds are needed over certain routes rather than at specific points, **travel times** can be used. Speed, distance and travel time are related: if you know two out of three, you can easily calculate the third. The advantage of expressing flows in terms of travel time is that the geographical factor is always clearly taken into account – a number of trajectories are defined. These can be main roads, certain popular routes or the entire ring road to get an idea of traffic fluidity. To cover road networks, the network can be split up into smaller segments for which travel times are calculated. Rio de Janeiro and the National Road Traffic Data Portal in the Netherlands use this. When a travel time is needed for a certain route, the values of all the smaller segments making up the route are added up. This study from the US Department of Transportation compares ten different ways to collect travel time data.

The importance of real-time travel times should not be overestimated. In particular, it is the reliability and predictability of the travel time that are important, especially when it comes to public transport and bicycles. This US Department of Transportation study, called "Does Travel Time Reliability Matter?" is entirely devoted to the subject.

Origin-destination data

Insights on where people are coming from and where they are going to (and when) is a good source of information for urban infrastructure and service development: planning new bridges, underground lines or ride-hailing services or finding optimal locations for Park and Ride sites or mobility hubs. Such origin-destination information also reveals where people are coming from that travel to a certain shop, school, event, neighbourhood or city. This can help authorities understand travel behaviour during a pandemic (like in Spain), transit accessibility to jobs or mobility equity. This study shows the potential by mapping out 1.1 billion taxi and Uber trips in New York. Origin-destination information can be condensed in origin-destination or trip matrices (TomTom's website gives an overview), which are fed into traffic models.

Origin-destination information is traditionally derived from surveys where people are asked about their trips, or based on datasets of home, school and work locations. The information is then extrapolated and sometimes cross-checked with traffic counts or ANPR camera data. Mobile phone data have become an excellent alternative (cf. section 1.6). This study compares the different ways of collecting origin-destination data.

Individual trip information

Mobile data, in particular tracking apps like Strava or Google Maps and connected devices like shared e-scooters, are capable of providing much more detailed information than just the origin and destination of trips. Other sources are on-board units for toll collection or data from Mobility-as-a-Service (MaaS) apps. They track each individual trip and show (in an aggregated way) which routes were taken. This information is extremely useful to detect missing links in a road or cycling network and monitor the uptake of new infrastructure. An example is this heatmap from the Bike Data Project. This kind of data is the subject of many discussions around privacy, because in its rawest form it shows the individual trips of individual users, indicating where a person has travelled at a certain moment and making it possible to extract, for example, the home location of the user (cf. section 3.2.2 "Data privacy").

Parking usage

Data on the occupation of off-street and on-street parking and Park and Ride sites are used to plan and monitor parking policy, to simulate or verify parking revenues, explain parking policies to residents and businesses, and guide cars to a vacant parking spot. It is interesting to know what the parking duration is and who is parking: residents, visitors, company employees and, for example, people with a disability.

Off-street car parks are usually equipped with a barrier and a payment system that keeps track of the occupancy and indicates when a garage is full. However, as a city it is not always easy to obtain or share these data with the public. Some parking operators consider it commercially harmful to share such data with cities and especially competitors. ITS Belgium has proposed contractual clauses for concession contracts that give cities insight into the data in a way that safeguards the commercial interests of the operator (cf. chapter 3). The occupation of off-street parking for bicycles is also being tracked, like in Utrecht and Belgium. The data are collected using sensors and increasingly by means of camera systems (such as those of LumiGuide or CyclePods).

Exact figures for on-street parking occupation were typically difficult to obtain. This has improved with digitally connected parking meters and parking registers (cf. chapter 2), scan cars (like in Amsterdam) and parking sensors (like in Kortrijk). Some navigation system suppliers also offer data that provides estimates of parking occupancy, like HERE and TomTom. Some companies specialise in on-street parking prediction, or what they call parking car data, like Bliq or Parknav.

Contextual data

There are a number of variables that impact urban mobility and can explain changes in mobility patterns: weather, economic activity, holidays, events, school, strikes, pandemics, accidents, recessions, etc. The most common example is the impact of weather data, such as rain, fog, frost and snow, on cycling behaviour and traffic, for example. The link between major events and traffic jams might also be logical. Less obvious are the (alleged) links between football matches and traffic deaths or drink driving laws and suicide. This information is mostly obtained through collaboration with other city departments.

Conversely, urban mobility (policies) also has an effect on life in our cities: housing prices, energy consumption, air quality, liveability, noise levels, economic activity, etc. For example, this study describes the impact of pedestrian zones and cycling on local businesses in ten cities around the world *("improve urban life, both by soft and hard factors. It raises local businesses' revenues").* There are studies on the effect of Urban Access Regulations and low-emission zones on air quality (*"Low Emission Zones have had a positive impact on air quality in many European cities"*) and on social equity *("This report demonstrates a spatial vulnerability assessment of the introduction of Low Emission Zones in Scotland").* The book Street Fight tells a compelling story of the effects of introducing more space for pedestrians and cyclists in New York. When monitoring the effects of mobility or implementing new policies, a cross-departmental approach is key, finds the study "Using the power of data to address urban challenges and societal change".

Human behaviour

Data are collected on the average speed of vehicles, number of cyclists, bus delays and number of accidents involving pedestrians. Note that in these examples, as is done in most cases, data are categorised by transport mode. For a policymaker and a traffic management centre, all cars are considered equal most of the time. In reality, a grandmother taking a car to a pharmacy, a football star driving to a fitness centre or a child in a wheelchair being driven to school have very different reasons for using a car. If policymakers want to change human behaviour in the future, there is a need to know **why** people are travelling, **why** they are using a certain transport mode or a particular route and **why** they are travelling at a certain time. This will give insights into what factors can change behaviour (introducing congestion charges or separate cycle paths), what alternatives can be offered to single occupancy car usage (carpooling, Park and Ride, better public transport), and if separate policies are needed for specific groups (people with a disability, residents, commuters).

It is hard to track the reason why people are travelling, apart from using surveys. Some travel patterns can be extracted from ANPR systems (cars driving past the same point each working day), public transport smart card usage, shared mobility usage or patterns in licence plates checked against a parking enforcement database.

More precise insights can be gathered by analysing the destinations of trips: kindergarten, gym, school, restaurant, theatre, etc. This information can be deducted from searches in route planners, but is mostly gathered through tracking phones, for example by analysing mobile data and comparing the end points of trips with nearby points of interest (like this company does). Another way is analysing floating vehicle data or aggregated information from activity tracking, navigation, Mobility-as-a-Service or insurance company apps. Even better information is obtained from apps that keep on tracking people once they get out of their car or park their bicycle (cf. "Mobile phone application location data" below)

System and infrastructure data

Urban mobility departments typically manage a number of roadside electronic systems, like traffic lights, parking meters, retractable poles, inductive loops, speed and ANPR cameras and bicycle or pedestrian counters. Generally, these systems can provide two kinds of information. Firstly, operational information on the functioning of the system itself. As the city counts on these systems to safely manage and enforce mobility, they should function correctly. A defective sensor at a traffic light can disrupt traffic – traffic lights not working at all can create unsafe conditions. Secondly, these systems can also give insights into mobility. Sensors to detect cars waiting at traffic lights give information on the number of cars passing by, parking meters can help measure the effect of new parking policies, and enforcement cameras can produce information on average speeds.

Such electronic equipment is generally capable of generating status messages and other data. All that is needed is a connection enabling the data from remote systems to be collected centrally. The equipment can be connected via fibre optic, Wi-Fi, analogue cable or a mobile connection. If this is not the case, it is sometimes possible to connect a 4G modem and read the data centrally. It is also important to ensure that whatever data are coming out of the systems can be read and understood by other software. For legacy systems, this can result in disproportionately high costs. For new systems, such compatibility can be enforced in tenders, contracts and agreements.

More and more cities – like Ghent – are starting to monitor these data streams using an internet orchestration platform like Microsoft Azure or IFTTT. Such a service constantly monitors whether the system is still sending data, and values are within an acceptable range. If this is not the case, triggers are activated that send messages to operators, or take action automatically.

In addition, cities and mobility departments have a number of IT or software systems to manage their internal processes: accounting systems, tools for geographic information systems (GIS), complaint management, infrastructure maintenance, enforcement and systems to manage permits for parking, low-emission zones, etc. Cities will also want to know if these systems are working properly and gain immediate insights into the data through a dashboard, for example, to follow up on parking revenue, the number of complaints received, the number of licence plates checked, etc. Therefore, compatibility with other systems and data extraction should be included in contracts and tenders for internal systems as well.

Service provisioning

Cities are not the only parties in charge of mobility on their territory. In some cases, urban authorities outsource tasks or give concessions for parking management/enforcement, public transport operations or even traffic management. In other cases, service operators like dockless bike sharing and free-floating electric scooter or car-sharing platforms operate independently from the local authority. In both cases, cities require operators to transfer data on the services they are providing in order to verify the performance of the operator or obtain a view of mobility patterns within the city. This is typically done through concession contracts (cf. chapter 3, and case study 5 in the EIB Technical Note on ITS Procurement for Urban Mobility) or mandatory data-sharing regulations (cf. section 2.6).

1.6 Data sources

This section provides an overview of data sources commonly used by cities to map and manage mobility flows. As specified, the data displayed in a number of mobility dashboards for cities are taken as a starting point. For some projects, these data sources have been described and documented and make a very interesting read for cities. This is the case for the MOMENTUM project, the TMaaS project and the Manyways dashboard.

The most relevant data sources for these dashboards have been identified. It is immediately noticeable that motorised vehicles generate more data, mainly due to the use of navigation services, on-board units and enforcement systems. However, this trend has been reversing for several years now, and increasingly more data are being collected about public transport and bicycles. This is partly due to the use of smartphone applications. Pedestrian data have long been a black spot in data acquisition, but this is also slowly changing.

Due to the nature of a dashboard, namely displaying variable information, the focus is on real-time, streaming or live data. After all, dashboard visitors want to see the current state of mobility at a glance, hence the information should be updated frequently.

Mobile phone application location data

While it was mainly navigation apps that initially started selling or providing data on mobility patterns (as the use of hardware-based navigation devices decreased), many other apps now also collect the location data of their users. Among them are apps where the user actively and deliberately tracks their activity using applications like Strava, Runkeeper or the Bike Data Project. There are also many apps that store locations in a manner that is less transparent to the user. A smartphone tracking industry is emerging, where app developers are attempting to monetise big datasets of millions of users. These data might also find many applications in urban mobility as advanced algorithms can define the transport mode used for a specific journey, or even predict it – platforms like Sentiance indicate their ability to forecast future journeys. This article by the New York Times shows that both the possibilities of these datasets and the risk of privacy violations are gigantic (cf. section 3.2.2 "Data privacy"). Currently, there are not many sources from which to buy such data, as they might be violating data privacy laws at large. Platforms like Datarade offer a number of mobile location datasets, of which some claim to be GDPR compliant.

In-vehicle data

Vehicles are being fitted with an ever-growing number of sensors that are increasingly connected to the car's main platform. The measurements from these sensors can be read by the car dealer during maintenance or sent in real time to the manufacturer. The data collected mainly provide information about the functioning of the car, but also reveal a lot about the context the car is driving in through sensors that collect data on rain, shocks, heavy braking, speed, etc., while the latest cars are also equipped with radar or cameras. Manufacturers are using such data for analytics and predictive maintenance, but also to try to monetise the data through platforms like Otonomo. Some of the data can lead to the creation of safety-related traffic information and fall under the EU ITS Directive (2010/40/EU) stipulating that the safety warnings based on such data should be provided free of charge (cf. chapter 3). The Data Task Force is bringing together different stakeholders to work on this. An extensive study on automotive data sharing ordered by the Norway Highway Administration contains a lot of information on the subject. Among others, key findings of the study are that the value of data increases exponentially, authorities are taking an active role and car manufacturers are increasingly positive about sharing data.

Bluetooth (or Wi-Fi) tracking

By installing Bluetooth sensors, devices that are Bluetooth enabled – like smartphones or car kits – can be detected. Some of these devices also broadcast a unique identifier. If multiple Bluetooth sensors are installed throughout a city, speeds (based on travel time between two sensors) and routes taken (the locations of sensors detecting a specific identifier) can be detected. The same principles can be used to track Wi-Fi signals instead of Bluetooth signals, for example from smartphones.

Ahmet Rasim Demirtaş, Parabol:

"Our installations in Turkey currently detect 8% to 30% of all vehicles, and this will only increase as 87% of new cars have in-car Bluetooth, enabled by default."

Bluetooth detection is considered a relatively low-cost solution that can be easily set up. It is suitable for defining travel times or the origin-destination information of cars between a limited number of points. The technology is unable to count the precise number of cars, as it only detects vehicles that are Bluetooth enabled. This can be solved by using another sensor in the measurement (such as an induction loop or a camera) at certain points and extrapolating the total number of vehicles. Bluetooth is also less suitable for collecting detailed information from Bluetooth-enabled phones to detect pedestrians or cyclists. Smartphones are increasingly designed to protect the privacy of the owner and to hide or frequently change the unique identifiers they broadcast.

Mobile/telco data

Mobile phones make connections to networks in order to make and receive calls and connect to the internet. These connections are logged by telecom (telco) operators, including in call detail records (CDR – read more about this on the website of the MoTiV project). Because there are so many smartphones and they are almost constantly connected to the internet, these log files provide good information on where their owners are located.

These data provide very good origin-destination information that in the past could only be collected through large-scale surveys on where people come from and go to. Nommon Solutions and Technologies indicates that these data hold even more detailed information, such as on the purpose of the trips (work, gym, culture, events) by matching the destinations with points of interest, the frequency of trips, the usage of toll roads, some socio-demographic information like the gender and age category (obtained from the telco provider), and an income estimation (by averaging the income in the home neighbourhood). The mode of transport can be derived mainly on interurban trips; for intra-urban trips this is still hard using only telco data, and fusion with other datasets is recommended for obtaining reliable results.

Miguel Picornell, Nommon Solutions and Technologies:

"A very big survey on mobility behaviour by the Spanish government had a cost of a couple million euro a few years ago. The study has recently been repeated based on telco data with a budget that is around ten times lower."

Mobile data do not give insights into intensities or volumes and are less capable of accurately defining the transport mode used for intra-urban trips. The collection of the location of individuals raises privacy concerns, but experts interviewed indicate that the data are always aggregated, so no personal information can be extracted (cf. chapter 3).

This article shows how the data collection works and includes some examples from Slovakia. Acquiring and processing such data is done by specialised companies, but some telecom providers like Proximus have built platforms where cities can perform their own analysis. Public authorities wanting to procure this kind of data can look at tender examples from Wales and Spain (in Spanish).

Inductive loops

After their introduction in the early 1960s, inductive loops were the most utilised sensor in traffic management systems for decades. They involve one or more turns of loop wire being placed under the road surface. When a big metal object like a vehicle passes over the surface, the electric effect is interpreted by a controller. Inductive loops can count vehicles or determine their speed (if the loops are placed in pairs). Other uses of loops include classifying vehicles, detecting congestion or serving as a sensor for steering traffic lights by detecting waiting cars or identifying a specific bus or trams.

The popularity of inductive loops has decreased due to the relatively high installation costs, the traffic disruptions caused by their installation or repairs, and the need to replace loops after roadworks. Inductive loops have in some cases been (partially) replaced with floating vehicle data, radars or other kinds of sensors, depending on the use case (cf. the National Road Traffic Data Portal case in section 2.7).

Floating vehicle data

A number of vehicles are equipped with devices that can transfer their real-time location, like navigation devices, on-board units and increasingly smartphone applications. These floating vehicle data (also known as floating car data or FCD) are collected, processed, aggregated and turned into speed information for road segments by companies like Inrix, Here, TomTom, Be-Mobile, Waze, Yandex and Google Maps. Some of these companies use the information to provide accurate routing information in their own products. Most of them also make this information commercially available in real time for traffic or event management or as historical data to see trends over time or analyse trends or policy impact.

Malin Stoldt, city of Gothenburg:

"We removed all ANPR cameras in 2016 and are now completely dependent on third-party data to collect average travel times. Initially a quality analysis was performed which concluded that third-party data were equivalent with ANPR in terms of quality. We are satisfied with the current level of service and the possibility to be able to change which roads we collect travel times from without moving or installing new roadside equipment."

For this study, a number of experts were interviewed, and cases studied involving floating vehicle data. This has led to some interesting insights. In the Netherlands and Sweden for example (see case descriptions in chapter 2), roadside hardware was (partly) replaced by procuring floating vehicle data to collect speeds. Tiffany Vlemmings from the National Road Traffic Data Portal indicates that the transition to procuring data had loads of benefits, but also had a profound impact on IT systems in their traffic control centres and made it difficult to make comparisons with data that were acquired through inductive loops, for example. Moreover, David Cunha from the city of Lisbon warns that although floating vehicle data are a good source for traffic analysis, they would not use such data over inductive loops for adaptive traffic control.

As no roadside equipment is needed, floating vehicle data can cover huge areas or road networks, like the entire country of the Netherlands. However, the quality depends on the number of individual cars that contribute to the dataset. In low-traffic neighbourhoods for example, there are generally not enough data points to provide representative speed information. This also goes for bigger roads at night. The Swedish road authorities indicate that they receive historical averaged data when the minimum real-time data threshold cannot be met. The experts interviewed from TomTom explained that in their datasets, the quality level of the data is indicated so that cities can see when historical data are used rather than real-time information. In the Netherlands, some roadside equipment remained in place to allow for data quality validation and to gather data on the exact volume of cars. This allows them to independently check quality, whereas in Gothenburg, they rely mainly on the quality reports from the supplier, except on the national roads where ANPR cameras are still in place. They estimate that approximately 5% of the 20 000 to 25 000 vehicles per segment per day should be connected to obtain a reasonably good picture of the traffic situation. Gothenburg indicates having stipulated in the tender the minimum level of real-time data required during peak hours; where this is not achieved, the system must indicate this. Following implementation, quality issues emerged from a different angle: as the supplier uses a different base map than the public authorities, road networks did not match. This can in theory be avoided by agreeing on the geographic coordinate system in advance or using map-agnostic location referencing systems like OpenLR or AGORA-C. However, Gothenburg uses OpenLR and manual adjustments were still needed.

Cameras and licence plate recognition

Some cities and police departments have been using cameras with a certain degree of intelligence built in to interpret licence plates, for example. Common use cases are enforcement of speed, urban access regulations like low-emission zones and increasingly parking access and control. This type of camera can collect a whole range of information, from simple vehicle counts to speeding, car colour, red light offences and vehicle occupancy. Furthermore, if the camera back office is connected to a number of registers, it is possible to identify the CO_2 emissions level of the car, the address of the owner, and whether the vehicle is blacklisted or has any open fines. These kinds of cameras are now also being developed specifically to detect cyclists and pedestrians.

If specific cameras do not have built-in intelligence, it is possible to connect their videostream to special software that can detect specific things like licence plates, or count vehicles, pedestrians or cyclists. Number plate recognition exists as open source software as well – this is also the case for software to count vehicles, bikes and pedestrians (cf. the Telraam case in section 2.8). The EIB Technical Note on ITS Procurement for Urban Mobility describes Brussels' procurement of licence plate recognition cameras to enforce its low-emission zone policy.

Ticketing systems and smart card data

Many transport operators around the world have installed automated fare collection systems using smart cards that enable passengers to identify or pay for a ride by "checking in" with the smart card at a station, stop or when boarding. In some countries like the Netherlands, passengers are also required to "check out" when alighting. The dataset with these records can give detailed information on the usage of public transport lines, along with the origin and destination of passengers. Making sense out of the data is explained in this paper assessing smart card data studies (conclusions: many of the studies do not take into account land use, sensitivity analysis is not applied, and in half of the studies the results are not validated). Some off-the-shelf solutions are available in the market to work with smart card data, like Cermoni or Biotron (the latter is used among other things to analyse public transport data in Bratislava).

Pedro Barradas, Armis:

"Transport providers a lot of the time don't know where demand is located and often don't serve the underserved – they are a part of the puzzle, so getting some of the data they are holding would help cities understand if their own offer is well managed."

Miguel Picornell assesses this as an undervalued data source: although cities and operators in theory have free access to the data, they are not using them to their full potential, like combining them with other datasets. Access to the data, however, depends on contracts and concessions, certainly when credit or debit cards are used to check in. Data from public transport usage could also be obtained from Mobility-as-a-Service providers. The Data4PT project supports urban authorities in applying (European) standards on exchanging public transport data.

Automatic vehicle location systems (public transport)

Public transport vehicles (buses, trams, shuttle buses) are generally equipped with a location tracking system. If they do not have such a tracker built in, these automatic vehicle location systems have become very affordable and easy to install. The systems are mainly used to inform passengers about punctuality of public transport, through apps or digital information signs at stops. Their data can also be used to adapt traffic lights or verify the service levels of public transport providers (EIB Technical Note on ITS Procurement for Urban Mobility holds a case on both the procurement of AVL systems and monitoring concessions through these data). The TMaaS Link dashboard cited above displays the real-time location of small shuttle buses and the OV zoeker dashboard gives an impressive overview of the real-time location of buses and trains in the Netherlands, including their punctuality.

Air quality

Air quality can be measured to monitor the effects of mobility, for example the exhausts of combustion engines. There are a number of different parameters that can be monitored, as some will be impacted more by urban transportation than others. This study goes into detail on monitoring the impact of sustainable urban mobility plans on background air quality in cities. Air quality data can be acquired through national environmental protection agencies, regional environmental organisations like Airparif for the Paris agglomeration or websites like the World Air Quality Index project that combines 15 000 measuring points in 132 countries, both from government and citizen-installed measurement stations. The site also stores historical data and offers an API.

Weather

As cited above with examples of impact on cycling behaviour and traffic, weather conditions like heat, frost, snow, rain and fog have a major influence on mobility. Cities can therefore store weather data to cross-check the impact of the elements on other mobility datasets. Weather data are relatively easy to obtain, store and retrieve as a service and can help interpret deviating numbers of cyclists, accidents or traffic jams. Services like OpenWeatherMap, Climacell and Accuweather all offer free data to some degree through APIs, and are also able to provide historical data going many years back.

Data collected by the public

As internet-of-things devices and sensors become accessible to larger groups of people, increasingly more members of the public are starting to collect their own data through sensors. Examples are:

- Luftdaten: people install their own air quality sensors
- Telraam: people stick a small camera to their window to monitor traffic
- Waze or OpenStreetMap: the map is created and updated by its users
- Mapillary: a crowdsourced version of Google Street View aiming to cover the whole world, not only streets (sold to Facebook in 2020)
- 311, a hotline service in the United States that evolved into a platform for citizen engagement

- Apps like FixMyStreet where people can report potholes, missing signs, defective traffic lights, illegal dumping, dangerous situations and more

Crowdsourcing as a data acquisition model is discussed in detail in chapter 2 featuring the case of Telraam and OpenStreetMap.

1.7 Budgets

Data projects can require quite high expenditure of public funds, for example when we look at the projects cited containing dashboard development. The Manyways project received €221 000 in funding, the TMaaS project €3.4 million, the Liège metropolitan area obtained a €188 000 grant and the city of Bergen secured about €1.14 million, of which €150 000 was spent on the mobility dashboard over a two-year period. According to the product owner from Bergen, most of the time and money is being spent on getting the data from the sources and transforming them, rather than on actually building the dashboard. This is an important insight confirmed by the Manyways and TMaaS projects, as these steps are part of almost any data project. Paris pays between €5 and €10 a month for average vehicle speed data per kilometre. The city of Lisbon indicates that building their data catalogue alone cost several hundred thousand euro, and sustaining their data ecosystem (cf. chapter 2) financially will be a challenge in the future.

Noam Maital, Waycare:

"There is an evolution from capital investments towards operational expenditure as we see more and more as-a-service solutions. Evaluating such a service makes it much easier to get insights into a possible return on investment than first making hardware-based investments for years."

What Lisbon is hinting at is that data projects are very rarely a one-off investment. As with all IT systems, they come with an operational cost that is far from negligible. Antwerp also advises taking into account operational costs: a €5 parking sensor comes with a totally different price tag if licence costs, maintenance, services, support, and breaking up streets and pavements are taken into account. Rio for example paid about $10 000 a month on integrating data from their traffic cameras into their own systems. The evolution towards services rather than capital investments also requires a different and cross-departmental budgeting mindset. For example, the road maintenance department of the Swedish road administration is currently paying for travel time data but holds the view that since there is no equipment, it should not be paying. Similarly, the planning department declines to pay on the basis that there is no system to pay for.

On the bright side, all companies interviewed indicate that it does not take huge budgets to get started. On the contrary, they are all convinced of the value they can add and even prefer to start small, see where the real benefits are for the city and scale up when the outcomes are positive. Most of the suppliers indicate that they can set up proof of concepts below the budgetary thresholds for direct award contracts. One supplier stated that they work for a village of only 800 inhabitants for a couple of thousand euro a year. Miguel Picornell remarks that it is also important for city administrations to understand the market and what factors influence the price. He quotes the competitiveness of the market, the number of times a provider can sell (access to) the same platform or service, the level of standardisation and how the data will be shared (private use, open source, etc.). The National Road Traffic Data Portal in The Netherlands confirms this: if a public authority wants to open up the data it procures, it can destroy the market for the supplier – and its competitors. In an interview for this study, TomTom highlighted the risk of allowing its customers to open up data: they do not want other potential customers to find the data on an open data portal. These kinds of contractual limitations are also seen in the contracts for floating vehicle data between the cities of Gothenburg and Paris and their suppliers.

Pedro Barradas, Armis:

"Cities are building up the foundations, and the value only comes at the level of the ceiling – we still have a long way to go. To get financial support, I recommend cities get involved in EU co-funded projects, not least for urban authorities to get the assurance their efforts are going in the right direction and are in line with the EU vision."

Moreover, cities can rely on funding from local, regional, national and EU institutions. For example, all the dashboards cited in this section have been built through subsidised projects. European project funding opportunities can be consulted here, and in many countries there are national agencies or consultancy firms that can help find the right kind of funding. André Ormond adds that cities should not only look for money when contacting development banks, but also for solutions. He states that these institutions can help build partnerships, foster social development, find opportunities and initiate collaborations with other cities.

Rob Roemers, STIB/MIVB:

"Private players often make the point that public sector data could generate multibillion euro industries when opened up but the moment we say 'ok, but could you make a contribution to the cost', we are quickly told not to overestimate the value of our data. It's a very black and white debate. There should be a give-and-take scenario where we both share a similar set of data with each other to optimise operations."

Finally, the study did not reveal any party that had made an extensive cost-benefit analysis for acquiring or using data. Most of the interviewees indicated that the underlying reasons for working in a data-centric way are technical and they see this as a next step, rather than a cheaper alternative. The only clear business case encountered is the Spanish Ministry of Transport using mobile data instead of large-scale household surveys for transport demand analysis – the results are available as open data here. In addition, the National Road Traffic Data Portal in the Netherlands indicates changing traffic flow data collection from inductive loops to floating vehicle data, because of the virtual elimination of almost all geographical limitations and a substantial reduction in costs. Moreover, the study did not encounter any cities that were selling data on a for-profit basis, but the findings from Copenhagen on setting up a data marketplace make an interesting read. Some in-kind data exchange partnerships were uncovered, which are described in more detail in section 2.2.4. Furthermore, some public agencies indicated a desire to explore data monetisation possibilities, including STIB/MIVB.

Chapter 2: Data acquisition models

2.1 Introduction to the acquisition models used

This chapter proposes a number of models for data acquisition which were derived from real-world data collection projects on urban mobility, a literature study and interviews with more than 30 experts from different parts of Europe and the Americas. An extensive list of initiatives and literature studied can be found in Annex III. Based on the information collected, a classification consisting of seven models is proposed in this document.

The division into seven categories is indicative and not conclusive. It is intended to represent the different perspectives and approaches to data acquisition and shed light on the alternatives to classic or straightforward procurement. This classification is one approach of many, for example the International Transport Forum uses six categories (including open data, which is not a core subject of this study). Each model described in this chapter is accompanied by two or three cases to illustrate the model in practice and comes with recommendations by the experts interviewed. The models described in this chapter are:

1. **Public procurement of data:** a public procurement procedure is used to buy data. In this document, a narrow definition is used to compare this model to others, especially the intermediaries model: this category discusses one-off or as-a-service procurement of a single-source dataset without advanced preprocessing by the vendor.

2. **Intermediaries – integrators, aggregators and marketplaces:** urban authorities in this model call upon a third party that offers services to (pre)process data, extract information, merge data sources or interconnect systems. Marketplaces for mobility data are included in this category.

3. **Financially compensated partnerships between the public and private sectors:** this model is an extension of straightforward public procurement – government and market parties collaborate and exchange on a deeper level. Such cooperation can be on a contractual basis or based on innovative procurement procedures.

4. **In-kind partnerships between the public and private sectors:** public authorities have a number of assets and exploit them in exchange for commercial data, or vice versa – a private company offers data in order to receive goodwill or data in return from cities.

5. **Mandatory data sharing:** in this model, urban authorities exercise their power to oblige service providers to share data in order for the latter to receive certain approvals or permissions to operate in a city.

6. **Collaborations between authorities:** cities work together with other (urban) authorities to exchange data, jointly procure data or build platforms, services or data standards.

7. **Crowdsourcing:** urban authorities collaborate with the public to collect data and information, check or improve data quality or even outsource some of their tasks to residents.

2.2 Public procurement of data

Definition

In this model, data are acquired through a straightforward procurement procedure. The urban authority procures data in the form of a dataset or a recurrent service from a service or data provider. This can also be an intermediary like a reseller or a consultant. The duration of the contract can vary depending on the use case. For traffic/travel analyses purposes, data are typically procured for a fixed period (e.g. two months) at different time intervals, for example before and after the implementation of a transport measure (as in the case of Ghent's traffic circulation plan) or at fixed intervals, for example the same month over a given number of years in order to measure flow changes (as described in the Paris traffic analysis case in this section). In the case of a recurrent service, a permanent data feed is needed to support an operational task, such as traffic monitoring and the provision of average travel times (cf. Gothenburg/Sweden case in this section). The data can be procured separately or obtained as a part of a broader contract, for example a public transport service obligation contract.

Application

Straightforward public procurement can be used to acquire data in cases where:

- There is good knowledge of what the market can offer, and the data needed is commonly available (from multiple vendors). If necessary, a market consultation is conducted (cf. the EIB Technical Note on ITS Procurement for Urban Mobility).
- The readily available platforms and data provisioning services on the market do not meet the specific needs of the urban authority, for example because the city's requirements differ from those of other public authorities.
- The city knows what data it needs and what the results of the tendering process should look like. This may be because a successful proof of concept has been undertaken, sample data have been evaluated or results from other projects or cities can be examined.
- The specifications of the data and/or services can be fixed in a tender before starting the procurement or project. There is little probability that the requirements will change throughout the project.
- The data needs to meet certain quality and availability levels and it is desirable to agree upon service level agreements in a contract (and penalties if those are not met). This may be because the data are used for critical processes within the city or the city is relying solely on this data source for this type of data.
- The city wants to be assured of a certain data format, a delivery frequency or certain privacy or security measures and wants those to be contractually obligated.
- Other possibilities to obtain the data have been considered, but these are not possible or desirable.

Procurement of floating vehicle data for analytical purposes by the city of Paris

The mobility department of the city of Paris has experience of procuring mobility data, specifically floating vehicle data and more recently positioning data generated by smartphone applications (cf. chapter 1) to support traffic flow and wider travel analysis (cf. section 2.3).

Data for traffic analysis: The traffic control centre, PC Lutèce, launched its first tender (administrative specifications – technical specifications) in 2016 for the delivery of (i) traffic speed data and (ii) origin-destination analysis, based on GNSS data. Mobile/telco data were excluded from the tender because they were not considered appropriate for capturing vehicle flows. The contract was awarded in 2016 to a consulting company and extended for a further two years in 2020. The company collects raw traffic flow data from public and private sector partners for the French road network and is specialised in traffic flow analysis. The tender specified the delivery of traffic speed data on the main Paris road network covering one month in 2017 and one month in 2018.

The traffic speed data are delivered raw to the city of Paris. The average traffic speed of a road segment is calculated every three to 15 minutes depending on the road category (three minutes for the higher capacity roads). A second type of speed data is also supplied: the actual speed of a vehicle at a given moment in time and the distance covered by that vehicle every 60 seconds. To be able to work effectively with these data has required the skills development of several traffic engineers within PC Lutèce.

Interviewed and consulted**: Mélanie Gidel* and Richard Nguyen**, city of Paris.*

Procurement of floating vehicle data in Gothenburg/Sweden

The procurement of floating vehicle data started in 2014 in Sweden, replacing the traditional method of generating travel times through ANPR cameras. The cities of Gothenburg and Malmö and the respective regional offices of the national road authority (Swedish Transport Administration) procured from one floating vehicle data provider whereas Stockholm held a contract with another. The aim was to achieve a more cost-effective way of gathering traffic data over a wider geographic area. A further motivation for Gothenburg was the need for improved traffic monitoring due to disruption from the implementation of the West Sweden Agreement, a massive transport infrastructure programme.

In 2018, the national road authority took the initiative to establish a national agreement with just one provider, leading to a joint procurement involving the three regional offices and their respective city authorities. There is a long history of strong cooperation between the cities of Gothenburg, Malmö and Stockholm and their respective regional roads office, which is evidenced by the joint traffic management strategies and control centres and the common traffic information service in Stockholm and Gothenburg.

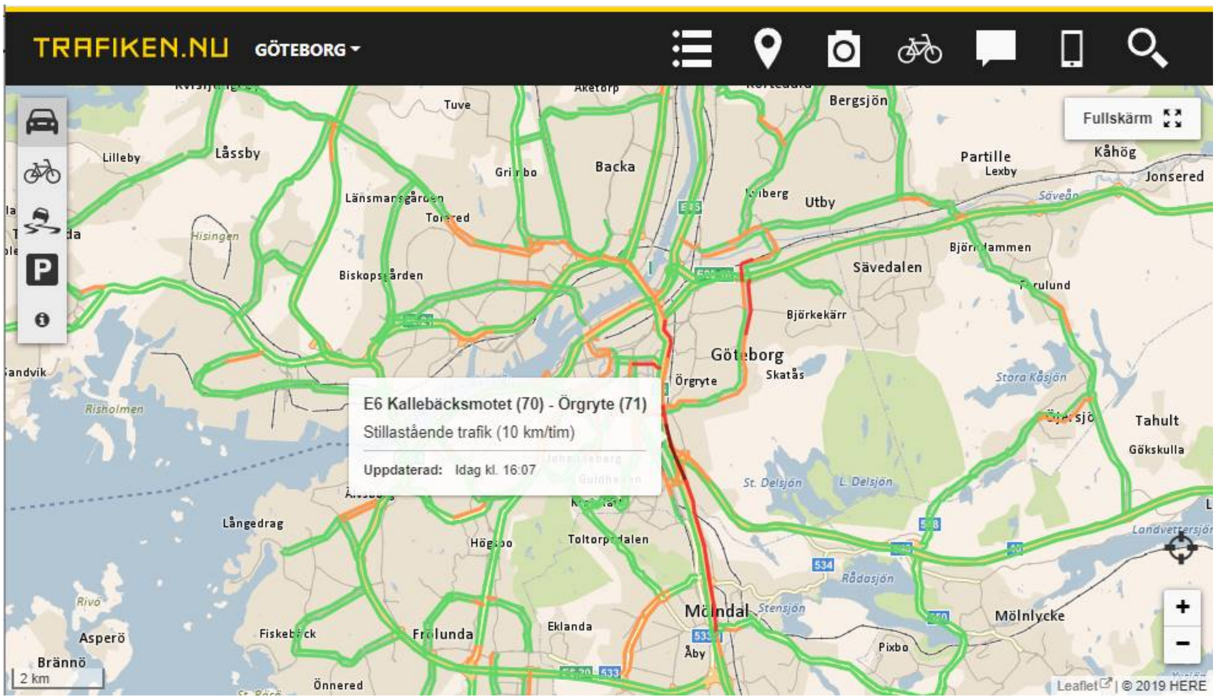
Given its technical expertise, the national road authority took the lead in defining the requirements through a dedicated tender group, with input from the respective cities. Each party defined the road network for which it wanted to receive travel times. This required careful consideration for financial reasons (more roads = more costs) and due to the limitations on extending road coverage during the contract itself (to take account of new needs). The legal department limited this to an additional 5-10%, beyond which a new tender would be required. The city of Gothenburg prioritised the roads carrying the most traffic where disturbances can heavily affect flow. The tender group did not consider free sources of floating vehicle data because it held the view that no free source of data could deliver the reliability and geographic coverage that was required. Nonetheless, traffic events coming from Waze are being used and Waze travel times may be added in the future as a complement.

Price was the main criterion for selecting the winning offer amongst the four bidding companies meeting all other tender conditions. Tender requirements included data source (GPS from smartphone or navigation system), delivery (frequency, information about proportion of real-time data and reliability of values, etc.), information about travel time calculation method, input source, volume of delivered data, etc. The tender document can be consulted here (and an appendix with map here). A two-year contract was offered with the possibility of two one-year extensions (2+1+1). While there is just one national agreement with the provider, for practical reasons there are four separate contracts: one for the Swedish Transport Administration's central planning department and one for each of the three regional offices (Gothenburg, Stockholm and Malmö). The cost is divided among the parties according to the kilometres of road requested by each party – there is a fixed price per kilometre. Each regional office shares the cost with the city authority according to the same principle. During the contract itself, it is possible for each party to change the roads for which it would like to receive average travel times, provided that the overall number of kilometres specified in the contract remains the same.

The travel time data have two main purposes: to inform drivers (by means of variable message signs and the travel information website trafiken.nu) and for traffic analysis and planning purposes. In addition, the traffic management centres also make use of such data for some operational matters. The contract does not stipulate what the average travel time data can be used for; however, it forbids the opening up of the data. Instead, the data received can be processed and combined with other data sources where available, and presented as average travel times on trafiken.nu.

The data are delivered every minute as a travel time value on short road segments. In the absence of sufficient real-time data, historic data are used. There are indicators showing the quality of data and the amount of real-time versus historic data for each segment. It is usually not possible to reach the minimum threshold for real-time data at night. Travel times are needed when traffic is most congested.

Interviewed: Malin Stoldt, city of Gothenburg, and Per-Olof Svensk, Swedish Transport Administration.



The Swedish dashboard showing traffic flows based on floating vehicle data. Source: Trafiken.nu.

Lessons learned and recommendations:

Aggregated vs. raw data: Based on its experience of acquiring and working with both raw and aggregated GPS data, Paris believes that it is easier for an administration to work with the latter and it is therefore increasingly leaning towards working with intermediaries in cases where aggregated data are needed, for example mobility flows/travel demand analyses (as described below in section 3.2). Working with raw GPS data requires specific skills and may give rise to privacy concerns (cf. section 3.2.2 “Data privacy”).

Data quality: Over the last four years of procuring floating vehicle data, road authorities in Sweden have learned the minimum level of real-time floating vehicle data that is required to deliver a realistic representation of the traffic situation. Minimum levels during peak times are stipulated in the contract. Where these cannot be reached, they can be compensated by historic data; however, the supplier must declare this. Other cities like Ghent have similarly required their providers of floating vehicle data to indicate quality levels (cf. section 1.6 “Data sources”).

Long-term budgeting: It is important to define a budget source for the long-term procurement of data as it does not sit neatly in one department. For instance, while the road maintenance department of the Swedish road administration has been financing the data procured for national roads to date (in place of buying, installing and maintaining equipment), it now holds the view that since there is no equipment, it should not be paying. Similarly, the planning department declines to pay on the basis that there is no system to pay for. Financial sustainability and a steady funding source/budget line are however critical as the supplier is the sole source of data for this specific function, as ANPR cameras have been removed.

Benefits of joint procurement. There are clear benefits in joining forces to procure data (services). Economies of scale have not been measured in financial terms. However, there have been other benefits. Gothenburg has always procured floating vehicle data jointly (with one authority for its 2016-2018 contract and three other authorities for the 2018-2020 contract). Teaming up with a (larger) authority can be beneficial in terms of skills capitalisation. Gothenburg was able to rely heavily on the technical expertise within the Swedish road administration for the tendering process and for data integration, which was technically challenging due to the different map base used by the supplier. Furthermore, during the 2016-2018 period with multiple suppliers, the administration had to work with different data flows and data quality and manage different supplier relationships. Since the national road authority now handles the data, it is far easier to deal with just one supplier. Finally, a joint approach enables travel times across national and city roads managed by different road authorities to be calculated.

Per-Olof Svensk, Swedish Transport Administration:

“As we have joint traffic centres in Stockholm and Gothenburg, some kind of joint procurement is needed and it is a big advantage to be able to use one and the same process for handling the data.”

Data reuse: Given its commercial value, sharing the procured (raw) data is mostly subject to strict terms. Open data in particular are prohibited, as is the case in both the Paris and Gothenburg examples described above. However, Paris may share the average speed data procured with other parties (e.g. consultants) for the purpose of traffic analyses only and according to pre-defined terms. In Sweden, the contract allows for the dissemination of aggregated data via the road authority’s traffic information website. Experts warn that requiring the possibility to share raw data (as opposed to aggregated data, as is the case in Sweden) in tenders or contracts can significantly increase the data price (see section 1.6 “Data sources” and section 2.9 under the heading “Market disruption”).

2.3 Intermediaries: integrators, aggregators and marketplaces

In the mobility data-sharing landscape, intermediaries are gaining importance due to the increasing number of relations needed between internal datasets, information systems of partners and data coming from external parties or vendors. In this study, these intermediaries are divided into three categories.

First of all, there are integrators. They focus on connecting or integrating data systems. For example, putting in place the interfaces for service providers or concessionaires to deliver contractually required data to a city (cf. section 2.6 "Mandatory data sharing"). Integrators can build direct data connections between systems and use their own platforms and services to transform and exchange data or implement proprietary or open source third-party solutions.

A second type of intermediary distinguished here are so-called aggregators. These companies collect data from one or multiple sources, process them and create meaningful information out of the processed data. This means that in most cases, the city or end customer does not need to perform advanced data processing and can obtain immediate insights into the length of traffic jams, air quality levels over time or modal split, for example. As a consequence, in some cases the intelligence and the raw data reside with the aggregator (cf. section 3.2.1 "Contracts and licensing").

A third category of intermediaries that is emerging are marketplaces for mobility data. These digital platforms bring together parties offering data and potential customers. Examples are Otonomo, which is a marketplace for in-vehicle data from multiple car manufacturers, the HERE marketplace and neutral server, Mobito and MDM, a national mobility marketplace set up by the German government.

In this document, a distinction is made between straightforward procurement (section 2.2) and acquiring data through intermediaries (this section). The integrators and marketplaces discussed here are clearly different from procuring datasets. For aggregators, the distinction might be vaguer. The line has been drawn as follows: the more traditional way of procurement is described in section 2.2, where a public authority procures a dataset from a supplier (being a data generator or a traditional intermediary like a reseller). In this section, an upcoming market segment of aggregators is highlighted that do more than just selling datasets. A number of such commercial intermediaries were interviewed for this study, which go beyond just offering data. Rather, they merge or fuse data from different sources (and modes), perform data orchestration and add an additional layer of (artificial) intelligence to the data, including visualisations, dashboards or services:

- Nommon Solutions and Technologies: processes and analyses raw/processed data from multiple suppliers (raw telco data, public transport smart cards, GPS data, etc.) and builds APIs and information systems on top of the data.
- Opendatasoft: positioned at the end of the information chain and focuses mainly on making open data easily available through its solutions, including APIs and visualisations on top of the data.
- Parabol: works with partners to collect data, has agreements with data providers like telco companies but mainly focuses on extracting (artificial) intelligence from the data.
- TomTom: apart from being a data provider itself, it also merges data from other sources and mainly works with intermediaries or resellers to turn data into information for cities.
- Vianova: helps cities integrate data, including through MDS, and builds dashboards on top of this (cf. case below in this section).
- Waycare: gathers data from multiple sources and offers a platform-as-a-service for traffic management centres, including predictive analytics.

In addition to these commercial parties, this chapter describes a number of initiatives in which authorities create such intermediary platforms, like the TMaaS project or the Dutch bike dashboard.

Intermediaries can be called upon to acquire data, process data or deliver data services in cases where:

- Cities do not want to get into agreements with a number of data suppliers, but prefer to have one contract with an intermediary who has pre-existing arrangements with several data sources.
- The city does not have the required skills or capacities to perform in-depth data processing, the work to be done is not part of the core competencies of an urban administration and/or the department does not see the need to expand its capacities in this field.
- The city's business case for using an intermediary's services (including maintenance and support) is beneficial compared to doing the data acquisition/processing itself.
- Working with the products, services or platforms of an intermediary is expected to be more beneficial than jointly building such a solution with a consortium of public authorities.
- There could be certain security or privacy issues processing the data and the city wants to mitigate these kinds of risks.
- The city has obtained access to data or tools through another acquisition model (e.g. the Waze for Cities partnership – cf. section 2.5) but is not able to process the data and market companies have developed tools or services to perform such tasks.
- Results need to be obtained fast, which can be done by using existing services or tools.

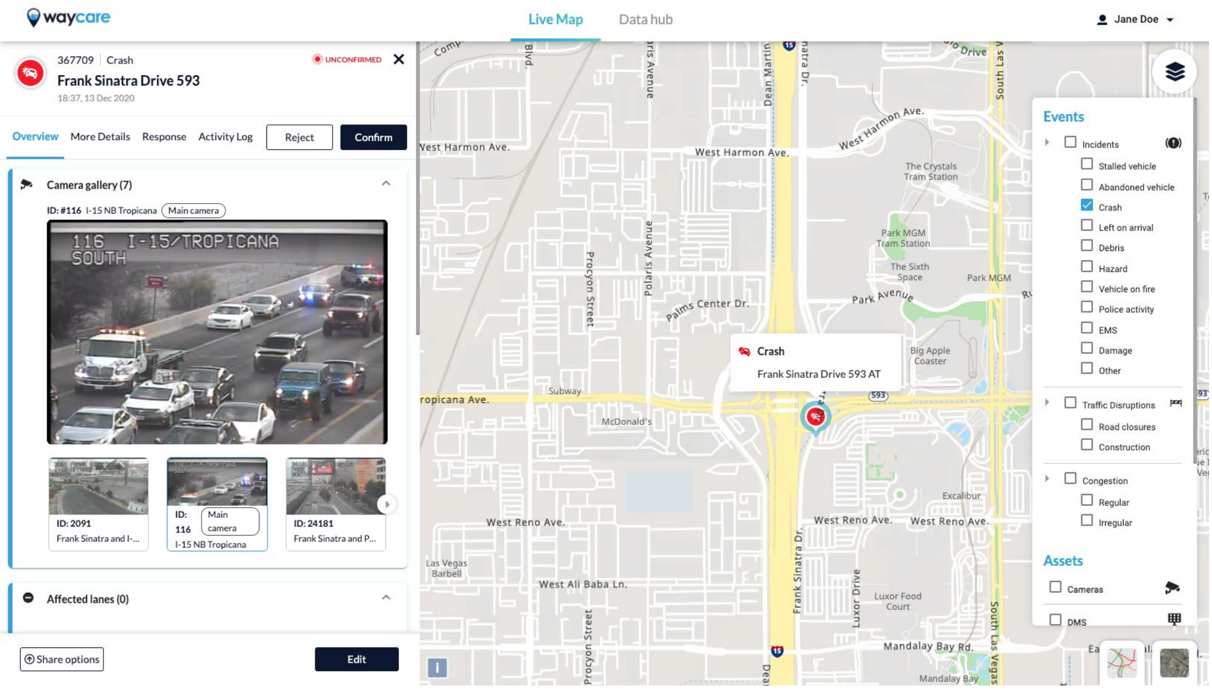
Case: Traffic Management as a Service (TMaaS)

The city of Ghent in Belgium, together with TomTom and a number of partners, is endeavouring to implement the Traffic Management as a Service concept through the EU-funded TMaaS project. This involves building a central traffic management platform that collects multimodal data from a wide variety of data sources. It is sometimes described as a "Spotify for mobility information". Instead of cities having to install hardware to collect data or purchase data separately, this is done centrally by the platform. Cities can then subscribe to the TMaaS platform as a service. This gives cities' mobility experts access to all data the platform holds on their territory, tools to analyse these data and means of communication to notify residents. Members of the public can register for free on the platform and indicate their preferences, for example their home location and daily route to work. Whenever something unexpected happens on their route, they receive personalised notifications via social media. For example, a cancelled train, traffic jams, an interrupted bike lane, etc. Finally, residents can also reply to these messages and inform the platform and the city's traffic managers about new events or incorrect information received.

The main advantage of this platform cited is that cities can easily register and start working with mobility data, regardless of their previous experience or investments. Economies of scale makes the TMaaS platform a lot cheaper than conventional traffic control centres. Ghent University has been studying the business case for the concept: *"Once the platform is built and has a lot of subscribers, let's say 200+ cities using it, we believe the annual subscription could be as low as €10 000 a year, after a one-off entry cost of also about €10 000."* The figure is low compared to the costs of a city building such a platform on its own (the project budget is €4.3 million) and in comparison to the costs for a traffic management centre, estimated between $111 000 and $3.1 million **annually** by the US Department of Transportation, depending on the size of the transportation management centre (TMC). However, it should be noted that the TMaaS platform is currently far more limited than a traditional traffic control centre and does not yet offer integration with legacy systems or traffic light controllers.

Waycare, a company offering a concept similar to Traffic Management as a Service, was also interviewed for this study. Waycare is a specialised traffic management platform, mainly active in the United States. The CEO, Noam Maital, cites predictive analysis (enforcement teams can be positioned at locations with a high risk of accidents) and integration with the police reducing intervention times by up to nine minutes as the biggest assets of the platform.

Interviewed: Peter Mechant and Timo Latruwe, Ghent University, Stephanie Leonard and Jeroen Brouwer, TomTom, and Noam Maital, Waycare.



An online traffic management platform showing a detected incident. Source: Waycare.

Case: Intermediaries between public authorities and shared mobility operators (the Dutch bicycle sharing dashboard and Vianova)

The rapid growth in micromobility services has created a need for data sharing to enable public authorities to monitor the usage of these services and to verify that operating rules are abided by. Gathering such data directly from a growing number of operators can be challenging for public authorities on many levels (organisational, technical and legal). A niche market of commercial data aggregators has emerged to respond to this challenge, which has largely built up around the MDS specification (cf. section 2.6) developed for free-floating mobility services. In the Netherlands, a government-funded project has led to the creation of a data aggregation platform for shared bike and moped services. The public agency CROW, which supports local authorities in transport policy and planning, took the initiative to build this platform to preclude the development of platforms by each local authority.

If required, these platforms can act as an intermediary between micromobility operators and the local authority, i.e. they can collect and process the data, meaning that there is no direct technical (data sharing) connection between the city and the operators. There is nonetheless a legal connection between the city (as data controller) and the platform (as data processor), meaning that the city is legally responsible for complying with the GDPR. (cf. section 3.2.2). The platform providers have agreements with the main mobility companies operating in a given area. A commercial platform company interviewed for this study, Vianova, confirmed that it has data-sharing agreements in place with most of the main electric scooter operators in Europe and reaches out to smaller micromobility operators on a case-by-case basis. The Dutch project currently involves eight providers of shared bikes and mopeds, comprising some 13 000 vehicles.

According to the data aggregator platform providers, working with them can benefit public authorities and mobility operators alike. Firstly, a relationship of trust has been established between mobility operators and platform companies, which benefits all parties. Secondly, public authorities are not required to ingest raw data from the operators, for which they do not always have the in-house capacity or in which they do not want to invest if they only require a report with the results of introducing a new policy measure or on the usage patterns of a new service. In such cases, using aggregated data can also circumvent perceived privacy issues, as the urban authority would not receive information that can lead back to the identification of individuals (cf. section 3.3.2). Thirdly, the platform tends to adopt de facto data standards (such as MDS developed by cities or GBFS developed by companies) thereby averting the multiplication of bespoke city specifications, which can be burdensome for mobility operators. Fourthly, an intermediary removes the need for operators and authorities to set up data-sharing mechanisms with multiple parties, saving time and effort on both sides. Vianova is moving towards standard contracts and data-sharing agreements in its dealings with operators and customers, which are adapted as required. Finally, since many cities do not have data-sharing requirements as part of their licensing conditions, data sharing is based on goodwill only, which a city alone would be hard pressed to deliver. Data aggregation platforms can overcome this challenge due to the agreements they have in place with mobility operators.

The data collected by commercial data aggregators (like Vianova, Blue Systems, Populus and Remix) is often delivered to customers in the form of customised dashboards (cf. section 1.4) that provide periodic monitoring of services (day-to-day, hourly, etc.) and historical insights, all of which can be viewed on a map. The platform may also allow the creation of geo-fenced areas (e.g. restricted/mandatory parking areas or vehicle caps in designated zones) and enable public authorities to verify operators' compliance with (local) regulations. In the future, the platform may also offer real-time orchestration whereby public authorities can communicate real-time demand/supply changes to operators.

On the Dutch shared bike/moped platform, a public authority can access mobility data from its own jurisdiction whereas an operator can see information related to its own services anywhere. The types of information a city can view include how long a bike has been unused, the start and end point of a trip and average utilisation rates. This information can help cities verify the usage rates of shared vehicles, which is usually a key criterion of permits issued to operating companies. All information is currently accessible via a digital map (based on OpenStreetMap – cf. section 2.8 in this chapter). CROW is working on a standard report that would enable municipalities to receive a monthly activity rate report at the click of a button.

The Dutch project is planned to finish at the end of 2020 but is likely to be extended. Building a business model is one of the remaining project tasks. CROW sees itself as a trusted third party but is uncertain whether it should continue or leave dashboard services to the market.

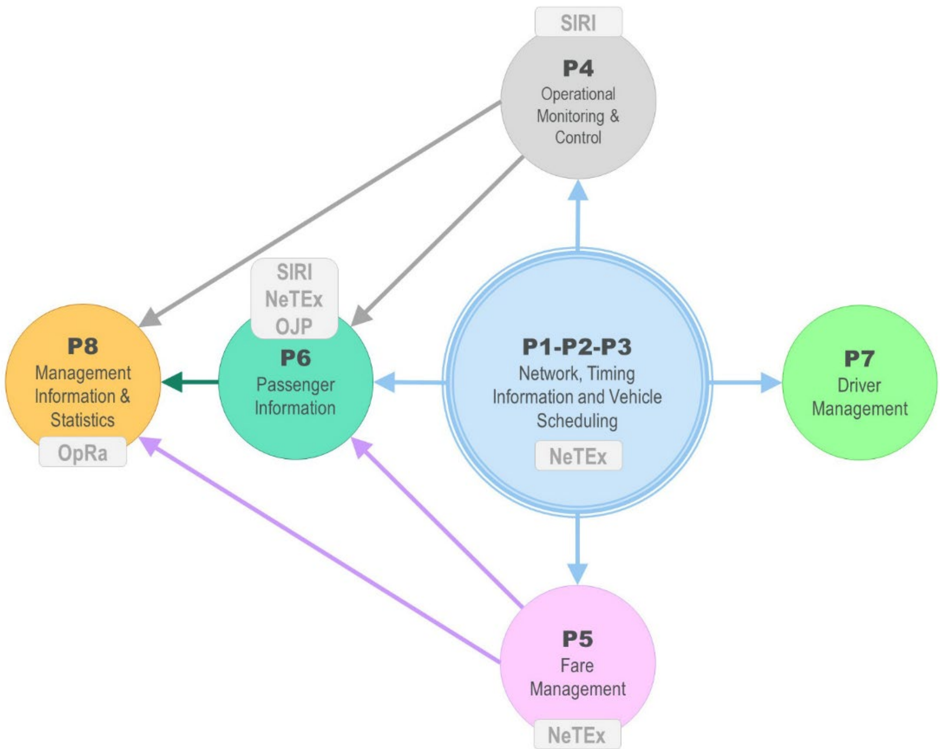
Interviewed: Otto van Boggelen, Kennisplatform CROW, and Thibault Castagne, Vianova.

Case: The Data4PT project

Data4PT is a project in the field of public transport that offers free and professional guidance on data integration and standardisation. The initiative was started by UITP (the International Association of Public Transport), ITxPT (a not-for-profit member association concerning IT systems for public transport) and a number of EU Member States: Austria, Croatia, Czech Republic, France, Denmark, Italy, Portugal, Slovenia and Sweden. Data4PT supports the development and implementation of data exchange standards and models to advance data-sharing practices in the public transport sector. Its objective is to enable EU-wide multimodal travel information services and in this way contribute to a seamless door-to-door travel ecosystem across Europe that covers all mobility services.

While the project is mainly aimed at fulfilling the needs of multimodal travel information service providers, the coordinators stress that they support a range of stakeholders like transport operators, Member States, software companies and also cities. Their main objective is to support the adoption of data standards to secure interoperability in the field of public transport. Data4PT supports Transmodel, NeTEx and SIRI implementations (data exchange standards for public transport, defined by the TC 278 WG 3 European standardisation body) and has, for example, been assisting organisations on many levels publishing public transport data on National Access Points (cf. chapter 3). In addition, Data4PT provides support to public transport stakeholders by answering technical requests, training experts and reporting feedback to standardisation bodies. The project's experts can help cities analyse what can be supported by a private company and what initiatives like Data4PT can assist with. The project can be contacted here.

Interviewed: Kasia Bourée and Emmanuel de Verdalle. Consulted: Christophe Duquesne and Anders Selling.



The Data4PT project propagates using the Transmodel data model, which covers most of the Public Transport domains. Source: Transmodel.

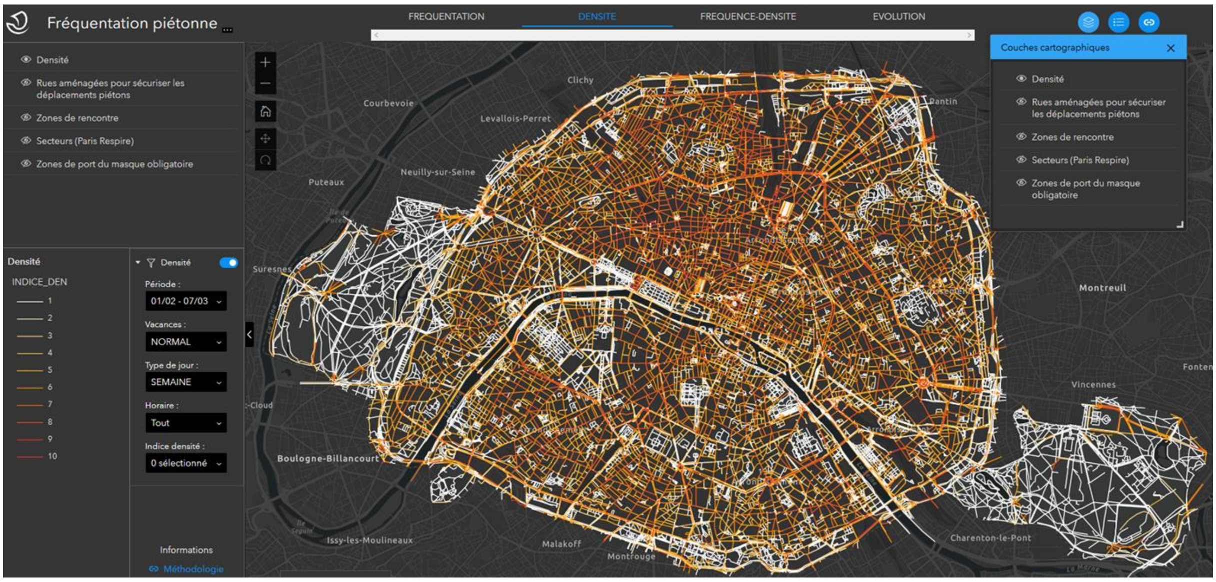
Case: Procurement of data for pedestrian flow analysis in Paris

In addition to the acquisition of floating vehicle data (cf. section 2.2), Paris' mobility department is also procuring data/information for other types of analyses. In a recent example, the city of Paris procured aggregated data (as opposed to raw data) related to pedestrian movements during the first COVID-19 lockdown in 2020. Due to the specific skill set required to process raw data (cf. floating vehicle data case in section 2.2), Paris decided to procure aggregated data from a consultancy specialised in transport data analysis. The information was based on GPS data from smartphones. Due to the urgency of gathering these data, the tender budget was limited to €40 000 (the maximum budget for a direct award). The contract was made up of two parts: the first part covered pedestrian density information for a specified period during and after the lockdown and the second part covered ad hoc analyses.

The positive pedestrian flow tender experience has generated interest within the mobility department for other types of analyses based on positioning data from phones. Paris plans to award a contract in 2021 aiming to understand the origin and destination of travellers, notably to determine the needs for new segregated bus lanes. Positioning data can effectively complement traditional surveys. The city of Paris is now considering a much larger cross-departmental tender for analyses based on this type of data, which would be managed by a horizontal team delivering support services across the municipality.

The city of Paris is leaning towards the procurement of information (or insights) rather than raw GPS data in specific cases (mobility flow/travel demand analyses) due to the specific skill set required to work with raw data. By way of example, processing raw phone-based location data requires certain skills to be able to differentiate between modes. Furthermore, procuring information can also circumvent privacy issues that may be attached to raw data.

Interviewed: Mélanie Gidel, city of Paris.



Dashboard showing the pedestrian density in Paris. Source: City of Paris/Kisio.

Lessons learned and recommendations:

Value creation: Calling upon an intermediary service or platform should not be done in isolation from the end-users, as the TMaaS project experienced. It is suggested to clearly define the end-user upfront to avoid a lack of focus later on: are the intermediary's services mainly aimed at the city administration to help increase mobility for inhabitants or should they be directed at and create direct value for the public? CROW is yet to start work on the business model. First, it must address the more fundamental question – should it be delivering such a platform/dashboard service or should this be left to the market.

Timo Latruwe, Ghent University:

"In TMaaS, we started off only from the city's vision without refining it along the way. I would recommend instead starting off small, iterating with end-users and expanding the aspects that create value for them."

Understand the business model and market: If you are using functionalities on top of data that a lot of other cities or customers need, it could be beneficial to use services from an intermediary, as the costs of development, new functionalities and updates should be shared amongst all the customers (as TMaaS and Waycare promise). However, if market research shows that the economy of scale is not reflected in the price of the service, cities could consider getting organised and jointly procure or build tools or services. Examples can be found in section 2.6 below.

Data ownership: Some of the intermediaries interviewed indicate that the city has access to the data they gather, process and store, even when the collaboration comes to an end. This should be part of any contract if the city sees any future use for the data or might consider moving them to another provider later on. However, the city of Antwerp warns (cf. chapter 3) that there is a need to specify how these data will be delivered and what they will look like (including accuracy and unique identifiers to be used). They have experienced receiving a large "data dump" that was unusable to them.

Online article at Bloomberg CityLab:

"Still, Reynolds, LADOT's general manager, recognises the potential for profit interests to interfere when cities depend on consultants and private sector expertise for technological needs, even as they're standing up to other companies like Uber."

Dependency: As for any supplier, cities might want to avoid depending too much on one intermediary for their integrations, data acquisition or strategy definition. In the example of Los Angeles, concerns have been raised about the conflicting interests of consultants and integrators involved in the creation of MDS. The ownership of the Mobility Data Specification was later transferred to a not-for-profit, the Open Mobility Foundation.

Ahmat Demirtaş, Parabol:

"We notice that cities prefer buying bundled solutions, and we support this, because the challenges are complex and no one can solve it all by themselves. This creates however the potential danger of creating vendor lock-ins. They can be overcome by using standard protocols and defining access in good agreements."

Data standards: These are at the core of this model. The Dutch shared bike dashboard's reason for existence is the GBFS+ standard, making it possible to collect and process data from multiple operators. The TMaaS project indicates that custom integrations are very time-consuming and states that the adoption and value creation of their concept depends on the availability of standardised data and interfaces, among other things. This also goes for commercial parties – the more standards are used, the broader their applications on top of data can be used without modifications, and thus at lower cost. It is recommended to exchange data with intermediaries in applicable standards where possible.

Slowly moving towards standards: Emmanuel de Verdalle from the Data4PT project explains there are different paths to standardisation. Some cities like Lyon have started with custom data exchange formats, taking a stepwise approach. He sees some advantages: it is an opportunity to gather local stakeholders and raise awareness of what can be achieved, before proceeding to a full standardisation effort. It is possible to convert existing datasets to standard formats later on using conversion tools. However, several experiences of Data4PT and ITxPT have highlighted the limitations of data conversion. One reason is because not all datasets describe all data features (e.g. a "column" or "field" in your initial dataset). For example, in public transportation many datasets lack descriptions of functionalities for people with disabilities. Therefore, de Verdalle recommends starting off with a native standard to enable sustainable interoperability. He stresses that a standard can be used with a limited number of features, and more can be added later whenever the organisation is ready or there is a need to expand functionalities or data exchange.

Access to an ecosystem: Compared to acquiring a specific dataset or service, intermediaries mostly collect and merge data from multiple sources, giving authorities access to data from an entire ecosystem. For example, Waycare indicates that they are the only traffic management platform that currently works with the police in the United States. The benefits of working with an intermediary can be better access, better coverage, and more diversified and less biased data sources. The big downside is that there may be heavy dependency on the intermediary.

Kasia Bourée, Data4PT:

"For the time being, unfortunately no certification organisations exist to check whether implementations of intermediaries are standard-compliant. This is left to contractual agreement with the implementing organisation."

Validations of implementations and integrations: Several software companies are skilled in implementing data standards or specifications in the urban mobility space, like DATEX II, MDS, SIRI or GBFS (cf. table in section 3.2.1 with an overview of data standards). However, in the majority of fields there are no certification organisations to check whether an implementation is standard-compliant for the time being. Therefore, the Data4PT project advises to carefully check and validate that implementations by integrators are standard-compliant. Third parties can be relied upon to perform such validations, and the Data4PT project itself can also provide support through tools and advice in the field of public transport data standardisation.

2.4 Financially compensated partnerships between public and private sectors

In some cases, public authorities partner up with the private sector and collaborate in a more intensive way than just procuring data (as described in section 2.2.1). For example, team members work together closely and can obtain a certain degree of insight into the internal affairs of the other partners. The partnerships can take various forms, like a collaboration contract or a European Innovation Partnership allowing for the combination of research and procurement. An example is the Mobilidata project, where the government and market jointly define the requirements for innovative cooperative ITS applications. In the case below where STIB/MIVB, Brussels' public transport operator, would like to obtain people flows, they used the pre-commercial procurement model for ICT-based solutions, which "enables the public sector to steer the development of new solutions directly towards its needs." Urban authorities can also partner up with the market to exploit mutual benefits. In the case described below, private companies resell on-street parking tickets through their applications and network and are compensated with a percentage of the income generated.

Financially compensated partnerships can be a good option to acquire data in cases where:

- No readily available solutions exist in the market or tailor-made solutions are needed.
- The requirements or objectives for the data project are not 100% clear and market companies are willing to define the specifications along the way together with the city.
- The specific technology or area is too advanced for or not part of the core mission of a city administration, and the city chooses to partly outsource this service or task.
- There is a risk that certain developments will not pay off, and market companies are willing to take on this risk in return for a share of the income.
- Certain activities are left to the market and some intellectual property rights remain vested with the technology vendors, but the city wants to remain in control of some core infrastructure and how these services are implemented and/or protect the rights of its residents.

Case: Parking rights database

Technology is evolving rapidly, and public authorities do not always have the capacity or the intention to get involved in (niche) areas where there is sufficient market supply. The parking register is an example in which the city makes innovation possible without having to develop anything itself. The city makes digital infrastructure available, in this case a database and interfaces to integrate into this database.

The city itself only sells on-street parking tickets through parking machines and leaves the sale of mobile or SMS tickets to the market. Any provider of mobile parking applications can enter into an agreement with the public authority to sell on-street parking tickets. These authorised providers can sell tickets by passing on their customers' licence plate numbers to the parking rights database, along with the start and end time of the parking session. The register automatically calculates the rate to be paid, and the supplier and the public authority settle the tickets at fixed times. The supplier can keep a percentage of the parking revenue as compensation for the services offered. In Ghent, for example, this amount is 5% of the ticket price, including VAT (cf. the regulations document – in Dutch).

This model is advantageous for the end-user: they can use the latest technology and a provider of their choice, at no extra cost. The city can use the existing user base and points of sale of various providers, does not need to invest in the development of apps, and is not taking any risks in terms of investing in non-profitable app development or outdated technology. There is a level playing field for the service providers, and with a popular app or service they can get a piece of the high parking revenues of cities (for Amsterdam, this is estimated to be €321 million in 2020). Moreover, such a platform also facilitates innovation because, for example, automatic systems such as scanning vehicles (as is the case in Brussels) can easily be deployed to perform parking controls using the interfaces of the parking database.

In Belgium, Ghent has a municipal parking register (the regulations make a very interesting read – in Dutch), while the cities of Brussels and Antwerp have their own versions. In addition, there is a commercially available parking register used by smaller cities in Belgium (Belgian Parking Register – in Dutch). The Netherlands has a National Parking Register, which avoids every city having to set up infrastructure and agreements with service providers.

This case was validated by the city of Ghent.

Case: Procuring R&D of solutions for monitoring people flow in Brussels – Muntstroom

Under the name Muntstroom, four Brussels public partners – STIB/MIVB (public transport operator), CIRB (IT department), Brussels Mobility and Parking Brussels – are working together with market parties and end-users to develop and test an IoT solution for people flow monitoring. This represents Brussels' first data-driven innovation procurement project.

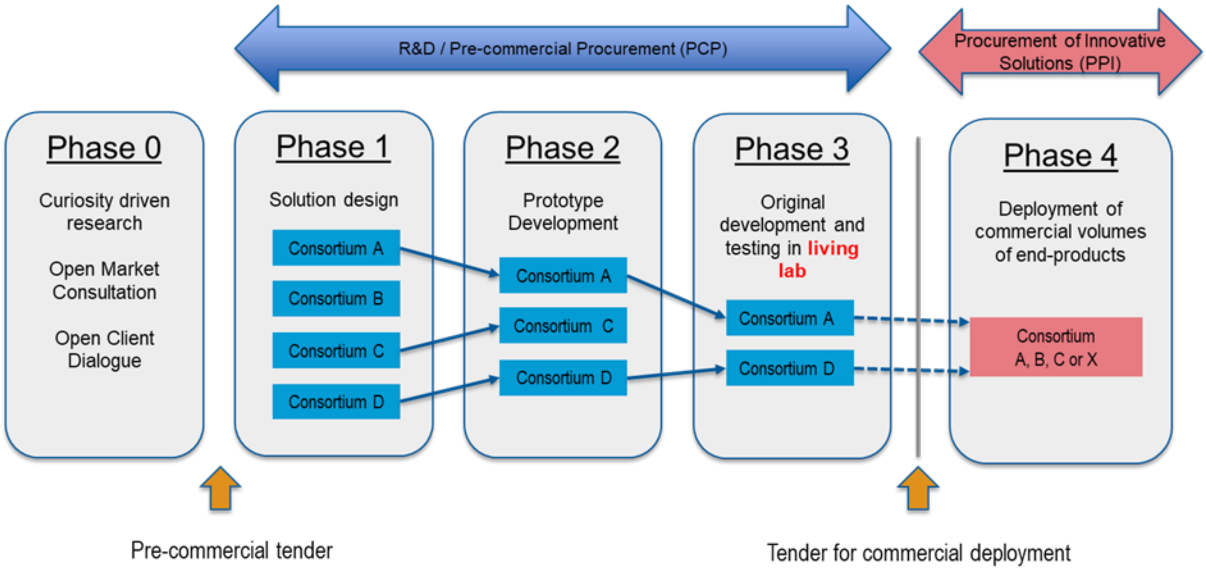
The public buyers chose to procure innovation/R&D instead of a product, because the solution sought does not yet exist in the market. The desired end-to-end solution should for example support complex analytics use cases and also data sharing between the public sector and private companies, while complying with the EU General Data Protection Regulation. To allow the market to propose state-of-the-art solutions, the innovation procurement approach deliberately avoids diving into technology and data definition. After a thorough preparation in the first phase, the development and testing of the solution is therefore a relatively low effort exercise for the public buyers.

Muntstroom is currently (December 2020) in phase 0 of pre-commercial procurement (PCP). An extensive assessment was made of the public buyers' needs, followed by research on existing solutions and patents. Via two market consultations, market parties were invited to give feedback on the feasibility of the ideas and the proposed planning. An open market consultation was organised to allow potential technology providers to respond to the technical challenges and an open client dialogue focused on market parties that were interested in using the future open and shared data for mobility apps, for example. The public buyers will fine-tune their requirements with the feedback from the market parties.

After the preparation in PCP phase 0, three PCP phases will follow. Following a pre-commercial tender, four or five consortia will be invited to design the solution. The results will be evaluated to determine which three consortia will continue to the next phase for the development of the prototype. Following an evaluation of that phase, two consortia will most likely be invited to test the developed prototype in a real-life environment (living lab). In this case, the area on, around and under Brussels' Place de la Monnaie (i.e. outdoor and indoor).

Given the risk averse nature of the Brussels' public authorities, had a traditional procurement approach been followed, the market would have no doubt proposed traditional solutions and there would have been no opportunity to engage with a wide range of different stakeholders. STIB/MIVB recognises that it is not only the end result that will bring the benefit; the project offers a great means for bringing people together to co-create a supported data framework and to build trust.

Interviewed: Rick Meynen, STIB/MIVB, and Stephan Corvers, Corvers Procurement Services BV.



The pre-commercial procurement process of the Muntstroom project in Brussels. Source: STIB / MIVB.

Lessons learned and recommendations:

Agreements: Good agreements are key and should set out clearly what the responsibilities are and what the financial compensation is in return. In some cases, cities can use existing contracts and agreements to start from, like the contract mentioned for Ghent or the agreement between the city of Brussels and the parking agency on a parking rights database.

Engagement: The importance of reaching out to players within the data ecosystem (IoT companies, potential data clients, public authorities, citizens and other data experts) is critical to defining the tender requirements in the procurement preparatory phase. As Brussels learned, this might be intensive and could require more effort than the actual tender process itself.

Stephan Corvers, Corvers Procurement Services BV:

"The PCP process in principle takes about 1.5-2 years but that misses out the initial research phase. Before starting a procurement, you need to know what you want – this is the needs identification and assessment phase. This requires a lot of desk research and understanding of what is available on the market."

Insights into internal processes: the experience of the TMaaS (cf. section 2.3) shows that the closer you collaborate with data providers, the more they will potentially want to have a say in what you do and especially do not do with their data in your joint project. By contrast, when procuring data, less accountability is required towards the supplier as long as the contractual agreements are obeyed.

2.5 In-kind partnerships between the public and private sectors

In some partnerships, urban or public authorities and private companies have found a way to exchange data or work together in a partnership without the need for financial compensation. This indicates that they both exchange items that are of approximately equal importance to both parties. This can be a bilateral exchange of data or data provisioning in exchange for less tangible benefits, such as reputation improvements, press attention or improved access to citizens or customers.

In-kind partnerships can be a good option to acquire data in cases where:

- The in-kind data are a cheap alternative to similar data that are commercially available.
- No hard guarantees are needed on data quality or availability and the data are not used for critical processes within the city administration.
- The usefulness of a certain type of data for the city's strategy is unclear or the outcome of working with such data is not certain and a low-cost proof of concept can clarify this.
- It is advantageous for a city to associate its reputation with that of a particular partner (even apart from any data exchange).
- The city is aware of what it can offer potential partners and what the implications are of engaging in a partnership with a specific partner.

Case: Waze for Cities

The Waze for Cities programme, originally called the Waze Connected Citizens Program, started in 2014 and was launched with ten partner cities from around the world. In 2021, more than 1 800 cities and other public sector entities are already part of the programme. Connected cities have access to the anonymous and aggregated data collected by the millions of users of the Waze app: travel times, road works, road closures and notifications from Waze users like accidents, potholes, missing signs and dangerous obstacles on the road. Waze, part of Google since 2013, offers free Google Cloud storage to connected cities, because over the years it became clear that cities cannot always cope with the large amounts of data. On top of the data storage, cities have access to a free dashboard as described in this Venturebeat article and to additional tools like BigQuery and Data Studio that enable partners to analyse the data.

In addition, the collaboration ensures a closer relationship with the local Waze community: local volunteers who keep the map up to date. Through such collaboration, adjustments to roads (such as the introduction of circulation plans or low-emission zones), events or new speed limits can be quickly implemented in the Waze map.

In return, cities are asked to share their data with Waze on roadworks and road closures through an automated API or more manual processes. This gives the navigation app insight into future road conditions – information that is not always easy to obtain without city involvement. In addition, a joint press release is drawn up about the collaboration, meaning that Waze enjoys extra publicity in its partner cities.

Finally, the partnership also goes further than just exchanging data. Occasionally, conferences take place where the partners come together and share experiences. The cities can indicate which specific policies or problems there are in their region, which leads to adjustments in the Waze app (such as the presence of high-occupancy vehicle (HOV) lanes, low-emission zones or the location of snow ploughs). In addition, new features or developments are also tested, such as Waze Beacons for tunnels or the Waze Carpool service.

Interviewed: André Ormond from Ormond Consultoria e Treinamento and one of the first ten Waze Connected Cities Program partners in in his former role as Traffic Engineering Director at the city of Rio de Janeiro. The case was validated by the Waze for Cities programme.

Case: A framework for multilateral mobility service data sharing (the creation of TOMP and CDS-M standards)

In the Netherlands, the public and private sectors have come together to jointly develop a standard for sharing data between transport operators (public and private) and MaaS providers. The impetus for a common data specification originally came from bike-sharing companies needing to develop interoperability by standardising the access to their bicycles for increased customer choice. This was required by the largest municipalities to obtain a permit to operate. Under the auspices of the Dutch Ministry of Infrastructure and Water Management, an open working group was set up to develop the technical interface. It involved over 20 different organisations, including bike and car-sharing companies and all major public transport operators. The outcome of this process, TOMP-API, is now implemented and used by several transport operators and MaaS providers in Europe.

The TOMP-API (Transport Operator to MaaS Provider - Application Programming Interface) is a standardised and technical interface between MaaS providers and transport operators. It allows all participating companies to communicate about planning, booking, execution, support, general information and payment of multimodal, end-user specific trips. After a development phase of nearly two years, in the summer of 2020 version 1.0 of the API was launched. Using the TOMP-API enhances the interoperability between parties in the MaaS ecosystem. A transport operator will be able to communicate with MaaS providers in the same manner, and vice versa for MaaS providers communicating with transport operators. The API enables any mobility operator to describe its asset offer (shared bike/moped), i.e. the real-time availability of a service/vehicle in a given region. TOMP works on a trip leg basis meaning that a trip comprising three different legs (e.g. shared bike to the station/train/bus) will require three API calls to different transport operators. Each mobility operator can then decide whether to offer a trip based on availability. The TOMP-API data flow is two-way, allowing any mobility operator to remain informed about its own vehicle that is planned, booked and paid through a MaaS provider.

A second specification has emerged besides the TOMP framework to enable operators to communicate their assets (vehicles) to cities, as opposed to MaaS providers. The need for a business-to-government (B2G) specification to complement TOMP's business-to-business (B2B) model has emerged to enable cities to monitor a growing market of shared mobility service providers, which are increasingly the subject of city regulation. The cities decided against adopting the widely used MDS specification (cf. case in section 2.7 below), because it did not meet their data needs and there were also concerns about data privacy.

The city of Amsterdam initially took the lead on discussions with the operators before it was joined by the G5 network of the five largest Dutch cities. The G5 established a working group tasked with building a data specification, CDS-M (City Data Standard - Mobility). CDS-M allows one-way communication of data about shared mobility assets, i.e. from transport operator to city authority; however, there are plans to develop a city authority to transport operator feature.

The development of CDS-M is still ongoing; the first implementation is expected to happen this year. The CDS-M working group overlaps with the TOMP-API in terms of members, but is independent of TOMP in order to keep focus on the business-to-government (B2G) aspect. CDS-M will become open source and the working group will be open to anyone. How the development of the specification will be financed is an open question.

Interviewed: Edoardo Felici, European Commission, Bon Bakermans, Dutch Ministry of Infrastructure and Water Management, and Tijs de Kler, city of Amsterdam.

Lessons learned and recommendations

Data dependency: In-kind partnerships (and data acquisition in general) can lead to dependency on the data supplier and create difficulties if the data supplier changes its business model, i.e. if it starts to charge for data or underlying services (as happened to Google Maps and created some disquiet). It therefore might not be in the interests of a city to rely exclusively on data through this type of arrangement.

In-kind vs. procurement: In-kind partnerships can compete with commercial services, requiring the latter to clearly differentiate their offerings, as TomTom confirms: *"If you want a viewer for traffic data, use Waze. But cities want an integrated solution, embedded in existing solutions they use, and that is something we can offer.*"

Shop around and check the small print: If a decision is taken to acquire data from third parties, it is worth checking whether there are free-of-charge data sources available. If this turns out to be the case, it would be useful to reach out to other cities also using these data to gather their views and experiences. It may also be necessary to check with (i) the local or regional data suppliers and be mindful of their arguments of possible market distortion; and (ii) the internal legal department to go through the terms and conditions of the in-kind partnership – there's no such thing as a free lunch. If the contract states that the city should provide data in exchange, it should at least be thoroughly considered to make these data accessible to other parties as well and avoid proprietary data formats, keeping in mind the good governance principles of equal treatment and creating level playing fields.

Big vs. small cities: The technical skills required to integrate data from an in-kind partnership such as Waze may also be dissuasive for smaller cities, where data skills are generally less developed. Smaller cities might also be less interesting for data providers to engage in partnerships with, which can be avoided by collaborating with other municipalities.

André Ormond, Ormond Consultoria e Treinamento:
"Over 800 cities have joined the Waze for Cities programme now, but I estimate that no more than 50 are actually using the data to their full potential because of a lack of expertise."

Need for capacity: Getting data is one thing, being able to process or feed them into existing systems is another. André Ormond states that a lot of cities are receiving Waze data, but many do not have the expertise to store or process them. Waze tries to accommodate this by offering free storage of the data and facilitating code sharing between cities, but the capacity gap between a tech giant and most municipalities remains huge (cf. section 3.3.1).

2.6 Mandatory data sharing

Cities have a lot to offer other parties, not least the power to allow or deny certain activities on their territory, which they are increasingly exercising. Cities have the authority to regulate certain services, for example taxis, parking, public transport and new shared mobility services (such as shared bikes and electric scooters). On top of this, cities regulate land use and the use of public infrastructure. In order to organise this properly and to monitor service providers, an increasing number of cities are requesting data from these service providers as part of an agreement, a concession or a licence to operate. In this section the cases of Paris and Antwerp are described, but some of the 90+ cities adopting the Mobility Data Specification could be considered as well (the case is described below in section 2.7). These examples demonstrate the legislative power of local authorities and should be viewed separately from the EU regulations on data sharing, which are explained in section 3.4.

Mandatory data sharing can be a good option to acquire data in cases where:

- Licences, concessions or approvals need to be given for certain services such as public transport or operating physical devices like e-scooters and taxis.
- It is necessary to maintain or create a level playing field for service providers to gain access to the local mobility services market.
- There is a need to verify the operations of service providers, or there are indications that the stipulations in the licensing agreements are not being complied with.
- Data from service providers can provide insights on urban mobility flows that cannot be obtained otherwise.
- Operational data cannot be obtained in a voluntary manner.
- The city has the legal capacity to create good licensing agreements and has the ability to enforce the stipulations in those contracts or licences.
- The city has the capacity to process and understand the operational data from service providers or is able to outsource this task.

Case: The mandatory sharing of micromobility data in Paris

Like many other cities around the world, Paris has adopted regulations (2019 – P16391) regarding the operation of free-floating vehicles (typically e-scooters and bikes) in the public space. These regulations include parking restrictions, an annual fee (depending on fleet size) and the provision of operational data. The data elements (format, transfer, types, storage, etc.) are described in an annex to the regulation. Currently, operators are required to transmit a predefined set of data (mainly vehicle ID, position, type and activity) to the city of Paris every three hours through an API in a format specified by the city of Paris. This format, called SIVU, is a simple model which was designed to respond to the data needs of the city. SIVU is available on the open source development platform GitHub. The data are fed into an internal dashboard that allows the city to monitor the fleet, analyse parking and verify whether the annual fee has been paid.

The city of Paris would like to move towards collecting data more frequently, which would then allow it to gain more detailed insights on the usage of these services and trip patterns. Until the city's IT server is fully operational, Paris' open data portal, which offers both data publication and collection functions, is offering an intermediate solution for ingesting data. Free-floating data are therefore transmitted via API to the Opendatasoft platform where they are stored and analysed internally but not published. Paris accepts that a long-term IT solution will be needed and that it may shift towards more widely used data formats such as MDS or GBFS, which are also specified in the regulation. Discussions are underway with operators about GBFS and initial tests reveal a difference in interpretation of the specification. Paris is participating in conversations on GBFS moderated by the French National Access Point.

The annex to the free-floating regulation cited above specifies that no API request should contain personal data.

Interviewed: Mélanie Gidel, city of Paris.

Case: Antwerp policies on mandatory data sharing for service providers

The city of Antwerp has made enormous progress in the field of data collection in recent years, following heavy investments in GIS and a dedicated team documenting the city's infrastructure since the beginning of the millennium. In 2016, these static data were topped with more dynamic datasets, due to a need for information to communicate about major roadworks in and around the city and to nudge behavioural change. Initially, data were purchased for a route planner and a smart map, in order to communicate with the public and road users. The initial approach where only data were procured by the city evolved towards the development of a level playing field ecosystem. The city has created regulations and established a framework for service providers to operate and to open up data with the city. For its part, Antwerp shares data with service providers, so they can estimate the compatibility and profitability of their services with the Antwerp mobility ecosystem. This further builds upon the Antwerp Open Data strategy that aims at opening up data for internal and external stakeholders and providing transparency on policymaking and implementation.

The city chose not to impose the framework and regulations unilaterally, but instead opted for a process of bilateral consultation with service providers. This has resulted in two agreements that are currently being drafted (cf. box in section 3.2.1). On the one hand, there is a generic data licence that determines which data must be supplied and for which applications and for what reason the city can use such data. On the other hand, a service level agreement is drawn up that determines the minimum availability and quality of the data, what support is expected and how long the data must remain available.

Data sharing is obligatory for micromobility providers. Soon this will also be the case for car-sharing companies that want to operate in Antwerp. Car sharing providers will have to provide the real-time location and availability of vehicles and individual trip data, as micromobility operators do today. For Mobility-as-a-Service platforms, which do not own physical vehicles, it is harder to oblige data sharing as they do not need a licence to operate. Antwerp tackles this by obliging each mobility operator (deploying physical vehicles or devices) to integrate with at least two MaaS platforms active in the city.

The city tries to operate as independently as possible from data standards, using the NGSI framework. This means that the data can be delivered in the service provider's preferred format, as long as it is an OSLO-defined standard (Open Standards for Local Administrations in Flanders) or a commonly used format such as MDS (cf. infra) or GBFS (General Bikeshare Feed Specification). Antwerp indicates that at the moment, they have to set up an individual pipeline for each data provider, because of the huge diversity in and interpretations of data standards. Combining an individual pipeline with a format of choice facilitates the process for service providers to share data. In the future, they would prefer to work with a uniform data pipeline and stress that clear data standards and/or policy regulations are a necessity.

Antwerp has a clear purpose for the data collection. They are building what they call M4: their Multimodal Mobility Manager. It is a piece of software that monitors service providers' adherence to the rules agreed in the the data licence and the service level agreement. In addition, M4 collects insights into travel behaviour and allows for policy planning and monitoring. Finally, M4 will share detailed insights back to the service providers and aggregated data with the entire industry. In the future, Antwerp is also planning to make its policy available digitally (like the MDS policy API, cf. infra) so that providers can obtain immediate feedback on the extent to which their policies are complied with.

Interviewed: Stijn Vernaillen, MaaS expert at the city of Antwerp.

Lessons learned and recommendations:

Not applicable to all service providers: Cities may not force (mobility) service providers to share data in cases where the city is not the regulator and does not have the power to issue operating licences. The city of Antwerp for example indicated that the model is not applicable for getting data from Mobility-as-a-Service platforms, and TomTom – which does not operate physical mobility devices – stated having never experienced mandatory data regulations. The closest they got is the EU regulation on the mandatory sharing of safety-related information when collaborating with vehicle manufacturers.

International Transport Forum – Data Driven Transport Policy (p. 33):

"For obvious reasons, however, the power (to compel data) should be limited to those instances where the data are necessary to carry out a public policy mandate. Even in those instances, the scope of data required should be minimised to that which is just necessary to carry out the public policy mandate."

Proportionate and purposeful: Both the ITF study and case studies show that the data that are requested should be in proportion to both the efforts needed by the providers to supply them and the right to privacy of users. Proportionality is also one of the governing principles of EU policies. Furthermore, the city should be able to clearly define the purpose of all the data requested, rather than collecting data for vague future uses (like creating a ground level version of air traffic control, cited by Ellis & Associates in a plan to help Los Angeles anticipate autonomous vehicles and other futuristic modes).

Data quality: Just because regulatory data sharing has been included in operating licences does not mean that they will be followed. The city of Paris has encountered problems with the quality of some data due to differences in interpretation of the data standard between the city and the operators. In Brazil, high capacity operators are contractually obliged to share data, but André Ormond states that the obligations are not specific and therefore not effective. The cities interviewed recommend checking the data quality, enforcing the regulations and building proper infrastructure to process the data.

Expertise required: The International Transport Forum stresses the need for capacity to extract information from the data received: *"(…) simply requiring regulated parties to provide data may not be sufficient for authorities to extract usable information from that data. The particular skill sets to understand, format, clean, parse and analyse large, unstructured or differently structured and high velocity data are not typically found in the public sector."*

Jascha Franklin-Hodge, Open Mobility Foundation

"There are always companies that will fight regulation and have arguments against it, there is nothing special about it. There is a political aspect to it, like other things cities do and they should be prepared for this."

Privacy discussion: Although MDS has been the subject of public discussions around privacy in Los Angeles, in most of the 90 other cities it was implemented no such fierce discussions arose. When asked, Jascha Franklin-Hodge, Executive Director of the Open Mobility Foundation (governing MDS), said involving communication departments intensely would be overkill, as cities are used to companies opposing regulation. Brussels' transport operator, STIB/MIVB, acknowledges being involved in privacy discussions and sees a need for a shared understanding: "*The private sector does not want to share,* *arguing they don't want to share customer data. We need to reassure service providers that we are not interested in customer data, only in movement data."*

A lack of standards: The city of Antwerp indicates that it would prefer to see European standards for the (mandatory) exchange of data with suppliers of shared mobility services. It chose the FIWARE NGSI interfaces (also used by the city of Vienna) over a specification like MDS, because these include smart city functionalities that transcend the smart mobility domain. Due to this lack of standardisation, Paris built its own specifications, called SIVU. In any case, Antwerp recommends offering multiple interfaces and standards to service providers and using standard or format-agnostic systems, with the possibility to change between standards if needed.

Consultation: The city of Antwerp created the mandatory data-sharing regulations in consultation with the service providers, which led to them adopting a wait-and-see attitude. In Los Angeles, the enforcement of data-sharing regulations led to ride-hailing companies comparing the city to dystopian authorities featured in 1984 and Brave New World. Other service providers have chosen a different approach: Voi Scooters for example profiles itself as "pro-regulation", offering a data-sharing dashboard and feeding cities data through their APIs. While most other data acquisition models described follow the free market logic of supplier and customer/recipient or setting up partnerships, in this model one of the parties uses its legislative power to impose its will upon the others. This could lead to tensions and therefore the city of Antwerp recommends consulting the service providers to alleviate concerns in advance and avoid imposing requirements that are resource-intensive or extremely hard for the service provider to fulfil.

2.7 Collaboration between authorities

In a number of cases, public authorities have joined forces to reach a certain scale, increase their impact, have a stronger negotiating position or jointly establish relationships with service providers. Collaborations can take the form of setting up umbrella organisations or data expertise centres, building platforms or services together or jointly procuring data.

Collaboration between authorities can be a good option to acquire data in cases where:

- Municipalities are not large enough to engage in relationships with big organisations.
- A group of cities can be brought together where each urban authority has a certain speciality or field of expertise.
- Several municipalities or public authorities have joint objectives or similar plans for urban mobility.
- Higher-level authorities possess data that are interesting for cities or are legally obliged to collect/share this information (for example through National Access Points).
- Other organisations are bound to the same legislation or processes (like tendering), which makes it possible to easily exchange contracts or tenders.

Case: Open Mobility Foundation (OMF)/Mobility Data Specification (MDS)

A standardised tool for two-way data flow between free-floating mobility providers and public authorities has been developed by the Los Angeles transit authority. Called Mobility Data Specification (MDS), this open source tool is used by public bodies throughout the United States and beyond, including Europe. The uniqueness of this standard is that it was initiated and steered by a public authority, and public bodies have a strong role in its governance today, through the Open

Mobility Foundation. Mobility providers are also members of this foundation and contribute to the development of MDS.

In essence, MDS is a set of three standardised APIs. The “provider” API is implemented by providers to allow a public body to query (pull) historical views of operations. The “agency” API sits with the authority to allow mobility providers to notify (push) data about events (e.g. trip start/end or vehicle status). The third and most recent “policy” API is implemented by the authority to allow providers to access information about rules and regulations affecting their services, i.e. it is putting regulation into code.

MDS is not a take-it-or-leave-it framework. The vast majority of MDS city authorities (95%) have in fact adopted just the provider API to generate an historical overview of trips undertaken. Even this API does not need to be implemented in its entirety – it contains a number of optional fields. However, customising the API is discouraged otherwise it loses its “standard” value. Most cities using MDS have procured the services of a data aggregator/dashboard service provider. This type of intermediary service can reassure cities reluctant to ingest raw MDS data due to potential data privacy concerns (cf. section 3.2.2). The OMF website has a section dedicated to privacy, including a link to its new privacy guide. The tool kit and aggregator service provisions offer a lower entry barrier for smaller cities with fewer data skills.

Interviewed: Jascha Franklin-Hodge, Executive Director, Open Mobility Foundation.

Case: The National Road Traffic Data Portal in the Netherlands

Around 2007, the idea emerged in the Netherlands to build a National Road Traffic Data Portal (NDW). There was traffic data at the time, but it was neither standardised nor centralised and there were a lot of blind spots where data were missing. The various authorities in the country decided to work together on an umbrella structure to manage traffic data for the whole of the Netherlands. 19 partners joined forces: the national government, the 12 provinces, the four largest cities in the Netherlands and two metropolitan areas. Every municipality in the Netherlands can rely on the National Road Traffic Data Portal and is represented at the provincial level. The National Road Traffic Data Portal was initially set up as a five-year project to centralise all traffic data.

At the time, the dominant way to collect data was by installing roadside hardware or sensors like induction loops and cameras. The project therefore started with the preparation of a number of tenders procuring this technology for area-wide traffic measurements over a five-year period. The new NDW project took care of the procedures, administration and centralisation of the data. The various authorities were consulted about the best locations to take the measurements.

Because it is simply impossible to equip every road in the Netherlands with induction loops, the measurements in this first period were mainly limited to the main road network. This changed with the emergence of floating vehicle data (cf. chapter 1), which offer the possibility to purchase data about traffic in an entire country. NDW extensively investigated the possibilities and limitations of these new data and replaced a large part of the data collection with floating vehicle data in new tenders. Today, the National Road Traffic Data Portal has nationwide coverage of floating vehicle data that are cross-checked with measurements from loops and cameras.

In the beginning, there were many visions of the organisational form that the NDW should adopt: private or public? Part of the national government, a separate organisation or a project? And what tasks can and may such a national register take on without disrupting the market? NDW evolved from a project to an independent organisation within the government and today no one questions its usefulness or raison d'être, says Tiffany Vlemmings, Project Manager at NDW.

NDW has an annual budget of around €15 million that comes from partner contributions and tailor-made projects and solutions on request. In return, all cities and governments in the Netherlands have access to NDW's nationwide traffic data and can rely on its expertise regarding data. Authorities have a say in the strategy and concrete projects. NDW has also engaged in partnerships or collaborations with Copenhagen and Waze for Cities (cf. above) and is involved in (European) standardisation bodies on behalf of cities.

In the future, the National Road Traffic Data Portal will expand its data to soft modes too (it has already started tendering bicycle data) and work more closely with national registers on parking and public transport data, in order to be ready for large scale Mobility-as-a-Service implementations.

Interviewed: Tiffany Vlemmings, Project Manager Innovation at NDW.

Case: Lisbon's urban data-sharing platform

To improve operational efficiency, the municipality of Lisbon took the lead in setting up an urban data-sharing platform. Systems operated by different departments of the municipality and other public agencies within the city (e.g. parking) are connected to the platform via the FIWARE standard or bespoke interfaces for legacy systems. The platform, formally known as the Lisbon Urban Management and Intelligence Platform, currently has 340 datasets from 40 different systems, which makes it a very powerful tool for integrated operations.

Access to the platform is limited to city hall services and external public entities (e.g. police, civil protection, fire services), with different levels of access granted. A protocol has been established setting out data sharing, ownership and updating principles, which is vital for building trust. Initially, the main function of the platform was to support an integrated operations centre, which opened in 2020. Now, the platform is the backbone for many other data-driven initiatives, including open data, a mobility catalogue under development (to enhance the navigability of open data related to mobility) and the urban data laboratory (more below). Under the supervision of the civil protection department, the integrated operations centre enables an integrated and coordinated response for all services, namely the municipal police and fire departments, the mobility department, the public transport operator (CARRIS) and the municipal mobility company (EMEL).

The urban data lab is another initiative that depends on the platform. The city of Lisbon has established a partnership with 11 universities to achieve a win-win approach for finding data-driven solutions to real-life problems. By partnering with academic and research institutions, Lisbon is able to access this expertise while researchers are able to access real data that are not publicly available. To gain access to the platform, the rector of each institution must enter into a non-disclosure agreement with the city council. A relationship of trust is crucial.

The first edition of the urban data lab commenced in 2019. Each of the ten municipal departments was invited to bring a real problem to the table and the universities were tasked with finding a solution to these problems using data. Five challenges, including a few related to mobility, were retained, leading to the formation of some 40-50 teams competing to find a solution. If a solution is found, the respective institute will retain the intellectual property rights but Lisbon will be entitled to use the solution free of charge. The first challenge is still ongoing and includes a challenge related to mobility.

Interviewed and consulted**: David Cunha* and Vasco Mora**, city of Lisbon.*

Lessons learned and recommendations

Public collaboration while leaving room for the market: While all the cases described in this section are based on collaboration between public authorities, there is room for and sometimes a reliance on market parties to develop tools to enable data sharing and to generate insights. NDW has deliberately avoided building out-of-the-box tools and visualisations to leave commercial opportunities to companies which make a business out of interpreting and processing NDW data. This has been a difficult decision and is in fact an ongoing debate because smaller municipalities in particular would prefer NDW to develop a tool centrally that they all can use instead of them all paying separately for the same services. The developer of the Dutch bike-sharing aggregation platform/dashboard (case described in section 2.2.2) is also faced with the same question: should a public agency be developing this or should it be left to the market?

Miguel Picornell, Nommon Solutions and Technologies

"The Ministry of Transport in Spain is considering buying data for the whole country and sharing it with interested agents (such as cities, transport operators). One challenge here is to define the scope of the project, because each agent might need slightly different information and/or different levels of detail – it's an interesting, as well as complex, problem to solve."

Aligning requirements: Collaboration inevitably creates the need to align requirements between the public authorities involved. In the case of Sweden's procurement of floating vehicle data (cf. case in section 2.2), Gothenburg was happy to leave the definition of technical requirements to the national road administration, which had the required expertise. In other situations, alignment might not be as easy, as confirmed by Nommon Solutions and Technologies.

Getting the buy-in: City authorities are renowned for being siloed, with each department operating independently and very little data sharing happening across and even within departments, often as a result of isolated system development that started several decades ago – a full replacement is considered expensive (in financial, labour and time terms), inconvenient (loss of service quality) and even risky (potential loss of data). While some effort was needed to convince the Lisbon council department managers to dedicate some time and effort to connecting their systems to the data platform, now that the data are integrated, these same managers are starting to see the platform's true potential and are proactively working with it. Where collaborations across administrations are concerned, reaching an agreement is that much more difficult and requires extensive discussion, as witnessed in the NDW case. A key challenge involved getting agreement from cities to let go or outsource some of their responsibilities.

Subject of collaboration changing over time: The cases offer very different examples of ways in which public authorities can share data in a collaborative manner. They all emerged in response to a well-defined need and have since scaled up to a level that goes above and beyond that original need. For instance, NDW is expanding its initial data focus from motorised vehicles to active modes. OMF/MDS is extending from e-scooters to ride-hailing and kerb management. In Lisbon, the original purpose of the data-sharing platform has evolved from an integrated control centre to include other projects.

2.8 Crowdsourcing

The accessibility of connected low-cost sensors and computers (like Raspberry Pi and Arduino), together with the rise of affordable long-range communication networks (like 4G/5G, LoRa and SigFox) have made it easier for people to collect and publicly share data from their surroundings, which is called crowdsensing. Some examples are the LuftDaten network and the Hackair project that both collect air quality data, the Making Sense project that is deploying networks to measure nuclear radiation, the Bike Data Project that allows people to publish their cycling activity as open data, and the Telraam/WeCount projects (see case studies below). This does not mean that technology should always be involved, however. The CurieuzeNeuzen project provided air quality measuring tubes to 20 000 people who attached them to their windows and could send them back by mail afterwards. The Straatvinken project asks people to sit in front of their house once a year (all at the same time) to count the number of cars and cyclists passing by. The city of Ghent is implementing new traffic circulation plans in its suburbs and requesting members of the public to help monitor and evaluate the plans by counting traffic or installing an air quality sensor.

The concept of crowdsourcing goes a bit further than collecting data: it means "outsourcing" tasks to the "crowd". This can be at the initiative of authorities, like the FixMyStreet app where people can signal potholes or broken streetlights, the VanConnect tool to report illegal dumping in Vancouver, traffic incident forms by the police, or the numerous 311 hotlines in the United States that evolved into citizen engagement platforms. Such specific platforms are also commercially available, for example Citizenlab and Civocracy. Private companies can also leverage the wisdom of the crowd directly through their users, as Waze is doing. There are also privately owned platforms to crowdsource microjobs like Amazon's Mechanical Turk or Microworkers. Cities could use these to count pedestrians in video footage, transcribe licence plates or take surveys, although the platforms are heavily criticised for providing unregulated, sub-minimum wage work. Finally, not-for-profit organisations also outsource knowledge gathering and verification to the public, like Wikipedia or OpenStreetMap (see case study below).

Crowdsourcing can be a good option to acquire data in cases where:

- The city wants to leverage the knowledge and skills of its inhabitants.
- The urban authority intends to closely involve the public as a means of participation in projects that concern or impact them.
- The data does not feed into critical processes/no service level guarantees are needed.
- There is an open political climate for discussing data and information that might not only be showing the positive effects of mobility (policies).
- Municipalities have the skills and organisational capacity for large-scale citizen interaction.

Case: Citizen-led traffic counting by the Telraam/WeCount project

Telraam is a project by Transport & Mobility Leuven, a research company supporting policy decisions on transport, Mobiel21, a non-profit organisation for sustainable and safe mobility, and Waan.zin, a web development company. The organisations jointly submitted a proposal in response to the Belgian government's Smart Mobility Belgium call in 2018.

Their proposal was to allow people to actively participate in the collection of mobility data. In this way, residents become more aware of the mobility in their street and they can objectify their feelings. There are often conversations or discussions about mobility between residents and local authorities, but there are rarely objective data to support these discussions. Sometimes residents themselves make observations or counts, but these are not always taken into account by the city.

Sometimes the city also sends out a city official to make an observation, but they only witness a snapshot.

That is why during this project a do-it-yourself kit was created with which residents can count traffic in their street and map speeds for cars. The kit consists of a low-cost computer (a Raspberry Pi), a camera and image recognition software written by the project. The resident tapes the camera to a window and connects the system to the home Wi-Fi. The image recognition counts pedestrians, cyclists, cars, vans and trucks and measures vehicle speeds. The recognition is done in real time, meaning that no images are stored or sent to a central server, in order to respect the privacy of passers-by.

Telraam's website shows an impressive number of cameras (1 881 as of January 2021) already attached to people's windows. The dashboard indicates a few hotspots and some isolated sensors throughout Western Europe. This picture can be explained by the strategy followed. The project started in Leuven, Belgium, and a number of devices were available free of charge for residents. Then a number of other cities also recognised the potential of the project and opted for a paid collaboration. In addition, there are quite a few individuals and neighbourhood committees that have purchased the device themselves. At the moment, there are 15 cities and municipalities offering a counting device free of charge. Everyone else can buy the kit for €85.

Due to the success of the initiative, the European project WeCount with seven partners was funded by the European Union for €2 million. The project aims to provide cost-effective data for local authorities and empower citizens to take part in the political discourse in an evidence-led way. The project is completely based on open source, and all data collected by citizens is publicly available through the Telraam API.

This case was validated by the Telraam team.

A tiny computer and camera attached to the window allow citizens to measure traffic in their street. Source: Telraam.

Case: Collaborating with OpenStreetMap to acquire, share and validate data

There are a number of proprietary map systems where the data are owned by a company (even if the maps can be consulted freely and without charge). The best known example is Google Maps. In addition, there are open map systems where the content is in the public domain that can be used by everyone and can also be adapted by everyone. The best known is OpenStreetMap, which does for maps what Wikipedia did for encyclopedias. Most people have used these maps without knowing it. The most popular apps using OpenStreetMap are Strava, Moovit, Foursquare, Trip Advisor, Pokemon Go, Snapchat, Facebook and even Apple Maps.

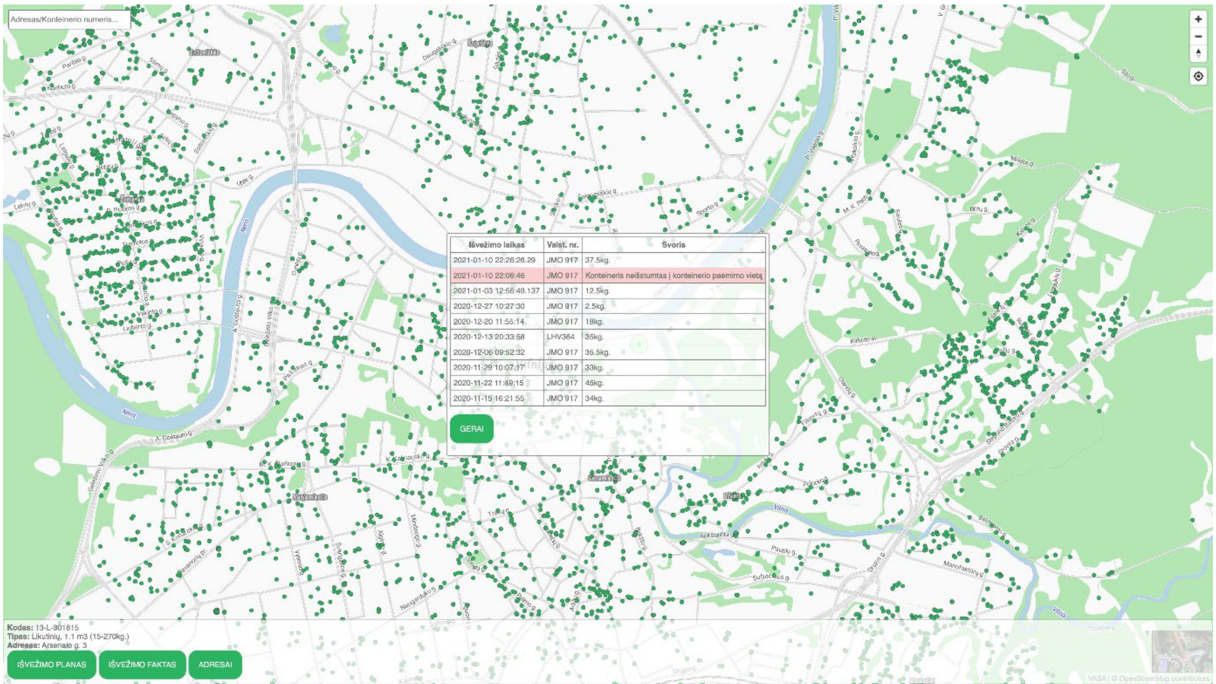
Once registered on openstreetmap.org you can make unlimited changes to the map. This also applies to cities and local authorities: in the event that certain errors are visible on the map, or the municipality has made adjustments to bicycle paths, directions, circulation, etc., the changes can easily be made to the maps.

Across the world, many public authorities use OpenStreetMap for internal applications or to share information with the public. For example, Vilnius publishes address data (house numbers), because these data are not publicly available in Lithuania. This way, anybody can easily look up an address. And because the maps are in the public domain, the information can also be used by other municipal agencies or anybody else. For example, in Vilnius, the waste management agency uses these data to manage household waste collection. People can check where to deposit their waste and can verify online if the information on kilograms of waste collected is correct. More (technical) info on the Vilnius-OpenStreetMap exchange can be found in this interview. Cities can also use OpenStreetMap data to build their own applications. Before implementing a new circulation plan, the city of Ghent created a route planner on a copy of the OpenStreetMap data that already reflected the future situation. This way, people could verify what the impact would be for them.

But if everyone can change the information just like that, won't the maps become a mess? The power of OpenStreetMap is just like Wikipedia and the community behind it. Almost all countries have a local OpenStreetMap community consisting of enthusiastic volunteers who take care of the quality of the data. In the example of Vilnius above, the community also plays a crucial role. Vilnius sends the information on to the OpenStreetMap community, and they compare the information with what is already on the map. All deviations are checked one by one and either adjusted if OpenStreetMap was wrong or sent back to the city if the source data were wrong. This means that the OpenStreetMap community not only publishes data for the city or others to reuse, it also improves data quality.

OpenStreetMap communities exist throughout Europe. You can reach them by looking up the contact details on the community overview page. These communities are fully staffed by volunteers, who work on the quality of the maps in their spare time. This means that they invest energy in projects they personally believe in, when they have time for it. If cities want to process large amounts of data or build applications with the OpenStreetMap data, they can call on a number of commercial software and services, which are listed on this website.

Interviewed: Tomas Straupis, OpenStreetMap volunteer in Lithuania.



OpenStreetMap as a base to manage household waste collection. Source: VASA - OpenStreetMap contributors.

Lessons learned and recommendations:

Tomas Straupis, OpenStreetMap community:

"One of the reasons for not opening data is the lack of data quality. Rather than trying to improve the data for years, municipalities should share the data, indicate quality concerns, and have the quality improved iteratively with help from the public. It will benefit cities as well."

Working with communities of volunteers: This requires a different approach to working with traditional suppliers. The experts working with you are not financially compensated, so it is good to know what motivates them and what you can do for them in return. The city of Ghent organised half-yearly meetups with the communities after working hours (because volunteers work during the day as well) with a sandwich and a drink to get to know the community members and see what the city could do for them. This is also happening in Vilnius: the OpenStreetMap community had a couple of open conversations with government agencies on what data they would like to work with.

Getting organised for public input: Crowdsourcing requires an entry point within the city organisation for people to contribute. Cities, suppliers and volunteers all stress this: people collecting information or helping to improve data quality need a short and responsive communication line to the experts or departments concerned. Volunteers sometimes have difficulties getting through to the correct people within administrations. André Ormond adds that this also extends to cities' social media channels. Processes need to be in place to make sure that feedback from Twitter, Facebook, hotlines and chatbots reach the right departments and feedback is provided to citizens. In addition, Opendatasoft indicates that if cities want broad feedback from their inhabitants, they need to create accessible information – not just sharing technical data files, but telling stories, sharing visualisations and putting the data into context.

Ensure data are used: If you share data with the public, as open data or otherwise, use the same data source for internal projects – the eating your own dogfood principle – as Tomas Straupis explains. It will ensure you get a good understanding of the data quality and availability. Creating different sources or copies of datasets can impede good collaboration, as are not working with the same source. Additionally, if public authorities encourage people to collect data, they should make use of these data, publicly share them and cite the source.

Who reaps the benefits/owns the data: If you work with a community of volunteers, invest in sharing data with organisations, support crowdsourcing or crowdsensing platforms or endorse specific systems, know who is benefiting from the efforts made and especially who owns the data. In both the Telraam and OpenStreetMap cases cited in this section, the data are published as open data under a Creative Commons licence. Most of the examples cited initiated by governments or not-for-profits publish the data in an open form or as public domain. This does not go for private companies, such as Waze, Strava or map edits sent to TomTom, Here or Google. The information collected becomes property of the company, regardless of whether it is publicly consultable or not.

De Standaard, article on Telraam (in Dutch):
"Finally, hard evidence that drivers are speeding here – From now on, residents in 330 Flemish streets will carefully monitor the speed of all cars passing by with a small camera. Citizens now have leverage for safer traffic."

Transparent performance of policies: Collecting crowdsensed data often happens without the involvement of the local authority. This means that the effects of policies and enforcement (or a lack thereof) will be clearly visible to the public: traffic jams, excessive volumes of cars, accidents, bad air quality, speeding, negative side effects of congestion charging or circulation plans, etc. For example, a number of smaller Flemish municipalities like Grimbergen, Oosterzele, Mortsel and Lochristi have been forced to defend their enforcement strategies after people started measuring traffic using a Telraam device. In a number of cases, the installation of such devices was actually facilitated by political (opposition) parties.

2.9 A comparison between the data acquisition models described

All of the seven models described have some advantages, drawbacks and contexts in which they are best used. This section relates the models to a number of parameters:

- **Control**: to what extent does the city have control over data specifications like data quality, availability, data formats, etc.?
- **Flexibility**: how much can the specifications, perspectives or outcomes of the data be changed throughout the process and to what extent can the data or services be tailored?
- **Budget**: what is the financial impact on the budget, and if the compensation is not financial, what is expected in return?
- **Organisational impact**: how heavily will the mobility (and other) departments be burdened engaging in this model?
- **Technical capacity**: what is on average the technical expertise needed to have or hire when using this model to acquire data?
- **Relationships:** does using this model impact relationships between the city and its stakeholders?
- **Implementation speed**: do some models tend to deliver faster results than others?
- **Legal capacity**: how much workload can using a certain model lay upon the legal department?
- **Market disruptions**: do some models have the potential to disrupt the local market for data?
- **Combining models**: some examples of combinations of the models described.

In the following overview table, the models used are compared against these factors indicating what capacity will be required from the city and roughly what the effect will be of using a certain model on the aspects listed above.

	Level of control	Level of flexibility	Impact on budget	Organisational impact	Technical capacity needed	Impact on relationships	Implementation speed	Impact on legal capacity	Market disruption
Public procurement	Orange	Green	Orange	Green	Yellow	Green	Yellow	Yellow	Green
Intermediaries	Yellow	Yellow	Orange	Green	Green	Green	Orange	Green	Green
Financial partnerships	Orange	Orange	Orange	Orange	Orange	Yellow	Green	Orange	Green
In-kind partnerships	Green	Green	Green	Green	Orange	Orange	Yellow	Yellow	Orange
Mandatory data sharing	Orange	Green	Green	Yellow	Orange	Orange	Yellow	Orange	Orange
Authority collaboration	Yellow	Yellow	Yellow	Orange	Yellow	Yellow	Green	Yellow	Yellow
Crowdsourcing	Green	Green	Green	Orange	Orange	Orange	Yellow	Green	Yellow

Green = low, yellow = medium, orange = high.

In the rest of this section, each of the models described in this chapter is compared according to the parameters listed above.

Control over the data quality and other specifications

- In procurement procedures there is more certainty on what the result of the procurement will be, as this is detailed in the tender. The descriptions will include requirements on quality and availability of the data or like in the Paris case, include specifications on data type, source, format, an implementation schedule and skills required. The city can expect to receive the data as agreed and impose penalties when the specifications are not met, as in the mandatory data-sharing model. This also goes for financial partnerships – the main difference here is that the specifications are mostly defined jointly with the supplier. The urban authority has a lot less control over the data specifications in models where data are mostly supplied on a best effort basis and without guarantees like in the in-kind partnership and citizen science models. In these cases, the city is in a weaker position to impose service levels than in models where it is paying and setting the contractual terms itself.

- Furthermore, when working with intermediaries, contractual terms and specifications apply. However, when using off-the-shelf services or products with pre-existing contracts, the city might have less of an impact on all the exact specifications than when writing its own tenders or contracts. In terms of ownership and transparency, cities might have less control over the data and algorithms used by intermediaries, as these are sometimes part of the proprietary processes or tools of the supplier or appear as a black box to the city. This is much less the case in models that are built around partnerships or collaborations.

- Crowdsensing can come with some specific challenges in terms of data quality (sometimes there is no centralised verification of whether the sensor is correctly positioned or calibrated) and data availability over longer periods of time (people need to keep sensors up and running, which requires a continuous effort; for example, the grey bars on the Telraam map mostly indicate sensors that need user intervention).

Flexibility

- In a procurement process, there are very few possibilities to change the specifications once the tender has been awarded. This need might arise for example where the data leads to new insights, the supplier proposes some modifications based on experiences in earlier projects or the outcomes are less useful than assumed. When using innovation partnerships or pre-commercial procurement (defined under the financial partnerships model in this study), the requirements are formulated in a dialogue with the market, allowing for more flexibility.

- When it comes to flexibility in terms of acquiring a tailored data service or solution, procurement provides the most possibilities to fulfil specific needs, as the city can define the desired outcomes in the tender. It is important to be conscious though of the risk that the more the specifications deviate from market standards or readily available services and products, the higher costs might be, and the lower the number of potential tenderers. In addition, mandatory data sharing puts the city in charge of the specifications, but requiring very specific efforts for a single city might complicate the relationship with (potential) service providers.

- Intermediaries (in the sense of data platform or service providers) mostly offer ready-made solutions (like TMaaS) that serve many customers and allow little customisation for a specific city. The same goes for in-kind partnerships and crowdsourcing, where one party engages in a one-to-many relationship with a number of authorities. An example of this is the Waze for Cities programme, where all cities get access to the same API or OpenStreetMap with thousands of people using the same mapping platform.

Budget and other compensation

- It is clear that in most cases, the in-kind partnership, mandatory data sharing, authority collaboration and crowdsourcing models do not require financial compensation for data. However, it must be taken into account that these models might require other returns like sharing city data, visibility for the data supplier or substantial legal, communication, technical or other efforts from the city administration.

- Crowdsourcing and crowdsensing offer possibilities to collect data on a scale that would be infeasible or extremely resource-intensive through commercial contracts. However, the city should consider whether this acquisition model may be more resource-intensive than commercially obtaining data.

- When working with intermediaries and their standardised products or services, urban authorities can sometimes rely on economies of scale obtained by the intermediary. Using these standardised services can be cheaper than building own platforms or performing data processing itself, as described in the TMaaS case above. Additionally, the cost of updates and new functionalities can be shared amongst the customers of the product.

- In models where the city acquires data without financially compensating the supplier, the public authority will have no or less entitlement to guaranteed support or assistance in processing or interpreting the data (like in the Waze for Cities case or when mandatorily requiring data). If this is the case, cities can appeal to and budget for intermediaries or consultants to provide support and process or interpret data.

Organisational impact

- As indicated above, the impact on the city administration might be rather limited when using straightforward procurement, as the city has experience with this acquisition model and the supplier will mostly follow the tender specifications. Calling upon intermediaries will probably also have a limited impact on the city's resources, especially when using readily available services. Innovative procurement, partnerships and collaborations might be more resource-intensive, as in many cases the administration will play a more active role in achieving results. Additionally, the exact impact on the mobility department might be more difficult to estimate when the efforts are not clearly agreed upon in advance or described in a contract.

- When procuring data, potential financial or other risks will mostly be borne by the contracting authority in cases where the tender does not deliver the desired results or return. In some other partnerships and collaborations, the benefits and risks can be shared. For example, in the parking rights database case described above, service providers earn a percentage of the parking income collected through their products, but have to make some initial investments themselves.

- In models where the urban authority takes the initiative, data are mostly collected with a specific direction in mind, like planning, validating or evaluating current policy. In the crowdsourcing model, people might collect data for various reasons, including potentially opposing or discrediting policy measures. In most models, the city is the main recipient of the data, and the administration has the chance to interpret and contextualise it before sharing. In the crowdsourcing model, members of the public collect the data and can potentially share their interpretations before the municipality has a chance to merge, verify, explain or appraise the data from all the different sources or sensors. Thus, this model might require an alternative collaboration and communication strategy to procurement or contractual partnerships.

Technical capacity

- In models like the in-kind partnership, crowdsourcing and authority collaboration models, public authorities are more likely to receive raw data or data that are not preprocessed to fit their specific needs, as municipalities typically do not pay for these kinds of services. This means that the city will need to hire or possess the technical capacity to process, manipulate and interpret the data. In other models like the procurement, intermediaries and financial partnership models, these services can be included in the tender or agreement.

- In models where the service providers supply technical, data processing or aggregation services, there is a chance that cities will have less opportunity to build their own capacity. For example, when working with intermediaries that use proprietary processes, the technical manipulation can appear as a black box mechanism to cities.

- In most models, the data supplier is expected to deliver neutral data that has as little bias as possible. For example, when procuring data, receiving data through an intermediary or collaborating with another authority, the city should expect to receive the most qualitative data possible. When a professional supplier or intermediary (pre)processes the data, the public authority should be able to count on this being done in a neutral and objective manner. In the mandatory data-sharing model however, the city sometimes uses the data provided by the service provider to evaluate the functioning of the very same service provider. In this model, it would be useful if the city had the technical expertise in-house to verify the data processing, or outsourced this task to a neutral third party.

Relationships

- When calling upon intermediaries that have pre-existing collaborations with other parties, cities might obtain access to a broader ecosystem, like in the case of Waycare (cf. section 2.3). This also goes to some degree for in-kind partnerships, like in the case of Waze for Cities (cf. section 2.5), authority collaborations and crowdsourcing, if those partners are already surrounded by an ecosystem prior to collaborating with the city.

- In some models, a dependency on certain data sources or suppliers might arise. This is more likely in models where the city uses existing services or platforms and does not have ownership of the data nor insight into the data processing, or when the supplier actively makes it difficult to switch to competitors by using proprietary data formats (which can sometimes be seen as a form of vendor lock-in).

- In the in-kind partnership model, the partner mostly expects something in return for the data: this can be city data, a better reputation, press attention, having an edge over competitors, etc. What exactly the partner's interest is might not always be as clear as in models with more clearly defined supplier and customer roles.

- In these in-kind partnerships, there is a possibility that the reputation of the data supplier will somehow be associated with the city. In some cultures, accepting "free" data could be perceived as approving the partner's activities or at least its procedures and processes for collecting data and safeguarding safety and privacy. For instance, according to this article, Los Angeles engaged in a data-sharing partnership with Waze, but did not renew it after two months because the navigation app was blamed for causing increased traffic in residential neighbourhoods, which would make it difficult for the city to defend such a partnership. A city's reputation might be less linked to a supplier when the municipality is solely fulfilling the role of a customer, rather than a partner. In addition, in some supplier-client relationships like procurement, the city might be in a stronger position to impose or require certain certifications or standards regarding data collection and processing.

- The mandatory data-sharing model is the only model where the data suppliers are not really (business) partners, but are required to provide data if they want to operate in the city. This could result in a different kind of relationship between the city and the data supplier, requiring a different approach to the partnership or procurement model. Additionally, one of the goals of the data acquisition is to verify if the service provider (like a free-floating e-scooter company) abides by the terms of the operational licence. In most of the other models, there may be checks on the data itself (such as the quality), but most of the time the data provider has no impact on the mobility patterns described by the data.

- Public authorities are already organised in a hierarchical way (EU-Member States-regions-municipalities) and have a lot of pre-existing collaborations between them on other topics, which could make it easier to collaborate on data acquisition and exchange as well, as described in the authority collaboration model.

- Crowdsourcing, as described above, may require another relationship with the data supplier, which in this model might be many individuals with an opinion on mobility in the city. As in the mandatory data-sharing model, some of the people collecting data might have an interest in a certain outcome of the data collection. In this model, the municipality will in many cases have to engage with individuals rather than well-structured organisations, which will require a different approach that could be more time-consuming. Even so, some crowdsourcing communities are very well organised, like the Waze and OpenStreetMap communities described earlier in this chapter.

Implementation speed

- Working with intermediaries with pre-existing services or tools in the market is sometimes a faster way to obtain results than through (innovative) procurement models where solutions still have to be developed or tailored.

- In some of the examples described of in-kind partnerships and crowdsourcing, well-established collaborations or platforms offer quick and easy opportunities for cities to join (like joining Waze for Cities and using OpenStreetMap or Telraam). In most of these cases, once the platform or standard (like the TOMP example) is developed and available, it can be used quite fast.

Legal capacity

- Legal work will be more limited in models cities have experience with, such as straightforward procurement. In financial partnerships (like innovative procurement) or mandatory data sharing, much will depend on existing contracts or legal material to build upon. For example, MDS (cf. section 2.7) does not provide sample contracts, but according to Jascha Franklin-Hodge, municipalities using MDS share legal documents and update them to match the local context, which is much faster than writing agreements from scratch.

- Some models are also more likely to come with a take-it-or-leave-it model contract with little margin for the city to impose certain changes or specifications. This may be the case for in-kind partnerships like Waze for Cities, with over 800 cities signing a similar agreement, or intermediaries with standard contracts for off-the-shelf services they offer.

- The authority collaboration and crowdsourcing models that publish data in the open domain (such as under a Creative Commons licence), as described in the OpenStreetMap and Telraam cases, are probably the most straightforward for legal departments. They use licensing agreements that are widely known and accepted.

- The mandatory data-sharing model might be slightly different than the other models, because it does not always involve a constructive (commercial) partnership between equal parties. Sometimes the city unilaterally imposes its data-sharing requirements upon the service provider. This can lead to situations where the agreements or contracts are questioned, like the discussion leading to Uber filing a lawsuit against the Los Angeles Department of Transportation.

Market disruption

- The in-kind partnership model can potentially disrupt or impoverish the market if one party offers in-kind data that other companies are offering for sale. Such disruption could theoretically be possible, but is much less likely for the mandatory data-sharing and crowdsourcing models, as the data exchanged mostly differs from what is offered on the market.

- While most models have a more indirect impact on urban mobility, mandatory data sharing can potentially have repercussions on the actual service offering. For example, Uber's permit was suspended in Los Angeles for violating the city's data submission rules. Apart from revoking licences and thus impacting mobility services, it is not unthinkable that some service providers might hesitate to request a permit in the first place if the city imposes obligations that are far-fetched or difficult to implement.

- Acquiring commercial data through any of the models and sharing them as open data can be disruptive as well, because neither the supplier nor its competitors will be able to sell such data for the given territory. This is the main reason why the National Road Traffic Data Portal in the Netherlands only shares aggregated data rather than the raw floating vehicle data they procure.

- Collaborations between authorities can also impact the market. For example, the National Road Traffic Data Portal in the Netherlands could easily build advanced tools on top of the data they collect to serve municipalities. But even though these municipalities are partners and partly finance the National Road Traffic Data Portal, the national government is reluctant to do so because it would disrupt the market for intermediaries and consultants.

Combining models

- The authority collaboration model can be combined with most of the other models, for example:
  - In-kind partnerships with the private sector: smaller cities teaming up or delegating their regional authorities to join the Waze for Cities programme.
  - Procurement: NDW buys floating vehicle data for all of the municipalities in the Netherlands.
  - Mandatory data sharing: Los Angeles and other cities joined forces with intermediaries to create and deploy MDS that is used as an interface by over 40 cities for mandatory data sharing.

- Intermediaries can be called upon for deploying (open source) tools or processing data resulting from acquisition through other models. For example, intermediaries are involved in implementing the open source MDS APIs, analysing and storing Waze for Cities data, operating platforms in the Talking Traffic Innovation Partnership and assisting in mapping city data on OpenStreetMap.

Chapter 3: Basic principles

3.1 Introduction

The purpose of this chapter is to address the wider policy, regulatory and organisational aspects that can inform, shape and influence the outcomes of mobility data sharing. The section on legal aspects of data deals with contractual issues and data privacy matters; it draws heavily on the cases collected for this report and the interviews conducted. Organisational aspects consider a range of areas that are specific to the culture of an administration, including data skills and expertise, data governance and data management. Mobility data (sharing) policy and regulation are considered primarily at the EU level, since there is some variety in the data policies, laws and cultures across the European Union. The final section of this chapter provides an outlook on future trends.

3.2 Legal aspects

3.2.1 Contracts and licensing models

Data clauses in concession contracts for off-street parking

Data on the occupancy of parking garages is important for managing traffic in bigger cities. Drivers being directed to a parking garage that is full can result in congestion at the entrances or the need to travel further in search of another parking spot. By sharing occupancy levels as open data and/or with navigation systems, these kinds of situations can be prevented. However, many cities have experienced difficulties accessing such data from private parking operators. The presumed underlying reason is that these operators, whose concessions may last from years to decades, fear that competitors will know exactly how many cars are parked and thus get an insight into the concessionaire's earnings, which would then lose a competitive advantage.

ITS Belgium has proposed a solution that takes into account the perspective of both cities and parking operators. It has drafted a number of paragraphs that can be included in a tender or concession contract, describing how the operator has to share the data with the city. A pragmatic solution has been found to serve the interests of both parties: only the data relevant to traffic management are shared with the city. The operator shares the status "available" until a certain percentage is reached, and only from that moment on are the exact numbers on free spaces given. If a city wants to use parking occupancy data for purposes other than traffic management, such as traffic models or monitoring parking policies, more detailed data will need to be exchanged.

Case validated by ITS Belgium.

Generic data licence and service level agreement used by the city of Antwerp

Antwerp has drafted two standard contracts for exchanging data. At the time of writing, the documents were not officially approved and therefore not yet published online. Once the licences are approved, all (shared) mobility providers wishing to operate in Antwerp as well as the city's data providers will have to sign both contracts.

1. A licence agreement for third parties that transmit mobility data to the city. This document begins by clearly indicating the city's vision: a 50/50 modal split for the entire region and the importance of accurate data in reaching this goal. The contract includes statements on intellectual property, asserting that all rights remain the property of the provider, but that the city has an unconditional right to use or distribute the data. Furthermore, the obligations of both the city and the data provider are described, mainly that they should act in a responsible manner. Regarding the GDPR, the contract states that the data exchanged should not qualify as personal data and should be lawfully collected. Finally, liability and sanctions are defined, as well as how to deal with situations not provided for in the contract.

2. An agreement on the levels of service, which also starts with the mobility policy goals. The contract stipulates that the data should be captured directly at the source (meaning that there are no processing steps outside the city's control), be of good quality and be kept up to date. Different types of possible incidents are described along with a practical example, priorities are given and lead times for incident response assigned to each category. The contract also contains brief specifications on how problems should be solved, contact information, specifications on availability, processes for updates and new releases and performance indicators. For example, data must be available 99.9% of the time and trip data must be accessible for the city within 48 hours after the end of the trip.

Interviewed: Stijn Vernaillen, MaaS expert at the city of Antwerp.

Data format/interface standardisation clauses in contracts with transport operators

The Data4PT project will define standard requirements specifications to support tendering procedures, and points out that the EU directive, the related EU delegated acts and national laws are starting to introduce (public transport) data standards as mandatory requirements, for example Transmodel and NeTEx in Commission Delegated Regulation (EU) 2017/1926, which also refers to DATEX II, INSPIRE and SIRI. Kasia Bourée from Data4PT adds that several data categories are published through the National Access Points according to the EU Delegated Regulations (priority actions A, B, C, D). Cities can refer to the EU regulations for standards for data exchange with public transport operators and further facilitate data processing, exchange and publication.

The Data4PT experts indicate that the ITxPT requirements specification (based on the standards from the European Standardisation body on ITS in public transport) is frequently used in public transport tenders. *(Note from the authors: ITxPT is a leading Data4PT partner and an implementing body to support the adoption and deployment of EU ITS standards).* Their specifications contain references to specific standards like Transmodel, SIRI and NeTEx and are under continuous development (specifications can be consulted here after free registration). Bourée observes that a growing number of cities are using the ITxPT specifications for ITS architecture, which she sees as an example of how standards are becoming an enabler for the digitalisation of mobility and the interoperability of IT systems, along with the European legal obligation for cities to use standards like NeTEx for static public transport data. Examples are Ile de France Mobilités having made NeTEx and SIRI mandatory for all their operators, which then made them mandatory for all their software providers. Another example is the adoption of SIRI in Lyon. In addition, RUTER in Oslo and TfL in London are using public transport tenders referring to data standards supported by ITxPT. Such specifications have also been adopted at a national level, like in Norway and the United Kingdom. More examples can be found on the Transmodel and NeTEx web pages and in the 2019 ITxPT activity report, which share practical experiences as well as lessons learned from implementation.

Interviewed: Kasia Bourée and Emmanuel de Verdalle from the Data4PT project. Consulted: Christophe Duquesne and Anders Selling.

Data standards

A number of standards have been mentioned in this document and in the tenders/contracts referenced. The table below provides an overview of these standards and indicates who is responsible for maintaining them:

Acronym	Full name	Origin	Maintained by	Domain
CDS-M	City Data Standard - Mobility	Amsterdam	G5 (five largest Dutch cities)	Exchange data between free-floating transport operators and public authorities
DATEX II	Data Exchange II	CEN TC 278 WG 8	CEN	Road traffic data exchange services
GBFS	General Bikeshare Feed Specification	United States	MobilityData	Bikeshare data exchange format
GBFS+	General Bikeshare Feed Specification Plus	Netherlands	12 Dutch organisations	Bikeshare data
GTFS	General Transit Feed Specification	Google	Google	Public transport data exchange format
MDS	Mobility Data Specification	Los Angeles	Open Mobility Foundation	Dockless e-scooters, bicycles, mopeds and carshare
NeTEx	Network Timetable Exchange	CEN TC 278 WG 3 (Public Transport)	CEN	Planned (static) public transport data exchange services (server to server) – includes network topology, timetables and fares
NGSI	Next Generation Service Interface	FIWARE	ETSI	Context information for FIWARE
OpenLR	Open Location Referencing	TomTom	OpenLR Association	Location referencing/maps
OSLO - Mobility	Open Standard for Linking Organisations	Flanders	Flemish Government	Public transport
SIRI	Service Interface for Real Time Information	CEN TC 278 WG 3	CEN	Real-time (dynamic) public transport data exchange services
SIVU	Service Interface for Vehicle Use	Paris	City of Paris	Free-floating vehicle providers
TOMP	Standard data flows for Transport Operators and Mobility-as-a-Service Providers	TOMP Working Group	TOMP Working Group	Communication between transport operators and MaaS providers
Transmodel	Public Transport Reference Data Model	CEN TC 278 WG 3	CEN	Data model for public transport network topology, timing information and scheduling, operations monitoring and control, fare management, passenger information, personnel management, management information and statistics

Contractual agreements are commonplace in data sharing, even in cases where there is no financial compensation, such as the Waze for Cities' programme described in chapter 2. While contracts bring reassurance to the data provider and the data user, they can also be an administrative headache. Several cities interviewed for this note recommended using contracts only when needed, including Ghent: "*Unfortunately, a lot of complicated contracts and clauses need to be signed, even for getting free data. This can be very time-consuming. Apart from the administrative burden, we have not had any problems with contracts or clauses backfiring."* Tomas Straupis from the OpenStreetMap community interviewed for this study indicated that the city of Vilnius' release of its transport data into the public domain – and hence no contracts needing to be signed anymore – has been an enabler for working with the data. Paris managed to simplify the administrative process by keeping the pedestrian flow project contract value under the national threshold for a direct award. This would not have been possible with an award for an ITS system, for instance, which tends to involve large sums of money and cumbersome procurement and implementation processes. Data acquisition can be scaled to the specific requirements of the customer.

The Gothenburg/Swedish case of floating vehicle data procurement allows for some flexibility. For instance, it is possible to increase the extent of the road network contracted for the delivery of floating vehicle data by up to 10% without recourse to a new tender. Furthermore, the selection of roads for which floating vehicle data are to be delivered remains flexible throughout the contract period, provided the overall number of contracted road kilometres is not exceeded. It would be an interesting exercise to further explore the degree of flexibility in data contacts, in comparison with ITS contracts for instance. Some of the ITS tenders examined in the EIB Technical Note on ITS Procurement for Urban Mobility revealed provisions for additional systems, services and developments that the procuring body was under no obligation to implement (section 3.6.3). This provision is useful where there is uncertainty about the need for a particular item at the outset of the procurement but the option is there in case it is needed, without recourse to a new tender.

The shift from system procurement to service procurement has an impact on the nature and the duration of contracts. The ITS systems/services examined for the EIB Technical Note on ITS Procurement for Urban Mobility typically included a reasonably long maintenance contract (five to ten years) following the usual one to two year system guarantee. Data contracts tend to be of a shorter duration – in the Gothenburg/Swedish case, the contract was for two years, renewable two times for a one-year period.

The contractual relationship/business model

Contracts are typically established between the data generator and the data customer, as demonstrated by many of the data-sharing cases described in chapter 2. As new data acquisition models emerge, the contractual relationships and business models are being adapted accordingly. The rise in data integrators/aggregators in particular is impacting the business models. The classic approach to procuring data integration/aggregation services is where a supplier of software (such as an open data platform or data visualisation tools such as a dashboard) has a contract with a city and the city must equally have contracts with the generators of the data needed to populate the software tools. This is confirmed by Opendatasoft: *"We have one customer, one contract, it's the city that has the contractual relationship with third parties."*

A model is starting to take off whereby the platform provider enters into a contract with the city on the one hand and with the (raw) data producers on the other hand. In this model, the city is not necessarily buying raw data, but rather a set of services, which may vary from aggregated data to travel insights. This is the model represented by the three cases described in section 2.3, whereby the data integrators/aggregators are actually providing a platform service to the cities based on data they have acquired from different mobility players. In this model, working with intermediaries removes a lot of technical hurdles compared to acquiring unstructured or raw data. Aggregators provide information needing no or little manipulation. On the one hand, using readily available services can speed up getting results. On the other hand, this means the city has little knowledge or control over the data sources or how the data are processed and typically does not build up a lot of expertise in this field. Furthermore, in some cases the (raw) data are kept by the intermediary, which could make it difficult for cities to retrieve the data for other purposes (cf. below "Data access in new data acquisition forms"). Thibault Castagne from Vianova indicates that this very much depends on the service provider contracted: *"I would encourage cities to look for open APIs, automatic dumps of non-sensitive data and a full technical documentation of analytics."*

Data ownership and reuse terms

Public authorities agree on the necessity of including access to data in any contract they enter into. This omission in former contracts has come at a high price for public authorities: either they pay extra to access the data or they have to do without them. Data reuse and ownership terms are central to data-sharing contracts too. The main difference relates to intellectual property: in data-sharing contracts, while the customer may own the data procured, the actual reuse of that data is often subject to terms set by the data provider.

In both the Swedish and Paris cases of data procurement (section 2.2), the opening up of the data is strictly prohibited. This would destroy the business case of data providers, as TomTom pointed out (cf. section 1.7). In Sweden, the information can be combined with other data, where available, and disseminated by means of variable message signs or traffic information websites. Similarly in Paris, the results of the traffic analyses can be made publicly available but not the raw data itself. There are some other interesting data use and ownership clauses in the Paris contract:

- The data can be used solely for the purpose of traffic flow analysis and cannot be commercialised in any way or integrated into the administration's database.
- The raw data becomes property of the city of Paris although reuse of the data is limited as already described.
- The city of Paris can share the raw data with another consultant provided that a confidentiality and non-disclosure agreement is signed and the raw data are destroyed at the end of the study/collaboration. The analyses must also mention the data source.

The Dutch National Road Traffic Data Portal (NDW) buys floating vehicle data on a national level and makes it available to all cities and provinces. The data procured are not raw, but are fairly detailed. The data that are being shared however are aggregated and the level of detail varies greatly with the data acquired from the supplier. NDW indicates that this is an informed decision. On the one hand, NDW feared a lot of existing open data streams on travel times and speeds based on Bluetooth and ANPR camera collections would disappear (as floating vehicle data would be cheaper) and this would impact the work of many companies relying on these open data. On the other hand, sharing the data as procured would have an effect on the price and the level of competition of the supplier – it would not be able to sell these data again in the Netherlands as anyone could access them as open data. Furthermore, competitors could have insights into the data and the data quality. Therefore, NDW has chosen to pay the contractor extra to allow the distribution of an aggregated dataset as open data, and not to share the more detailed source data. It should be noted that the larger public authorities in the Netherlands contribute towards the cost of running NDW. The approach described can fuel disagreements because local and provincial authorities first partly finance the data acquisition, then the aggregation process, to only receive the aggregated data.

To avoid paying twice for the same data, the city of Antwerp stipulates reuse clauses in its contracts. Previously, different departments procuring the same data each had to pay for the data because the contract did not allow the data to be used by other departments within the municipality or for purposes other than those included in the contract (as stipulated in the Paris contract mentioned above). This is no longer the case today in Antwerp. Data reuse terms are described in detail in each contract, of which there are essentially two types:

- The data can be used for other internal purposes, without necessarily specifying that purpose.
- How the data will or could be made available to third parties.

André Ormond recommends that cities include clauses on data in any contract or tender. Ultimately, a lot of urban mobility services or systems will include a data aspect or somehow generate data. Ormond gives the example of parking meters, which were analogue and stand-alone before, but are now mostly connected and generate a lot of useful information on parking behaviour and availability. Cities can avoid paying extra for the data generated by the system by making good agreements upfront and/or including clauses in tenders or concession agreements on data usage and ownership. Such clauses can be obtained from the tenders of peers, knowledge networks or (European) projects (cf. the boxes above on the sample specifications from the Smart Flanders and Data4PT projects).

Data access in new data acquisition forms

Thibault Castagne, Vianova:

"Starting with a standard contract can accelerate the contractual process. A city designing its own data-sharing agreement typically needs six to nine months."

As pointed out above, there is a tendency to work with intermediaries that capture and process the raw data and only feed information to the city. Establishing data access and ownership clauses in software-as-a-service contracts for example is especially important, as Antwerp points out: "*Getting your data is difficult with SaaS solutions, the data they have is their business model. If you enter into a four-year contract and then stop, you will lose four years of data."* Eight companies offering such services that were interviewed for this study indicate that cities always have access to the data, even if they stop using the service. Stijn Vernaillen from Antwerp, however, indicates that such data dumps can sometimes be hard to work with, because the intermediary does not provide the city with insights into its data structure and specifications. Therefore, Antwerp wants to be in the middle of the data chain, rather than at the end: all data should be processed and stored in its systems, so it can easily switch (or combine) data sources if needed. To put this into practice, Antwerp has developed ACPaaS

– Antwerp City Platform as a Service – which also defines how software-as-a-service solutions should be implemented.

Other terms and clauses to consider

Smart Flanders: sample clauses on open data for urban authorities to use in agreements

The Smart Flanders programme from the Flemish government in Belgium has made what they call a scenario document (in Dutch) with example clauses on (open) data for tenders and contracts. The guide was created after observing missing, unclear and widely varying specifications on data in government agreements with suppliers. The content is designed to help urban authorities better define roles and responsibilities for data when tendering or (re)negotiating contracts.

Some of the topics on which example clauses are provided:
- City data vs. personal data
- Open licences and legal implications
- Machine-readable and open formats
- Linked open data
- Selection and award criteria
- Ownership and access to data
- Metadata

Other types of interesting administrative and technical clauses contained in contracts include:
- Quality clauses: The public authorities in Sweden have stipulated minimum levels of real-time data to be provided at key times (such as peak hours) in their procurement of average travel times (tender document – cf. section 2.1).
- Detailed data requirements: Paris has described in detail the data it wishes to procure, in terms of source, format, definition, quality and application (tender document – cf. section 2.1).
- Skills profiles: Paris has established minimum conditions for the skills profile of its contractor. The team must have at least one expert with a minimum of five years of experience in data processing and traffic engineering.
- Market clauses: To accelerate the development of MaaS services, any mobility operator wanting to enter the Antwerp market must at least have "deep" integrations with two MaaS apps. This means that the services can be procured inside a MaaS app, without having to leave the application or the customer having to interact with the app of the service provider. This has led to an active ecosystem, with a lot of integrations between companies and a richer offering for consumers.

Tendering

In the fast-growing data-related sector in which cities are also requesting innovative solutions, the traditional way of public tendering is not always suitable, as is reflected in this opinion piece by Ethar Alali: *"As of 2015, we started to see many more contracts appear wanting 'agile' skills, 'cloud computing', 'DevOps', 'SaaS' or 'Internet of Things' etc., but the one thing that remains constant is the procurement process."* Tender documents serve as a contract between the local authority and the supplier. On the one hand, they should ensure upfront that all loopholes are closed; on the other hand, the provisions cannot be too tight in order not to hinder the usage of new technologies or creative and innovative ways to reach the goals set. One way to overcome this is by using innovation procurement schemes, which STIB/MIVB – Brussels' public transport provider – is doing to mutually define specifications (cf. the Muntstroom case in section 2.4). This is however not a solution to procure as-a-service solutions. Parabol indicates that it is possible for cities to subscribe to such schemes using traditional tendering procedures, but it requires a lot of administrative hassle and cities sometimes lack expertise.

Stephan Corvers, Corvers Procurement Services BV:

"One reason why the US was successful earlier in technology development and deployment is its procurement instruments pushing for research and development. Public authorities in the EU wanted to procure without risks, sometimes resulting in old-fashioned technologies and services."

The mobility data ecosystem is also characterised by startups. The classic tendering procedures might disadvantage such startups, says Noam Maital, CEO of Waycare, one such startup. In his view, which is shared by other startups interviewed, governments on the one hand set requirements that are hard or impossible for startups to live up to, such as providing a solid financial history for over five years. On the other hand, there are often requests for heavy contracts concerning more than ten years or multiple millions of euro. Most experts interviewed recommend going for an agile approach: smaller pilots, iterative wins and fails, and involving multiple vendors. This might require substantial effort from the public authorities and entails the risk that one of the partners ceases to exist. But this way of working has clear advantages: cities can be more flexible and iteratively pinpoint the challenges they are building solutions for. To foster such collaborations, it might be beneficial for cities to pursue tendering strategies that are inclusive for startups as well.

In comparison with the ITS sector, data seem to lend themselves to joint procurement, as evidenced in the Paris, Brussels and Gothenburg/Swedish cases for this study. The benefits of jointly procuring floating vehicle data among four different public bodies in Gothenburg/Sweden are many. The national road administration, which ingests and transforms the data for the other parties, no longer has to deal with multiple suppliers, data flows and quality. Travel times can be provided across road jurisdictions and the city authorities partnering with the national road authority are able to take advantage of its technical expertise and infrastructure.

André Ormond, Ormond Consultoria e Treinamento:

"Whenever you buy or procure something, be it parking meters or new streetlights, include data in the contract. Preferably, some standard clauses or paragraphs can be made about data (ownership, standards, quality, availability, etc.) and be included in any tender or contract."

3.2.2 Data privacy

Personal data sharing and portability by MyData

Many systems collect personal data: social media, route planners, smartphone applications, telecom providers, etc. Even though the GDPR makes it possible for people to request and to a certain degree control these data, it is not transparent to the user what system is collecting which data. And having the data erased or even getting insights into what is collected where is very hard or almost impossible.

Some new schemes are advocating "data portability": this means that the user can choose where their data are stored, manage this storage and grant or deny access to applications or platforms. Some systems only advocate consent management (the data can stay where they are, but users should be able to easily change their preferences or relocate their information). Others also propagate the storage of personal data in neutral data vaults. Amongst the new schemes are Solid (by Tim Beners-Lee, "founder" of the web) and MyData.

The non-profit global MyData organisation referred us to Onecub for an interview. This startup is a platform for data portability and a MyData operator, meaning it implements the MyData principles, which are also mentioned in the European data strategy. Onecub CEO Olivier Dion explains that his platform should be seen as critical infrastructure like 4G or 5G networks. On top of this, operators can build services.

The Onecub platform in essence manages consent from end-users: it keeps track of what data are stored where and what they can be used for. It does not store the data itself, as it wants to apply the separation of powers principle. For example, let's say you are using a popular navigation app that is also tracking your location in order to register travel times, traffic jams, etc. In the traditional model, your location data would be stored on the servers of a private company and shielded from the rest of the world. With the MyData model, you have full power over your personal data that are being stored and you can give or refuse consent for them to be used. This means that you can use the data, share them with your city or a competitor of the navigation app, or revoke access to the data at all times. Meanwhile, infrastructure like Onecub covers more data portability aspects like APIs, consent mechanisms, and legal and commercial services.

When asked for the role of cities in this, Olivier Dion states that cities are best positioned in the ecosystem to be at the centre of data sharing because they are at the crossroads of many daily activities and interactions with private and public entities. He adds that today, mobility data are already stored in a lot of MaaS platforms. Therefore, there is no need for people to individually store all their data; good consent management enforced by the local authority would in most cases suffice.

Interviewed: Olivier Dion, CEO, Onecub.

In its Managing Mobility Data paper, the National Association of City Transportation Officials (NACTO) explains the challenge of reconciling the city's and business' need for data with the individual's right to privacy. It argues that a positive outcome can be achieved for cities and individuals through the use of thoughtful tools and principles; in other words, a good data governance policy (cf. section 3.3.2).

The need to think twice

As explained in section 3.2.5, the introduction of the General Data Protection Regulation has transformed the way in which organisations deal with personal data. The city of Lisbon, for instance, has set up an independent GDPR unit to advise the city council on data privacy issues. The mobility sector has not escaped this overhaul. It is forcing organisations to think twice about working with any data that may be considered personal and to be overly cautious about data sharing, even where the purpose has strong merit. By way of example, the city of Lisbon had wanted to correlate air quality data with health data (mainly respiratory problems, such as asthma). The health authority refused the request for data, so the city of Lisbon took a reverse approach: instead of requesting data about people with respiratory conditions, it offered the air quality data to the health authority to make the analysis.

Data aggregation and anonymization

An effective way to comply with privacy regulations is to modify data so they cannot lead to the identification of individuals. This can be done through data aggregation, i.e. combining data and presenting them in a summarised format. For example, when visualising origin-destination information for an event, information might only be shown for a certain origin when a minimum of 50 individuals begins their journey at the same point. Another method of "de-identification" is anonymisation: removing those parts of the datasets that can lead back to individuals. For example, anonymisation of licence plates in video footage (example from Germany) or removing the parts of individual journeys between home locations and arterial roads. This website gives some guidelines for de-identification and this large-scale study on 30 billion mobile phone call records warns that *"sharing anonymised location data will likely lead to privacy risks,"* especially due to the fact that the dataset contains multiple data points linked to one person and can be identified as home or work locations.

Data aggregation can be performed by an intermediary or by the city itself. Cities tend to acquire aggregated data from intermediaries, either because they lack the capacity to process raw data, the use case only requires information on a more global level, or out of an abundance of caution and uncertainty due to perceived data privacy risks. Intermediaries/data providers generally have more experience of mobility data de-identification or are already processing the data to some degree. Nonetheless, it should be noted that cities already stored a huge amount of personal information long before the GDPR, and treated them with due care. In the case of mandatory data sharing (cf. section 2.6) and information provisioning in concession contracts (e.g. for off-street parking, cf. section 3.2.1), requesting aggregated data can lower the threshold for service providers to exchange information, due to less business-sensitive data being obtained by the city.

Ultimately, the choice between acquiring aggregated or disaggregated (raw) data and the involvement of an intermediary depends on the use case, i.e. the purpose for which the data are required. For long-term planning and analysis purposes, aggregated data usually suffice. For regulatory/enforcement cases, raw data are often needed, for example where a city has prohibited parking free-floating mobility devices in specific areas, detailed parking data per device are needed to verify the compliance of mobility operators. The recently launched Polis report on micromobility data confirms that raw data are mostly used by public authorities to monitor compliance with rules and that "individual vehicle status" and "individual vehicle location" are the most commonly used data elements. According to NACTO, data aggregation is a key tool for managing the balance between access and privacy. While it is accepted that the aggregation level goes up where population density goes down, there is no agreement/standard on the aggregation threshold. The NACTO policy document also adds that the retention period is an important factor: *"In general, cities may choose to hold individual trip records for brief periods of time, for example until enough data can be gathered for processing or aggregation or until specific violations (e.g. a parking ticket) are addressed."*

A further important consideration is the reliability of aggregated/anonymised data. There have been examples of mobility companies providing misleading or incomplete data to regulators, including the Uber Greyball programme and micromobility examples in Chicago (p. 75-76) and Nashville. It is recommended that if the data aggregation or de-identification is carried out by a third party, the urban authority should have transparent insight into the manipulations. This is in order to verify the quality of the result, but also to obtain consistency with other datasets and ensure a correct interpretation. If certain datasets are aggregated at another level or processed in a different way to other information used in the city, comparisons would be hard, and interpretations might be incorrect.

Finally, it is considered good practice for cities to require intermediaries to demonstrate their compliance with data privacy rules. Nommon confirmed that Spanish public authorities procuring information based on mobile data requested that the telecoms provider carry out a GDPR compliance check. In its procurement of aggregated GPS data from smartphones for travel demand analysis, the city of Paris believed it had exonerated itself from any potential privacy conflict, even if the primary reason for selecting aggregated data was for technical reasons – they are easier to interpret than raw data.

The end justifies the means

The heated discussion about data privacy in the mobility domain may have been influenced by the heavy media coverage surrounding the data specification MDS (cf. MDS/OMF case in section 2.2), which some mobility operators are challenging for data privacy reasons. Against this backdrop, the growing market of free-floating data aggregators/integrators sees itself as an effective and trusted intermediary between the data generators and the public bodies requiring the supply of vehicle movement data. According to Vianova, in such a service model, the city (as data controller) is nonetheless accountable for the correct implementation of the GDPR (regardless of who is actually processing the data), whereas the data aggregation platform assumes the role of data processor, which involves more limited compliance responsibilities. Vianova believes that while MDS does indeed constitute indirect personally identifiable data, it should not prevent a public authority from gathering such data. It claims that the GDPR does not prohibit the collection and processing of personal/personally identifiable data. However, there has to be a clear and justifiable reason for gathering the raw (personal) data as opposed to aggregated data, which some operators are calling for. These reasons include verification of data quality and veracity, input for transport planning and enforcement. This is in line with the GDPR concepts of necessity and proportionality: *"proportionality requires that only that personal data which is adequate and relevant for the purposes of the processing is collected and processed."*

Thibault Castagne, Vianova:

"Data are needed for the enforcement of micromobility regulations. Micromobilty data, for instance, can help cities verify the compliance of operators and users, related to parking, service usage, maintenance and zonal vehicle caps, among others. Data can alert a city when an infringement of a regulation occurs."

The city of Antwerp would welcome further guidance on which data level (aggregated or individual) is best suited to which use case. Meanwhile, the city prefers to start with aggregated data (like floating vehicle data) to get a feel for mobility patterns and to see how people are moving. Individual vehicle data will be acquired if and when there is a need, one important need being to monitor new mobility providers to ensure they abide by the rules set for that particular type of service. This view is opposed to that of Los Angeles, which states that it needs such data to pace up to be prepared for autonomous vehicles and other future mobility trends.

Privacy by design

The city of Antwerp advocates the privacy by design principle, which essentially integrates data protection with technology design. In a data-sharing context, this implies that data protection should be incorporated into the data processing procedures at the design phase of a system.

Privacy by design is also promoted in the data-sharing principles paper of the World Business Council for Sustainable Development (WBCSD). In addition to incorporating privacy by design principles into data repositories, access mechanisms and sharing protocols, the paper recommends building trust with stakeholders (including citizens) and creating governance structures that administer and regulate privacy frameworks, such as neutral third parties and data trusts. It also advocates clearly defining and communicating the scope and purpose of the data to be collected, and considering consent-based data sharing to enable people to view and redact data.

International Transport Forum – Data Driven Transport Policy:

"In the vast majority of cases de-identification will strongly protect the privacy of individuals when additional safeguards are in place."

The International Transport Forum published a study on data-driven transport policy that discusses privacy by design for the use of geo-location data and that argues in favour of data minimisation and de-identification in system design.

3.3 Organisational aspects

3.3.1 Capacity and expertise

When it comes to acquiring and working with (urban mobility) data, a lot of cites are wondering what internal expertise is needed and how to attract the right profiles. And if certain skills are not present within the administration, can and should such expertise be found externally? This section includes a number of examples and recommendations from both cities and suppliers. It is interesting to note that commercial companies recommend that certain competencies are present within cities, in order to be able to set up partnerships and to achieve good results.

Stephanie Leonard, TomTom:

"Since 2008 we have been offering data services, but we still see a lack of understanding of the data ecosystem. Even today, cities need to be convinced that traffic data are a cost-efficient way to gather insights in urban mobility. We would have thought that we'd be much more advanced in 2020."

Recruiting data experts

Chapter 1 outlined the importance of adopting a strategic approach to data, particularly defining the objectives/use cases and the data that are needed to fulfil those purposes. Being able to work with such data to support mobility management or to gain insights into travel patterns does require skills. There is no point in acquiring huge swathes of data if the skills and resources are not there to make use of them. Having the right set of skills within an organisation is therefore crucial. This has been recognised by many public authorities, and indeed many larger cities are setting up data teams. For instance, the Brussels public transport operator STIB/MIVB recruited a data expert some years ago who has since set up and is heading a Data and Analytics unit. Some cities are even prioritising IT specialists for transport posts, such as the former head of traffic management in Rio de Janeiro.

Nonetheless, recruiting and retaining data specialists is not an easy feat for local governments, due to rigid recruitment and career development policies and an inability to compete salary-wise with the private sector. This situation was acknowledged by many interviewees, including Tuğçe Işık who admitted that a recently graduated data specialist would progress faster in the private sector.

In-house skills development

Cities are finding innovative ways of addressing this recruitment challenge, including the city of Lisbon, which has partnered with universities and seconded local government staff to universities to learn from data specialists and to work with data tools. The outcome of this in-house skills development process is a "business-agnostic intelligence team", offering its services across the many "business" areas of the municipality. When the team's expertise is called upon, the business area (department) must make resources available to explain the rules of the business area and to work on the problem/need as a team. Lisbon acknowledges that this takes time and requires commitment. However, the value of this data work is becoming increasingly understood at management level to the extent that the intelligence team's services are now over-solicited.

David Cunha, city of Lisbon:

"Lisbon had to build a data team. It partnered with universities, embedded city staff in universities working with city datasets to enable them to use data tools. Now the city council has data capacity."

Another interesting experience comes from the city of Paris' traffic control centre where traffic engineers had intensive training to be able to work with raw traffic data procured (cf. case study in section 2.2). However, in subsequent procurements of pedestrian and bus passenger flow data, the city mobility department decided to opt for aggregated data/insights rather than raw data, partly because of the specific expertise needed to work with raw data. These positive experiences of working with third-party data have led Paris to consider a cross-departmental tender for data to support its internal analysis.

Mainstreaming the data culture

Beyond the actual data specialists themselves, there is a need to develop data skills more broadly across the organisation. Many employees throughout the municipality work to some extent on data, in particular inputting data and database management. It is important that city staff improve their understanding of data and the context of the data and can check the quality and correctness. All employees need to commit to the skills development process to enable the organisation to properly harness the data potential. The city of Antwerp is promoting this through use cases that show the added value of mobility data analytics and the importance of entering mobility data correctly.

Miguel Picornell, Nommon Solutions and Technologies:

"Whenever tendering or outsourcing, include capacity building for your team. In our collaborations with public authorities we provide information about the characteristics of the raw data, the methodology followed to obtain information from these raw data, the main limitations encountered working with this type of data and how they may be overcome. This way administrations are aware of the technology and possibilities, and it benefits our mutual understanding and current and future cooperation."

ITS and data companies also see the value of skills enhancement in the public sector. Some companies have organised formal training sessions and others offer informal guidance and support. For instance, the free-floating mobility data aggregator and dashboard service provider Vianova has acquired much expertise in the area of GDPR and is now coaching its customers on this topical and sensitive matter. Miguel Picornell advises to include training and build capacity whenever outsourcing specific tasks. Pedro Barradas adds that there are a lot of congresses and webinars available for administrations to build capacity, along with a lot of online content, and he advises to make time and engage in online and offline capacity building.

The resource challenge

Beyond a vision and a plan, the availability of resources is another prerequisite for skills development, as well as for infrastructure. While many bigger cities are in a position to build capacity, the challenges are that much greater in small and medium-sized cities. In the case of limited resources, Antwerp advises that efforts be focused on interpreting the data rather than on data manipulation. For example, knowledge about the data mobility market, about what information you can get from what sources and how to spot deviations. Stijn Vernaillen adds that above all it is important to know how mobility works, get out there and get a feel of how this actually looks like in the actual city. David Thoumas from Opendatasoft also stresses substantive knowledge about local mobility: *"to work with us, public authorities need time and subject knowledge."*

Miguel Picornell, Nommon Solutions and Technologies:
"It is always possible to rely on external expertise, but the core knowledge should also be held within the public authority."

Miguel Picornell adds that even when cities call upon consultants or specialists, administrations should hold some general knowledge about where data come from, how they are processed and how they can help achieve strategic objectives. This basic understanding is needed to specify the city's needs in tenders and evaluate the work done, which is confirmed by Malin Stoldt from the city of Gothenburg: *"Public agencies need a combination of IT and infrastructure skills as well as procurement skills, particularly to procure a service. They need to know what to procure and what the requirements are."*

The EIB Technical Note on ITS Procurement for Urban Mobility describes the case of Timișoara relying on consultants to write a tender and the administration not holding the expertise to fully understand the tender specifications and their implications. Tuğçe Işık and André Ormond encourage cities to hire expertise where needed, but recommend that the city always remains in control.

Even where there is data capacity within a public authority, consultancy services are often needed for specific tasks. Bergen, for instance, has recruited developers and analysts for the construction of its data lake, in particular to create the data flows to the data lake and the flows within the lake itself. The city authority hopes that it will eventually have the internal capacity to meet all its data needs. The city is getting there on a vision level, but acknowledges that finding a budget and defining the business cases are challenging. Antwerp similarly tries to do as much as it can itself and is very cautious about outsourcing to avoid tying itself to a particular supplier. It is for this reason that the city did not buy an off-the-shelf data solution but is building the NXTMobility platform to its own specifications.

Building an internal and external data ecosystem

The cases presented in chapter 2 show the range of stakeholders involved in data sharing: mobility and other municipal departments, data providers, integrators, aggregators, other public authorities, platform operators, consultants, the academic community, etc. In many cases, municipalities will be involved in multiple data projects over time and be surrounded by internal and external partners that jointly work on achieving the city's mobility strategy. Miguel Picornell recommends "creating an ecosystem of companies, like an integrator, an analytics partner, a data processor, data providers, etc." He adds – and other experts interviewed confirm this – that in such an ecosystem it is extremely important to avoid monopolies or strong dependencies on a single partner. Building an ecosystem will be helpful to supplement skills present within the city council and to exchange expertise, but it should be an open ecosystem to allow for innovation and progress. Having multiple partners that can provide similar services is one solution to creating an open ecosystem.

STIB/MIVB's new Data and Analytics team is taking the lead in setting up a data ecosystem among Brussels public agencies and with the private sector, which has been named "Data moves Brussels". Through the Muntstroom project (cf. section 2.4), STIB/MIVB and other Brussels agencies are also creating a wider ecosystem with the private sector to explore the many facets of data sharing, including business models, privacy and security. The city of Antwerp is taking another approach: it requires each mobility provider to offer its services in two Mobility-as-a-Service apps, which has led to an active MaaS ecosystem, with a lot of integrations between companies and a richer offering for consumers.

3.3.2 Data governance

Data governance refers to the data handling processes and procedures in place within an organisation. It is the act of managing the data once they have been acquired and keeping them available, secure, accessible and up to date: *"it ensures that data are consistent and trustworthy and do not get misused."* With organisations increasingly relying on data while at the same time facing more and more data privacy regulations, governance is key to keeping track of and managing the data stored and used across different departments.

The applications of data governance are very broad. It can mean implementing master data, for example using one single list of street names across all departments and keeping it up to date, or using the same definition of what a street is. Governance can include setting data standards and requirements on quality, update frequency or when data should be made available externally. Above all, data governance intends to standardise the usage of data across departments and ensure it is possible to get a grip on what is going on in the data ecosystem.

The city of Bergen for example assigns different roles, such as the "product owner", who is in charge of the infrastructure like the data lake or a use case on top of it. A very interesting role is the one of "data owner": this person oversees the usage of a certain type of data (such as human resources data from city employees) and when the data are involved in a certain use case or system, they can decide who gets access to the data.

While many private companies have been practising data governance for some time, the GDPR has been an enabler for cities to apply some aspects. For example, the not-for-profit aNewGovernance was established in 2018 to advocate personal data spaces, including one on cities and mobility. In addition, NACTO has put forward a set of mobility data governance principles for handling sensitive data centred around storage, sharing, access and oversight. Data governance is also one of the key aspects of the European data strategy and a data governance act was proposed by the European Commission on 25 November 2020.

3.3.3 Integrated approach to data management

The value of mobility data is not limited to mobility functions but can also serve other functional areas within the city administration (and beyond) and vice versa. While this data value is acknowledged by cities, creating the policies, structures and procedures to enable more systematic data sharing within an administration and across public agencies is proving to be a challenge, as acknowledged by Rob Roemers (STIB/MIVB): "*Brussels is fragmented, each administration has its own priorities,"* and Sergio Fernandez Balaguer (EMT Madrid): *"Quite often, data sharing is not a part of internal procedures. The importance of sharing data and open data policies is acknowledged, but from a practical point of view there are barriers, even between departments, as data management mostly depends on transversal IT departments. This is changing slowly over the years."*

André Ormond, Ormond Consultoria e Treinamento:

"Mobility data are very important for other areas: economy, commercial purposes, etc. If the government has these data, they can conduct their own studies with them. I have worked with a bank to use mobility data to make future studies on GDP."

Access to the data of other departments and agencies has nonetheless been facilitated by open data, whereby datasets from one department can be drawn down by another department through the open data portal. This is precisely what the Brussels public transport operator STIB/MIVB is doing with traffic flow data, which it takes from the Brussels Region's open data platform and integrates into its bus and tram fleet system management as context data. Integrating these data enables the operator to understand whether bus or tram delays are caused by traffic. In future, STIB/MIVB would like to get the data directly from the data feed rather than via an open data link; it is promoting that every new data pipeline built should incorporate data sharing.

Rob Roemers, STIB/MIVB:

"To improve inter-agency data sharing, STIB/MIVB is promoting that all new data pipelines incorporate data sharing to overcome the ad hoc situation today."

Lisbon is addressing the data fragmentation issue through the creation of its urban data-sharing platform (cf. Lisbon case described in section 2.7), which contains more than 300 datasets from across the city administration and other public agencies (notably the emergency services) and is the foundation of its integrated operations centre that opened in 2020.

As regards data that are not public, Bergen's aim is to cross-departmentally implement the innovative data-sharing roles between city departments, as described in section 3.3.2. Lisbon and STIB/MIVB have found ways of providing access to non-public data to universities, which desperately need these real-life data for research purposes. STIB/MIVB can recruit students on a temporary basis, which means they are covered by the STIB/MIVB non-disclosure agreement. Students can then access the data, following approval by the organisation's data privacy officer. Only the research findings leave STIB/MIVB, never the data. Through its urban data lab (cf. Lisbon case in section 2.2.6), Lisbon can share non-public data with researchers following the signing of a non-disclosure agreement by the respective institute's rector.

3.4 EU policy framework

The European Union is making use of various instruments (legislation, standardisation, guidance, etc.) to increase the availability and accessibility of mobility data with the aim of creating economic value from these data and contributing to mobility goals. This section will provide an overview of the most important instruments that are relevant for cities and give a taste of the broader EU data policy.

3.4.1 The EU ITS Directive

The ITS Directive 2010/40 EU was adopted in 2010 with the aim of accelerating and coordinating the deployment of intelligent transport systems (ITS) and promoting continuity of services across borders. It provides a framework and mandate for the European Commission to adopt specifications and standards in four priority areas: traffic and travel information, freight, road safety and vehicle-infrastructure connectivity.

The Directive has led to six Commission delegated regulations, including two that have particular relevance for city authorities: real-time traffic information (RTTI) (2015/962, 18 December 2014) and multimodal travel information services (MMTIS) (2017/1926, 31 May 2017). A third piece of legislation worth mentioning but not described further in this section is the safety-related traffic information (SRTI) delegated regulation (2013/886, 15 May 2013), which requires all service providers (public and private) to share data about safety-critical events. A data-sharing ecosystem was launched in December 2020 through the Data for Road Safety partnership, involving vehicle manufacturers, traffic information service providers, automotive suppliers and public authorities (mainly national level).

The real-time traffic information and multimodal travel information services regulations are designed to improve access to the wide range of road, traffic and travel data and multimodal service data in view of providing better information services to users. In essence, they specify the types of data to be made available in a standardised form and where they should be accessible – on the National Access Point (NAP). These regulations therefore constitute a form of mandatory sharing of data that are available in machine-readable format, from the data generator/holder, which may well be a city authority, national authority or private company, to the NAP.

It is important to note that the regulations only currently apply to data that already exist in a machine-readable format, meaning that it is not required to create data or make information machine-readable. Furthermore, the legislation does not require open (i.e. free of charge) data and therefore does not interfere in the commercial aspects of data exchange. In reality, since most public authorities practise open data, much of the public authority data made available on the NAPs is free of charge.

The RTTI delegated regulation mandates the publication of approximately 40 types of standardised road and traffic data on a National Access Point to enhance digital maps and driver information services, including road attributes, road works/closures and real-time traffic conditions. The prescribed data format is DATEX II, which is a data exchange standard with traffic management origins – it was designed to enable traffic data exchange between traffic control centres, in particular across borders, and has since become the transport industry standard for road and traffic data. The regulation currently applies to the TEN-T network, other motorways and designated priority zones only; however, its application to other roads (including urban and rural roads) is under consideration in a potential revision of the regulation (as of January 2021). A revision could therefore place data-sharing obligations on city authorities. In addition to a geographic extension, the potential revision may include the mandatory creation of specific machine-readable datasets as well as a data type extension, in particular urban vehicle access regulation (UVAR) data, recharging/fuelling station data and in-vehicle data.

There is now a concerted effort, sponsored by the European Union, to “urbanise” DATEX II within the DATEX II community and the European standardisation body CEN. European projects are also supporting this effort, including the UVAR Box project, which is building a framework to enable city authorities to transform data about urban vehicle access regulation schemes (e.g. low-emission zone and road pricing) into a machine-readable and standardised (DATEX II) form.

The MMTIS delegated regulation covers the main modal sectors (air, road, rail and maritime) and all types of services (scheduled, demand-responsive and personal) to enable improved multimodal travel information services to users. All the static data listed in the annex to the regulation are mandatory, meaning that where this dataset is available in a machine-readable format, it has to be standardised and made accessible via the National Access Point. Data types covered include public transport routes, timetables, fares and stops/stations, bike-sharing services and pedestrian and cycling infrastructure (bike network, bike-sharing stations), etc. It therefore concerns the operators of mobility services, many of which are operating under a public service obligation contract to a passenger transport authority (which city authorities are typically a part of) as well as the infrastructure managers, which are mainly city authorities.

By way of example, data about Lille’s publicly funded bike-sharing service V’Lille is available on the French NAP and the Brussels region’s pedestrian network can be found in digital form on the Belgian NAP. In contrast to the previous regulation (RTTI), MMTIS is not limited geography-wise and it applies to both the public and private sector. Its mandate is restricted to static data, although the European Commission is currently consulting with Member States and stakeholders about a possible regulatory revision to mandate real-time data and the resale of transport tickets by third parties (as of January 2021).

Should a revision go ahead, it is expected to introduce more competition in the provision of mobility information and booking services. In terms of implementation, the obligation is on the Member State (or competent administration) to put in place the mechanisms (support structure, data converters such as the French GTFS-NeTEx conversion tool, laws such as the French mobility law 2019-1428, etc.) to enable data generators/holders (including city authorities) to comply with the regulation.

Given the strong European policy push to improve the availability of mobility data (through the NAPs) and to make it more easily accessible (through standards), coupled with the general digitalisation drive, public authorities are encouraged to:

- Include in relevant system and service tenders clauses related to (i) data access/ownership and (ii) generation of data in formats that are compatible with the delegated regulations (specifically DATEX II for traffic data and NeTEx/SIRI for multimodal transport data). For instance, if procuring a new road works management and information system, specify that the data output should be in DATEX II in addition to locally required formats. The cost of generating data in an additional format is far lower if included in the system procurement phase than doing so retroactively.
- Move progressively towards the digitalisation of the infrastructure (i.e. putting it into a standardised, machine-readable format), particularly road links and attributes and road/traffic regulations. There are many initiatives and projects underway around Europe, including the UK government-sponsored digitalisation of traffic regulation orders and the European urban logistics digital twin project LEAD.

3.4.2 The National Access Points

The MMTIS and RTTI delegated regulations require the creation of a National Access Point (NAP) in every Member State where the data mandated by the regulations must be made available. The purpose of the NAP is to make it easier for data users to discover the range of data available. The regulations do not specify what form it should take. In some cases, the NAP builds its own data storage, as is the case in France where more than 300 transport datasets are already available to download from its platform. In other cases, such as Belgium, the NAP may simply offer an inventory of all the data that are available and provide a URL link to the dataset. Another model is the German Mobility Data Marketplace, which offers a virtual marketplace for data providers and customers to enter into a data-sharing agreement.

The construction of the NAPs is still underway in some countries. The NAP and National Body harmonisation group created under the EU-EIP project is facilitating cooperation and coordination among the NAPs on a range of issues, including the harmonisation of metadata. Other data challenges may be country-specific, such as the absence of data harmonisation among Belgian bus operators: each bus operator has adopted a different set of station codes in its fleet planning system, meaning that the code given to a specific station is different from one operator to the next. These issues often come to light once data are being prepared for the NAP and demonstrate the challenges of moving towards a harmonised way of doing things.

In some cases, a specific mobility dataset may well be held on both the local open data portal as well as on the NAP. The main difference is typically in the format of such data, as the NAP requires the dataset to be published according to the standard described in the delegated regulation. For instance, V'Lille bike-sharing data are available on the French NAP in GBFS and also on the Lille open data portal in their original form. Once fully operational, the NAPs should bring greater transparency to the mobility data sector and hopefully lead to more data exchange, for the benefit of the public and the private sector and users.

Among some commercial data generators there appears to be a need for reassurance that the obligation to publish data on the NAPs will not lead to a loss of data revenue. Armis for instance is supporting motorway concession holders in defining terms and conditions for data sharing on the NAP. The delegated regulations do not actually require open data (as mentioned in section 3.4.1) and do not interfere in any way in commercial agreements. The purpose of the NAP is simply to make data easier to find. If there are any technical gaps and development to take into account to publish the data (in the field of standardisation of data formats, interface, exchange protocols, etc.), Edoardo Felici points out that the European Commission is addressing these through programme support actions and specific funding to support the community.

3.4.3 Other EU legislation relevant to mobility data sharing

Among the other European legal obligations related to mobility data (sharing), the most important one is the Directive on open data and the re-use of public sector information, which superseded previous legislation setting rules on the re-use of public sector information. As its name suggests, the 2019 directive has extended its remit from setting information re-use rules to mandating open data. This directive empowers the European Commission to adopt implementing acts in six high value data areas, one of which is mobility. Once a mobility implementing act is adopted, public authorities (including city authorities) will be compelled to open up, through an API and without charge, the datasets listed in the act. It is not clear whether this piece of legislation would apply to data that are procured from a commercial party and are therefore subject to strict re-use rules (cf. section 3.3.1). No implementing act has yet been adopted (as of November 2020) and therefore the types of mobility data are not yet known.

3.4.4 EU data strategy

March 2020 saw the launch of the EU data strategy, which lays out the European Union's vision for creating a European data space (i.e. a single market for data) and enabling Europe to become a leading global player in the data economy. Through legislation and governance frameworks, the strategy will tackle data availability, standards, tools and infrastructure as well as data skills. A European data governance act was proposed by the European Commission at the end of 2020 with the aim of enhancing the availability of data for use by increasing trust in data intermediaries and by strengthening data-sharing mechanisms across the European Union.

The data strategy proposes the development of nine common European data spaces in strategic sectors and domains of public interest, including one on mobility. Each data space is expected to build data pools, technical tools and infrastructure for data use and exchange as well as governance mechanisms. They may also give rise to sectoral legislation and policies to stimulate the use of data and demand for data services. Regarding the mobility data space, it will cover all modes (road, rail, maritime and air) and is expected to involve a review of different legislation. This data space will build heavily on the data infrastructure and governance foundations already laid by the ITS Directive (cf. section 3.4.1) and will include a review of this piece of legislation.

The slow development of Europe-wide standards in the fast-growing data sector was noted by the city of Antwerp. It has led to situations where non-European standards are adopted, such as GBFS or MDS in the new mobility services domain, and to the adoption of a standards-agnostic approach in the case of Antwerp. Kasia Bourée from Data4PT points out the advantage of CEN (European Committee for Standardization) standards (where available) over industry or de facto standards: *"DATEX II and SIRI are CEN standards, but this is not the case for GBFS, for example. This is an important message to convey to EU cities: CEN standards offer sustainability and governance, take into account multiple European needs and are integrated into EU law, whereas some de facto standards like GTFS and GBFS do not present these advantages."*

3.4.5 General Data Protection Regulation (GDPR)

Putting people in control of their data

The new EU regulation on data privacy, the General Data Protection Regulation (GDPR), entered into force in May 2018 and has had a profound effect on many aspects of public life and business. It provides individuals with enforceable rights to access, rectify, erase and object to personal data held by others and provides for greater transparency and data portability (personal data can be shared across different services to avoid data silos). Consequently, it has forced public and private organisations to thoroughly review and adapt their data collection, usage and storage processes and to communicate clearly to their constituents/customers what personal data are held, how they are processed and how they can be accessed.

Personally identifiable data

The notion of personal data in the GDPR also includes personally identifiable data, which refers to non-personal data that, when manipulated, could lead to the identification of an individual. According to NACTO, *"geo-spatial data is, or can become, personally identifiable information through recognisable travel patterns and combined with other data."* An example is the open data on London's bike-sharing scheme that reveals great insights, but also unwillingly disclosed personal details of its users. Geo-spatial data, aka location/movement data in the mobility domain, is generated by connected devices such as a mobile phone or a vehicle. Location data, collected by service providers (telecom, mobility or public transport operators for example) and combined to produce individual trips, is widely sought by transport authorities for the purpose of managing and monitoring the movement of people and goods. It appears in many of the data-sharing cases described in chapter 2. To avoid any potential conflict with the GDPR, city authorities appear to be moving towards acquiring aggregated data based on individual trips, as opposed to raw data. When in doubt, the advice of the administration's data privacy officer should be sought. According to the OMF, *"There are no US standards on what movement data are personal data. There is no clear line, therefore most cities treat them as personal data."*

Getting cold feet

The ambition and unprecedented nature of the GDPR has created a sense of unease among many organisations and in some cases a reluctance to work with data that may be considered personal. Armis confirms that it is a sensitive issue and can be viewed in a black and white manner. The GDPR does not prevent the collection and processing of personal (or personally identifiable) data but rather stipulates the rules and principles that must be abided by in any organisation working with such data. By way of example, the recent coronavirus contact tracing apps set up in many countries have been developed in full compliance with the GDPR, which shows the flexibility of the regulation.

Pedro Barradas, Armis:

"There is always a way to accommodate GDPR, and ways to go deeper, wider and further and still live up to GDPR."

Section 3.2.2 provides insights into the ways in which cities are dealing with data privacy, especially in the context of the GDPR.

3.5 Future considerations and trends

This document was finalised in early 2021 and efforts have been made to include best practices from cities at the forefront from different parts of the world. These practices should remain relevant for a number of years, as many other cities are still starting out with similar efforts. In order for this document to remain relevant, this section includes trends that have the potential to become dominant in the next decade. These trends have been identified in two ways by the authors. Firstly, the seeds of innovative systems that have already begun to sprout in the mobility landscape, but that are still in the embryonic stage. Secondly, some trends becoming dominant in other sectors will most likely expand to the mobility domain in the years to come. These trends include:

Use of app-based data collection: As indicated in section 1.6 ("Data sources"), many popular mobile phone applications collect large amounts of location points without the active intervention of the user (as opposed to active tracking apps like Strava and Runkeeper). These datasets hold extremely detailed information on the behaviour of smartphone users, including mobility patterns. If a completely transparent and legal marketplace can be built for such datasets while considering and observing privacy regulations, this data source has the potential to replace many existing ways of collecting data on mobility behaviour.

Digitisation and automation of policies: While some cities are already publishing their policies as open data, this trend goes a step further and publishes policies through a machine-readable interface. This way, any provider or system can obtain an instant view of policies and remain up to date if they became more dynamic (such as stricter emission standards when the air quality drops below a certain level). Examples to be noted today are the MDS policy API (cf. chapter 2) and the UVAR Box project mentioned by Armis.

Towards an API economy: Application programming interfaces (APIs) are computing interfaces for orderly (and standardised) interactions between IT systems. They overcome the need for custom-built integrations for each new function or connection. As systems become increasingly integrated, APIs will become key for connecting internal systems too. An example is Antwerp's NXTMobility platform that will also make its route planner available through an API.

In-vehicle data: The data collected by sensors and cameras in vehicles will become very important data sources. Today the data can already be acquired through intermediaries like Otonomo or IBM's IoT marketplace for connected vehicles. Interesting work in this field is being done by the data task force on safety-related traffic information.

More specific policies: While cities are currently shaping policies around modes of transport (one policy for cars, one for bikes, etc.), more specific mobility data will enable policies to be created based on the type of user (disabled person, inhabitant, visitor, etc.) and why this user is travelling and using this specific mode of transport. Early signs of this are urban vehicle regulations, information services and pricing policies (discounts for certain groups) that do not treat all cars equally.

Predictive data: As cities become acquainted with real-time data, this will shift towards predictions of mobility patterns in the near future, by using artificial intelligence and predictive models, and data from navigation systems and planned or ordered trips from on-demand services.

The rise of intermediaries: As extensively illustrated throughout this document and specifically in section 2.3, the number of integrations between systems is rising exponentially. As cities often lack technical expertise but share similar needs, a new market has already emerged with companies integrating systems or distilling data into information that is easy for cities to use and understand.

Privacy and data ownership: As described earlier in this chapter, the GDPR has introduced an emerging awareness of the protection of sensitive data, both by public authorities and leading companies like Apple. This awareness is expected to grow, further legislation will no doubt be put in place (such as the proposed European data governance act) and new governance models that put the user at the centre should flourish, like MyData and aNewGovernance mentioned above.

Citizen science and IoT: While cities had the monopoly on mobility data collection for a long time, this has changed with the advance of cheap sensors that can easily be connected to the internet. Citizens are already connecting data on traffic volumes and speeds through Telraam (cf. section 2.8) and dangerous spots for cycling through projects like Ping if you care. Cities will have to adapt to this technically and organise themselves in order to enter into discussions with the public based on such data, which might challenge policies.

Digital twins: Many projects are currently emerging around digital twins. This can be a digital 3D replica of the city, where the effects of measures and policies on air quality, noise, congestion and liveability, etc. can be visualised. An example is Virtual Singapore. A much more modest example is Remix, where authorities and people can visualise street redesigns.

Semantics over standards: Although many discussions have centred on data standards and putting them in place, semantic interoperability might become predominant. This basically means that systems do not just exchange data, but they actually understand what they are processing. Examples are the Semantic Web and Linked Open Data, where systems share the same understanding of a bicycle parking space or bus stop. The SPRINT project and yearly Sem4Tra workshop are working on this.

Marketplaces: With cities being given more choice on where they can acquire their data, marketplaces are emerging like HERE's neutral server and Germany's Mobility Data Marketplace. The European regulations on National Access Points could reinforce this trend.

Chapter 4: A possible roadmap for data acquisition

This chapter concludes the study by summarising the conclusions and recommendations from the previous chapters in a roadmap that can be used by urban authorities. The intention is not to draw up an exhaustive checklist that offers detailed guidance. Rather, this chapter is intended as a summary and starting point for cities, based on the literature review and interviews conducted, combined with the experience of the authors.

1 WHY – Always start from the urban strategy: the strategic mobility objectives or sustainable urban mobility plan (SUMP), and the data strategy of your city. Do you need data for policy planning, monitoring and evaluation or operational mobility management? See section 1.2 “A data strategy”, which references some examples and stresses that data are not a solution if there is no clear problem statement.

2 WHAT – Determine which data type you need to contribute to the strategic objectives set. Do you need traffic volumes, numbers of cyclists, insights into parking usage, air quality information or behavioural changes? Section 1.5 contains some data types and examples of how they can be applied.

3 WHAT – What data sources are available to provide the data types needed? For example, if you need origin-destination information, what data sources can deliver this? Section 1.5 contains a table linking data types to data sources and section 1.6 describes a number of common data sources used in the field of urban mobility.

4 WHAT – What data do you already own within the city? Many experts interviewed stress it is important to create an inventory of the data that are gathered through different systems the city already owns like cameras for enforcing speed or low-emission zones, parking guidance systems or public transport ticketing data. It is also worthwhile looking to other departments and consulting with the regional and national authorities, as they might already collect and share such data, for example through a National Access Point (cf. section 3.4.2). It might be beneficial to access existing data or open up silos rather than acquiring the same kind of data again.

5 LESSONS LEARNED – Look for cities, networks or organisations that have rolled out similar strategies, measures or projects and enquire about the lessons learned. Section 1.2 contains a number of organisations, networks and knowledge bases you can refer to.

6 WHAT EXACTLY – Define the exact scope of your project. It is recommended to make it very specific, small and feasible. Once the objectives have been met, it will be possible to extend the buy-in, improve and scale up. Define the budget for the project – section 1.7 is dedicated to this. The bottom line: with small budgets, a proof of concept can be built that can demonstrate the added value for a city and its people.

7 HOW – Once you have defined what data you need, think about how you will acquire them. Straightforward procurement is an option, but there are many others as well. Chapter 2 proposes a number of data acquisition models, along with lessons learned, recommendations and 17 cases from different parts of the world.

8 DATA/INFORMATION – When acquiring data, define what you actually need. Would you prefer to receive preprocessed, aggregated *information* or is it raw sensor *data* that are needed? Read section 1.3 “From data to wisdom” and the paragraph “Data aggregation and anonymisation” in section 3.2.2 for more information.

9 CAPACITY – What in-house expertise does the city have, both on urban mobility and on data processing and interpretation? Consider finding external expertise, but ensure the capacity is available to define the specifications and evaluate the work done by third parties. Make sure that city experts are involved and the organisation is learning from the expertise hired. More on this topic in section 3.3.1 "Capacity and expertise".

10 WHO – Find parties to work with and acquire data from. Try building an ecosystem of partners for the long term rather than always sticking to a customer-supplier relationship. The public should be at the centre of the ecosystem, supplemented with data providers, aggregators, universities, startups, and others. Make things redundant: do not depend on one data source or partner for important processes or data. See the paragraph "Building an internal and external data ecosystem" in section 3.3.1.

11 CONTRACTS and AGREEMENTS – Check what agreements need to be in place, what legislation is in force and what paragraphs to insert, specifically on data ownership, data usage and privacy. Check section 3.2.1 for examples and recommendations. Consider adding standard paragraphs on data to all contracts and tenders.

12 EVALUATE – Check that the data, information and project meet the strategic objectives. Check with experts and citizens if the results make sense. When the outcome is or could be satisfactory and useful, reiterate, adapt and scale up. Replace temporary or improvised processes and systems with solid and sustainable variants.

13 MAINTAIN – If new processes have been created, platforms have been built or data are being acquired as a service, transfer the product of the project to the operational organisation and ensure maintenance, support and monitoring are in place.

Annex I: List of case studies

The table below contains a list of all the cases referenced in this study. All the cases are plotted against the seven data acquisition models discussed in chapter 2, which are shown in the columns. The bold Xs mark the model for which the case is used as an example, but all cases relate to some extent to some of the other models as well.

	Traditional procurement	Intermediaries	Financially compensated partnerships	In-kind partnerships	Mandatory data sharing	Collaboration between authorities	Crowdsourcing
CHAPTER 2							
City of Paris: data procurement	**x**	x					
Gothenburg and Sweden: procuring floating vehicle data	**x**					x	
Traffic Management as a Service		**x**					x
Bicycle data-sharing platform (CROW) and Vianova		**x**			x	x	
The Data4PT project	x	**x**	x			x	
Pedestrian flow data procurement in Paris	x	x					
Parking rights database		x	**x**				
Brussels “Muntstroom”: people flow data			**x**			x	x
Waze for Cities		x		**x**		x	x
The Netherlands: TOMP and CDS-M standard		x		x	**x**	x	
Micromobility data in Paris		x			**x**		
Antwerp data-sharing policies		x	x	x	**x**	x	
Open Mobility Foundation and MDS		x			**x**	x	
The National Road Traffic Data Portal	x	x				**x**	
Lisbon’s urban data-sharing platform				x		**x**	

Telraam/WeCount: citizen-led counting		x					x
OpenStreetMap				x		x	x
CHAPTER 3							
Data clauses in off-street parking concession contracts			x		x		
Antwerp: generic data licence and service level agreement	x	x	x		x		
Data standardisation clauses in agreements on public transport	x	x	x		x		
Smart Flanders: sample clauses on (open) data for agreements	x	x	x		x		
Data portability by MyData	x	x				x	x

Annex II: Literature list

Recommended reads

- OECD/International Transport Forum: Data-Driven Transport Policy (2016): https://www.itf-oecd.org/sites/default/files/docs/data-driven-transport-policy.pdf

- NACTO: Policy guide – Managing mobility data (2019): https://nacto.org/wp-content/uploads/2019/05/NACTO_IMLA_Managing-Mobility-Data.pdf

- Bloomberg CityLab: For Ride-Hailing Drivers, Data Is Power (2019): https://www.bloomberg.com/news/articles/2019-08-22/why-uber-drivers-are-fighting-for-their-data

- Bloomberg CityLab: This City Was Sick of Tech Disruptors. So It Decided to Become One (2020): https://www.bloomberg.com/news/articles/2020-02-21/as-l-a-plays-tech-disruptor-uber-fights-back

- World Business Council for Sustainable Development (WBCSD): Enabling data-sharing – Emerging principles for transforming urban mobility (2020): https://www.wbcsd.org/Programs/Cities-and-Mobility/Transforming-Urban-Mobility/Digitalization-and-Data-in-Urban-Mobility/Resources/Enabling-data-sharing-Emerging-principles-for-transforming-urban-mobility

- The New York Times: Twelve Million Phones, One Dataset, Zero Privacy (2019): https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html

- Alali, E.: Tender Irony – Barrier to Innovation is Innovation? (2017): https://medium.com/tech-sojourna/tender-irony-barrier-to-innovation-is-innovation-1594752d4dd4

- City of Copenhagen: City data exchange – lessons learned from a public/private data collaboration (2018): https://cphsolutionslab.dk/media/site/1837671186-1601734920/city-data-exchange-cde-lessons-learned-from-a-public-private-data-collaboration.pdf

- The World Bank – Open Traffic Data to Revolutionize Transport (2016): https://www.worldbank.org/en/news/feature/2016/12/19/open-traffic-data-to-revolutionize-transport

Full reading list

- MOMENTUM project: deliverable D3.1 – Data Inventory and Data Quality Assessment (2020): https://h2020-momentum.eu/wp-content/uploads/2020/03/MOMENTUM-D3.1-Data-Inventory-And-Quality-Assessment-Issue-1-Draft-5.pdf

- CATAPULT: The Transport Data Revolution (2015): https://static1.squarespace.com/static/565ebf94e4b0f2a77109071e/t/56631604e4b0672d1446accf/1449334276748/Intelligent+Mobility+Data+Report+-+Final.pdf

- The Urban Mobility Daily: U.S. Micromobility Battleground: Data, Privacy, and Regulation (2019): https://scott-shepard-20711.medium.com/u-s-micromobility-battleground-data-privacy-and-regulation-18d3c22686b8

- Nature/de Montjoye, YA., Hidalgo, C., Verleysen, M. and Blondel, V.: Unique in the Crowd: The privacy bounds of human mobility (2013): https://www.nature.com/articles/srep01376

- Slate: Cities Can See Where You're Taking That Scooter (2019): https://slate.com/business/2019/04/scooter-data-cities-mds-uber-lyft-los-angeles.html

- Workday Technology: Data-as-a-Service – a Global Data Platform (2018): https://medium.com/workday-engineering/data-as-a-service-a-global-data-platform-ba8d1572d268

- Urban Mobility Company: Why the Urban Mobility Data Debate Matters to Public Transportation (2020): https://www.autonomy.paris/why-the-urban-mobility-data-debate-matters-to-public-transportation/

- Urban Mobility Company: In the Quest For Open Data, How Do Cities and Transport Operators Move Forward? (2019): https://eukalypton.com/en/2019/04/19/in-the-quest-for-open-data-how-do-cities-and-transport-operators-move-forward-2/

- Consumer Reports: The Cameras in Your Car May Be Harvesting Data as You Drive (2020): https://www.consumerreports.org/automotive-technology/the-cameras-in-your-car-may-be-harvesting-data-as-you-drive-a3473812015

- LEMO project: Leveraging big data for managing transport operations – Report on Economic and Political Issues (2018): https://static1.squarespace.com/static/59f9cdc2692ebebde4c43010/t/5b88ef82b8a0457103352e09/1535700879604/20180831_D2.1_Report+on+Economic+and+Political+Issues_LeMO.pdf

- Rock Your Data/D'ma: Getting Started with Data Lake (2019): https://medium.com/rock-your-data/getting-started-with-data-lake-4bb13643f9

- Chowdhury, M., Apon, A. and Dey, K.: Data Analytics for Intelligent Transportation Systems (2017): https://www.sciencedirect.com/book/9780128097151/data-analytics-for-intelligent-transportation-systems

- Forbes: How Blockchain is driving new opportunities in mobility (2020): https://www.forbes.com/sites/forbestechcouncil/2020/03/11/how-blockchain-is-driving-new-opportunities-in-mobility/

- MoTiV project: Studying human mobility through cellular data: where we are and future opportunities (2018): https://motivproject.eu/news/detail/studying-human-mobility-through-cellular-data-where-we-are-and-future-opportunities-1.html

- UK Department for Transport: Local transport data discovery report (2018): https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/730787/local-transport-data-summary.pdf

- Populus: A practical guide to mobility data sharing and cities (2020): https://static1.squarespace.com/static/59bec4a212abd994ed2743ac/t/5ec35fd76cf43a2f5020fb51/1589862398665/Populus_MobilityDataSharingAndCities.pdf

- TravelSpirit: An open future for cities – Preparing cities for the necessary transformation and organisational changes needed for an open future (2019): https://travelspirit.io/wp-content/uploads/2019/06/TravelSpiritWhitePaper8-An-Open-Future-for-Cities-June-2019.pdf

- Los Angeles Magazine: Waze Hijacked L.A. in the Name of Convenience. Can Anyone Put the Genie Back in the Bottle? (2019): https://www.lamag.com/citythinkblog/waze-los-angeles-neighborhoods/

- European Commission – JRC Technical Reports: JRC Digital Economy Working Paper 2018-06 – Access to digital car data and competition in aftersales services (2018): https://ec.europa.eu/jrc/sites/jrcsh/files/jrc112634.pdf

- US Department of Transportation: Travel Time on Arterials and Rural Highways: State-of-the-Practice Synthesis on Rural Data Collection Technology (2013): https://ops.fhwa.dot.gov/publications/fhwahop13029/ch2.htm

- European Innovation Partnership for Smart Cities & Communities (EIP-SCC): Rethinking the city: using the power of data to address urban challenges and societal change – A guide for city leaders (2017): https://smart-cities-marketplace.ec.europa.eu/sites/default/files/EIP_Leadership_Guide.pdf

- Geo Awesomeness: What can SIM cards tell us about the tourism industry? (2019): https://geoawesomeness.com/telco-data/

- US Department of Transportation: Does Travel Time Reliability Matter? (2019): https://ops.fhwa.dot.gov/publications/fhwahop19062/fhwahop19062.pdf

- A European strategy for data (2020): https://ec.europa.eu/info/files/communication-european-strategy-data_en

- Traganos, K., Grefen, P., den Hollander, A. and Turetken, O.: Business model prototyping for intelligent transport systems: a service-dominant approach (2015): https://www.researchgate.net/publication/313580800_Business_model_prototyping_for_intelligent_transport_systems_a_service-dominant_approach?channel=doi&linkId=589f004d92851c7fb4bdbd5d&showFulltext=true

- National Cooperative Highway Research Program: Framework for Managing Data from Emerging Transportation Technologies to Support Decision-Making (2020): http://www.trb.org/main/blurbs/181365.aspx

- National Cooperative Highway Research Program: Guidebook for Managing Data from Emerging Technologies for Transportation (2020): http://www.trb.org/main/blurbs/180826.aspx

- Intelligent Transport: Three cities turning e-scooter data into infrastructure improvements (2020): https://www.intelligenttransport.com/transport-articles/110464/three-cities-turning-e-scooter-data-into-infrastructure-improvements/

- Esri: Mapping transit accessibility to jobs (2018): https://www.esri.com/arcgis-blog/products/arcgis-desktop/transportation/mapping-transit-accessibility-to-jobs/

- Tolouei, R., Psarras, S. and Prince, R.: Origin-Destination Trip Matrix Development: Conventional Methods versus Mobile Phone Data (2017): https://www.sciencedirect.com/science/article/pii/S2352146517308669

- Schneider, T.: Analyzing 1.1 Billion NYC Taxi and Uber Trips, with a Vengeance (2018): https://toddwschneider.com/posts/analyzing-1-1-billion-nyc-taxi-and-uber-trips-with-a-vengeance

- The Atlantic: Exploring the Curious Connections Between Traffic Deaths and NFL Home Games (2013): https://www.bloomberg.com/news/articles/2013-08-29/exploring-the-curious-connections-between-traffic-deaths-and-nfl-home-games

- Shlayan, N., Kurkcu, A. and Ozbay, K.: Exploring pedestrian Bluetooth and WiFi detection at public transportation terminals (2016): https://www.researchgate.net/publication/311919537_Exploring_pedestrian_Bluetooth_and_WiFi_detection_at_public_transportation_terminals

- van der Loop, H. et al.: Validation and usability of floating car data for transportation policy research (2019): https://www.kimnet.nl/publicaties/publicaties/2019/5/23/validation-and-usability-of-floating-car-data-for-transportation-policy-research

- Future Place Leadership: The effect of pedestrianisation and bicycles on local business (2017): https://futureplaceleadership.com/wp-content/uploads/2017/05/Tallinn-High-Street-Case-studies-Future-Place-Leadership.pdf

- US Federal Highway Administration: Traffic Detector Handbook (2006): https://www.fhwa.dot.gov/publications/research/operations/its/06108/02.cfm

- Li, T., Sun, D., Peng, J. and Yang, K.: Smart Card Data Mining of Public Transport Destination: A Literature Review (2018): https://www.researchgate.net/publication/322510689_Smart_Card_Data_Mining_of_Public_Transport_Destination_A_Literature_Review

- The World Bank Group: Big Data and Urban Mobility – Background Document (2014): https://www.worldbank.org/content/dam/Worldbank/Feature%20Story/mena/Egypt/Egypt-Doc/Big-Data-and-Urban-Mobility-v2.pdf

- Forbes: When Will Cities Begin To Monetize Their Residents' Data? (2018): https://www.forbes.com/sites/kalevleetaru/2018/07/19/when-will-cities-begin-to-monetize-their-residents-data/?sh=1ac03a146618

- Sorbal, T. et al.: Semantic integration of urban mobility data for supporting visualization (2017): https://www.sciencedirect.com/science/article/pii/S2352146517303873

- KPMG Norway: Automotive Data Sharing Report (2020): https://home.kpmg/no/nb/home/nyheter-og-innsikt/2020/11/automotive-data-sharing.html

- Rescot, R.: Feasibility of Bluetooth Data as a Surrogate Analysis Measure of Traffic (2011): https://kuscholarworks.ku.edu/bitstream/handle/1808/8381/Rescot_ku_0099D_11736_DATA_1.pdf

- CLARS/Sadler Consultants Ltd.: (The impact of) Urban Access Regulations in Europe (2017): https://urbanaccessregulations.eu/low-emission-zones-main/impact-of-low-emission-zones

- EIB Technical Note on ITS Procurement for Urban Mobility

Annex III: Data sharing initiatives in urban mobility

1. Waze for Cities

A programme set up by Waze to exchange data with cities. Public authorities get free access to Waze data in return for sharing data on planned road closures and roadworks. The initiative exceeds mere data exchange – a community of cities and interested parties is built that exchanges best practices on data sharing, but also on urban mobility in general.

- https://www.waze.com/ccp

2. Traffic Management as a Service (TMaas)

The city of Ghent is building a virtual traffic control centre as a service, completely powered by using third-party data providers. The idea of the TMaaS concept is that cities will be able to subscribe to the platform and immediately get access to data about their city and tools to analyse these data, manage traffic equipment and send notifications to residents through personalised messages.

- https://uia-initiative.eu/en/uia-cities/ghent

3. Mobility Data Specification (MDS)

A significant number of cities around the globe have adopted MDS, a standard originally produced in Los Angeles for cities to exchange data with mobility service providers (currently focused on dockless scooters, bike sharing and car sharing but has the capacity to expand to other transportation modes and means). The governance of the MDS has been transferred to the Open Mobility Foundation.

- https://www.openmobilityfoundation.org/

4. Socrates 2.0

A number of governments and service providers/original equipment manufacturers are collaborating to give personal route advice to car drivers that is composed and approved by traffic control authorities. This means that public governments and private navigation services are joining forces to give route advice that serves the overall traffic system, taking residential areas into account. This strongly contrasts with current systems that are mainly built to guide individual users as fast as possible from A to B. Socrates 2.0 is a partnership of public authorities and commercial navigation services (in-car GPS systems and navigation apps).

- https://socrates2.org/

5. Worker Info Exchange

"Worker Info Exchange is a non profit organisation dedicated to helping workers access and gain insight from data collected from them at work usually by smartphone. Whether you are an Uber driver or a Deliveroo rider, we aim to tilt the balance away from big platforms in favour of the people who make these companies so successful every day – the workers."

- https://workerinfoexchange.org/

6. Telraam/WeCount

How can people help cities collect data? Cameras attached to windows count (and measure the speeds) of cars, cyclists and pedestrians and share the data through a unified API. Some cities have (partly) subsidised the purchase of such devices by residents in order to gather mobility data.

- https://www.we-count.net/

7. Parking rights database

Some cities (in the Netherlands this is organised on a national level) have introduced a parking register allowing market parties to sell parking tickets on their territory without having to develop applications themselves. The external providers place the licence plates of their plants in a city database and receive financial compensation for each ticket sold.

- https://nationaalparkeerregister.nl/home.html
- http://ip-mobile.be/ons-aanbod/bpr/?lang=nl

8. Dynamic Anonymous Bicycle Register

In the Netherlands, a national cycling organisation has created an anonymous register so cyclists can register their bike and be contacted if their bike is stolen, found or parked badly – but without the sender knowing their identity.

- https://www.fietsberaad.nl/Kennisbank/Je-Swapfiets-terug-dankzij-het-Dynamisch-Anoniem-F

9. The National Road Traffic Data Portal (NDW)

The Netherlands has built a National Data Warehouse for traffic information and acquires/procures data on a national level so that it can be used by any city in the country. The initiative is a collaboration between 19 public authorities in the Netherlands. What are the advantages of this way of working? And how does it cooperate with cities?

- https://www.ndw.nu/

10. World Bank floating taxi data – Project Leapfrog

The World Bank set up a project to gather its own floating car data by collecting data from the smartphones of 500 000 drivers using a taxi hailing app called Grab, starting in the Philippines.

- https://www.worldbank.org/en/news/feature/2016/12/19/open-traffic-data-to-revolutionize-transport

11. Flemish Open City Architecture

The region of Flanders is working on an open city architecture allowing cities to easily connect and interconnect data sources.

- https://vloca.vlaanderen.be/

12. Talking Traffic

The Dutch government is finishing a €90 million cooperative ITS project to connect people and vehicles with infrastructure (mainly traffic lights).

- https://www.talking-traffic.com/en/

13. Mobilidata

The Flemish government is running a €30 million cooperative ITS project to connect traffic lights, vehicles, cyclists and pedestrians, including a policy tool based on data.

- https://mobilidata.be/en

14. DOVA (real-time public transport data)

The Dutch government built a platform to collect data from public transport providers. It processes these data to inform end-users and to follow up on the performance of public transport providers.

- https://www.dova.nu/ (in Dutch)

15. Velopark

Fietsberaad, a Flemish cyclist organisation, built a platform to centralise and standardise cycling data as Linked Open Data so that cities only have to edit bicycle parking information in one place and it can be imported automatically into other systems.

- https://www.velopark.be/en

16. MOMENTUM project

"The goal of MOMENTUM is to develop a set of new data analysis methods, transport models and planning support tools to capture the impact of these new transport options on the urban mobility ecosystem, in order to support cities in the task of designing the right policy mix to exploit the full potential of these emerging mobility solutions."

- https://h2020-momentum.eu/

17. BITS project

BITS is a project on bicycles and ITS. They have created an inventory on data and are building a Cycle Data Hub in order to centralise and open up data on cycling, like floating bike data, bike counters and camera detection for cyclists.

- https://northsearegion.eu/bits/

18. Mapillary

Mapillary is a collaborative platform for street-level imagery (best known in this field is Google Street View). Anyone can upload images, which are stitched together to create a street-level view of the world. Artificial intelligence can be applied to the images to detect street signs, etc.

- https://www.mapillary.com/

19. OpenStreetMap

OpenStreetMap is a collaborative project to create a free editable map of the world. Anyone can contribute to the map and mapmaking communities all over the world join forces to keep the map complete and up to date. The maps can also be used by everyone (including urban authorities) to display data on.

- https://www.openstreetmap.org/
- https://digitransit.fi/en/ (Helsinki's public transport route planner, based on OpenStreetMap)

20. MUV – Mobility Urban Values

"MUV levers behavioural change in local communities using an innovative approach to improve urban mobility: changing citizens' habits through a game that mixes digital and physical experiences. Rather than focus on costly and rapidly ageing urban infrastructures, MUV promotes a shift towards more sustainable and healthy mobility choices by engaging in a positive way local communities, local businesses, policymakers and Open Data enthusiasts."

- https://www.muv2020.eu/

21. National Access Points

"EU Member States are setting up their National Access Points to facilitate access, easy exchange and reuse of transport related data, in order to help support the provision of EU-wide interoperable travel and traffic services to end users."

- https://transport.ec.europa.eu/transport-themes/intelligent-transport-systems/road/action-plan-and-directive/national-access-points_en

22. Mobility-as-a-Service (MaaS) pilots

The Dutch government has set up a partnership with cities and private companies to roll out seven MaaS pilots to investigate the benefits and needs of rolling out MaaS on a larger scale. Data are key here, both for identifying transportation options and for the exchange of tickets and commercial agreements between MaaS providers and operators.

- https://dutchmobilityinnovations.com/attachment?file=7qczeMbWTcRrUzL2ExA8ug%3D%3D

23. StreamR

StreamR is a network that allows individuals or groups of users to share real-time data through a global peer-to-peer network. By using cryptography, the messages can be shared safely, and the publishers of the data can monetise their data and be rewarded in cryptocurrencies.

- https://streamr.network/case-studies/machine-witness

24. Urban Transport Data Exchange

"Working together, REEEP, the Institute for Transportation and Development Policy (ITDP) and the World Wide Web Foundation intend to harness the power of Linked Open Data to improve decision-making and planning to ensure low carbon, sustainable transport systems. The Urban Transport Data Exchange will initially build on ITDP's existing work in developing the Bus Rapid Transit Standard and sharing data on design characteristics of BRT systems in 36 cities."

- https://www.reeep.org/urban-transport-data-exchange

25. European Data Task Force

The European Data Task Force is a consortium of Member State vehicle manufacturers and service providers. They signed a memorandum of understanding, the basis of a trusted partnership based on the principle of reciprocity. Safety data will be offered in return for safety services.

- https://www.dataforroadsafety.eu/

26. Traffic Management 2.0

The Traffic Management 2.0 platform brings together 40 members from both the public and private sectors to focus on new solutions for advanced interactive traffic management.

- https://tm20.org/

27. 6Aika

The six biggest cities in Finland have joined forces to collaborate on smart city challenges, including mobility and data.

- https://6aika.fi/en/projects/

28. ITS Belgium parking work group

Stakeholders from all kinds of backgrounds (cities, service providers, public transport industries, the automotive industry, etc.) joined forces in a parking working group, which resulted in standard components on parking data that can be used for agreements with parking operators and concession holders.

29. Uber Movement

Uber aggregates the data from its ride-hailing app and shares them with a number of cities, together with a number of tools to interpret the data. The data are part of a partnership for the company to maintain relationships with cities around the globe. At this point, it is not clear if Uber expects anything in return for sharing the data.

- https://movement.uber.com/

30. Analysis of data from freight tolling

The Belgian regions have had a freight toll in place for a number of years, which requires vehicles over 3.5 tonnes to install an on-board unit to collect data on the location of the vehicles. The data are processed by a third party that also provides services to transport companies. This is an interesting case for a government to collect data on mobility patterns as a by-product of tolling.

- https://www.vlaanderen.be/en/mobility/kilometre-fee-heavy-goods-vehicles

31. The Synchronicity project

"Opening up a global market, where cities and businesses develop IoT- and AI-enabled services through pilots to improve the lives of citizens and grow local economies."

- https://synchronicity-iot.eu/project/active-travel-insights/

32. World Business Council of Sustainable Development (WBCSD) – Data Sharing Principles

"Data is a key enabler of the mobility transformation. With this project, WBCSD is bringing business together to agree on the principles of data-sharing that will enable sustainable urban mobility management across public and private stakeholders"

- https://www.wbcsd.org/Programs/Cities-and-Mobility/Transforming-Urban-Mobility/News/WBCSD-releases-foundational-data-sharing-principles-for-mobility

33. Urban Transport Data Analysis Tool

"The Data Analysis Tool for Urban Transport is a simple Excel-based tool that enables users to compare several urban transport related indicators in a city with similar indicators in peer cities. Such a comparison would allow users to identify areas where the city under study is performing well or is performing poorly. The output is a report presenting how the city is performing vis-a-vis peer cities with respect to a set of performance indicators."

- https://www.worldbank.org/en/topic/transport/publication/urban-transport-data-analysis-tool-ut-dat1

34. Mobility in Cities Database

"The Mobility in Cities Database (MCD) provides key insights into transport patterns and trends for more than 60 metropolitan areas worldwide. The wide range of indicators collected span areas ranging from demography, economy, traffic, transport infrastructure and mobility to public transport supply and demand data. The latest MCD publication gives a snapshot of 2012, the most recent year for which data is available, identifying groups of cities with similar performances. Building on the previous editions, the evolution of key urban indicators is shown, spanning a 20-year period."

- https://www.uitp.org/MCD

35. US Department of Transportation: ITS Data Access and Exchanges

A platform from the US Department of Transportation focused on data exchanges in the field of transportation. The partnership focuses on exchanging data and the platform offers several tools, for example:

- The ITS DataHub – a single point of entry to discover the US Department of Transportation's publicly available ITS research data, including connected vehicle data.
- The Secure Data Commons – a cloud-based analytics platform that enables traffic engineers, researchers and data scientists to access transportation-related datasets.
- The ITS CodeHub – a resource for the ITS community to discover open source code, software and more.

36. Flanders – OSLO vocabulary for mobility and Flemish Open City Architecture (VLOCA)

The Flemish government, alongside experts and companies in the field, has created a semantic vocabulary to exchange data in three fields: mobility policies, traffic signs and the occupation of public domain. Combined with the Flemish Open City Architecture, it will allow cities to freely exchange data through common tools and standards.

- https://data.vlaanderen.be/standaarden/kandidaat-standaarden/vocabularium-mobiliteit/vocabularium-mobiliteit.html
- https://www.imeccityofthings.be/en/projecten/vloca

37. OpenTraffic

OpenTraffic is a global data platform to process anonymous positions of vehicles and smartphones into real-time and historical traffic statistics. The initiative was started by the World Bank, among others, and uses fully open-source software.

- http://opentraffic.io

38. Open 511

Following the example of other formats like GTFS (for transit), Open511 proposes a specification for road incidents, construction and much more that matches open data criteria.

- https://www.open511.org/

39. NUMO Policy Data Tool (for micromobility)

NUMO has created a tool to leverage data to achieve policy outcomes, focused on collaborations for micromobility. On the portal, you can choose a policy goal to achieve, such as equity, safety or enforcement. You can then click through to find a matching objective, and the tool will inform the user what data are needed to achieve this goal.

- https://policydata.numo.global/

40. Travel time map

This comprehensive tool allows you to define a time range and check how far you can travel by walking, cycling, driving and public transport.

- https://app.traveltime.com/

41. MyData Global

MyData is an international non-profit organisation that empowers individuals by improving their right to self-determination regarding their personal data. The core idea is that everyone should easily be able to see where their personal data go, specify who can use their data, and alter these decisions over time.

- https://mydata.org/
- In Finland, student information is kept in one place, and can be shared for other purposes, such as automatic attribution of student discounts on public transport: https://www.hsl.fi/en/news/2019/student-tickets-and-auto-renewing-season-ticket-subscriptions-now-available-hsl-app-17105

42. SAE Mobility Data Collaborative

"The Mobility Data Collaborative is a multi-sector forum where mobility partners gather to establish a framework for mobility data sharing. Our partners share a joint vision of leveraging mobility data to promote safe, equitable, and livable streets for all."

- https://mdc.sae-itc.com/

43. TfL: Road danger reduction dashboard

Transport for London created a comprehensive dashboard containing (open) data on locations and details of road accidents and collisions in the London area.

- https://tfl.gov.uk/corporate/publications-and-reports/road-safety

EIB Technical Note on Data Sharing in Transport

European Investment Bank
The EU bank

European Investment Bank
98 -100, boulevard Konrad Adenauer
L-2950 Luxembourg
+352 4379-22000
www.eib.org – info@eib.org