



**Gemeente  
Amsterdam**

Version, version number (if  
any)  
8 February 2021

# CDS-M Principles Document

Towards standardizing data sharing between cities and shared mobility operators

# City Data Standard for Mobility (CDS-M) Principles Document

This document sets out the principles for designing a City Data Standard for Mobility. It is an initial version and thus provides a starting point for principles on which to develop a CDS-M.

## 1. Background

The mobility chain is changing at the speed of an express train: transport is becoming increasingly smart, and infrastructure, modes of transport and citizens are increasingly interconnected through technical applications. Data sharing is increasingly the basis for the operation of the mobility chain, and public authorities are increasingly adopting a digital approach. They are increasingly requesting data from market parties (commercial operators) in the mobility sector and other government/parastatal agencies in order to gain an understanding of, and control over, the use of the public space. This data is being used as input to policy, in order to monitor, manage and plan the public space and enforce the rules. It also creates opportunities to learn and innovate.

People travel beyond city limits, and mobility operators offer their services in various cities. Dutch cities currently set their own requirements for data sharing in collaboration with commercial operators. They would be able to learn from one another if data handling were based on the same principles. This can be done more effectively and efficiently by collaborating and standardizing, e.g. by setting standards.

## 2. Objective

The initiative to develop the standard has been taken by Amsterdam and other large cities – under the working title of a City Data Standard for Mobility (CDS-M): in collaboration with the private sector they would like to create a standard for data sharing between shared mobility operators and cities. Using a data standard would improve data sharing and thus provide better information on the use of shared mobility in cities.

The purpose of this document is to set out the principles for designing a CDS-M. It also describes the value of and need for a standard of this kind and how it should be developed and can be applied. Two further documents are being drafted, a 'CDS-M Blueprint' and a 'CDS-M API'. The CDS-M Blueprint describes the technical principles and the CDS-M API the actual code.

These documents serve as an initial starting point for the development of the CDS-M and its use based on use cases. The present Principles Document goes on to set out the following aspects that will contribute to the drafting of a data standard:

- Urgency and need
- Desired effect
- Our approach
- The ingredients of a data request
- The organization of data sharing
- Legal legitimacy

### 3. Urgency and need

There is a clear trend for public authorities to increasingly back up decisions on policy agendas with data. A good set of qualitative and quantitative data is needed for informed decisions to be made. Also, public enforcement responsibilities will increasingly be digitized, requiring up-to-date, correct information on the use of the public space. Enforcement increasingly goes hand in hand with control, as the authorities do not want undesirable situations to persist. As traffic becomes heavier in cities, the need to control it in various ways grows, and this will require information on how the public space is being used. To summarize, Dutch cities are requesting data from mobility operators to meet three needs:

- Policy-making and evaluation
  - What is available where in the public space (what is/is not working well)?
  - Vision and adjustment (short/long term)
- Enforcement
  - Undesirable behaviour in the public space
  - Fines
- Planning
  - Tactics (pricing, geofencing, behaviour)
  - Strategy (organization of the public space, permits and concessions)

Cities are not yet requesting data from mobility operators to meet these needs in a uniform way. Each municipality sets its own policy rules/permits for shared mobility (whether at the planning stage or in place), and requesting data is part of this. The authorities are not yet being specific about precisely what data they need, in what formats and at what frequency, and how they intend to handle it responsibly in relation to privacy (the GDPR) or economic sensitivity. As a result, mobility operators face different rules and data requests in each city.

This is resulting in inefficient processes in both mobility operators and public authorities, and these are getting in the way of scaling up shared mobility faster and achieving an MaaS ecosystem. How can we achieve shared information on the use of mobility in cities based on data? How are we currently learning from what in many cases is being permitted as an experiment? And how can we organize cities so as to encourage the use of shared mobility?

It is time-consuming for mobility operators to have to supply data based on different requirements each time. This is inefficient, and the data is consequently not comparable between cities, it is non-compliant and it creates delays and high costs for all concerned. Having to meet different data requirements and adapt their technical systems accordingly makes life difficult for mobility operators and increases their costs. It is undesirable for them to have to organize data sharing differently for each permit. Using a data standard would mean that the standard would be developed more quickly and it would enable better information to be obtained. Improvements can be achieved sooner with more organizations using the same standard. Also, many of the operators are small companies (start-ups or scale-ups), which do not have a host of data analysts on their payroll, with the result that supplying data without a standardized API costs a lot of unnecessary

effort, time and money. This has been shown in particular by the implementation of the TOMP-API and data submission.

The current approach is also inefficient for municipalities, which are each having to reinvent the wheel. Cities would like to know what requirements to set in invitations to tender and how to plan for this, enforce the rules and obtain information on traffic use in their city. In addition, as a result of lack of data standardization, cities are unable to learn more broadly how shared mobility is contributing to the mobility system, or how different sub-modes are responding to one another.

At the same time there are developments in response to this, a good example being the Centre for Transport and Infrastructure (CROW) Shared Bike Dashboard. At the moment, however, these are not really meeting the needs of municipalities, e.g. in the areas of policy-making and enforcement. The CDS-M working group is developing a data standard to provide additional data on shared mobility, and it may be worthwhile to incorporate and develop existing standards.

What is needed, then, is a standardized data request from municipalities to shared mobility operators, which will enable them to supply computerized data in the same way. This will make it easier to analyse the data on a large scale and learn how shared mobility works best. This will result in shared mobility being adopted sooner and scaled up more quickly, and the authorities can learn from shared mobility policy and pilots in other municipalities. The shared data needs to be able to be accessed where technically possible and properly processed, provided this is legally legitimate and meets the statutory requirements and standards under the GDPR. This may be done by municipalities or by a third party. This is one of the questions that will need to be examined as the standard is developed.

#### 4. Desired effect

A national standard will enable standardized data sharing between commercial operators and public authorities and will be used in partnerships between mobility operators and municipalities. The effects will be:

- Effective agreements on data sharing standards for commercial operators wishing to roll out nationally or internationally. A data standard will make it clear what has to be supplied to whom, and proper agreements will provide a toolkit showing how the data should be processed so as to be GDPR-proof.
- It will be easier for public authorities to exchange information from data: for instance, they can learn from pilots in other municipalities and adjust their policies accordingly. The actual use of data and the legal legitimacy of data sharing will be developed in specific use cases.
- Better planning and understanding of the use of shared mobility and the possible effects in the public space (e.g. accessibility, liveability and inclusivity)
- Faster scaling-up of the exchange of shared mobility data
- A standard will enable mobility operators to make a one-off investment in their technology and application-building for data sharing with municipalities.

Designing standards is not something new, and there are already various mobility standards. Some of the existing standards have their origin in the MDS standard: this was devised by the Open Mobility Foundation (OMF)<sup>1</sup>, which developed from an initiative in Los Angeles in the USA. The MDS standard was based on American legislation, which is not 100% in line with European privacy legislation. Various European cities (Brussels, Helsinki, Espoo, Stockholm, Faro and Basel) are now running pilots using it. The standard is applied in various ways: functionalities can be deactivated under the MDS. These cities feel the need to explore the possibility of a European standard as well.

The CDS-M being developed in this way is actually a development of and addition to existing standards, including the MDS. The aim is to bring it more into line with existing laws and regulations in the Netherlands and Europe, e.g. the GDPR.

The CDS-M also links up with current data sharing projects, e.g. MaaS, bike data, TOMP, Openbike2/GBFS+. Standards have already been identified in the MaaS learning environment and bike data project. The current project is looking at how all the data requirements can be standardized as far as possible, for example at what level aggregation should be standardized.

The approach is to involve commercial operators in the development of the new standard, so that it commands widespread support and can be applied in practice.

---

<sup>1</sup> <https://www.openmobilityfoundation.org/>

## 5. Approach

The G5 and the Dutch Ministry of Infrastructure and Water Management (in the MaaS programme) have been discussing the standardization of data sharing between cities and shared mobility operators since May 2020. One of the reasons for this is the setting-up of permit systems for shared mobility in the public space and new experiments in some large cities involving data requests. It was therefore decided to jointly explore how we can achieve a uniform standard in collaboration with the commercial operators.

This stage focuses on developing a number of products detailing the legal, technical and policy aspects as the basis for a standard.

Given the desire for greater standardization of data sharing, the G5 have set up a working group to develop this. The G5 working group has therefore drawn up a road map to examine what is needed more broadly in order to develop a data standard. This involves delivering five products:

- Principles document
- Survey of policy needs<sup>2</sup>
- CDS-M Blueprint
- CDS-M API
- Fact-finding survey of European cities' needs

The products are being developed in parallel, as they each deal with the same problem from a different angle, as outlined in the introduction. The data standard can be applied and improved in a use case, based on the development questions set out in the documents, in collaboration with cities and mobility operators and any companies accessing and processing the data.

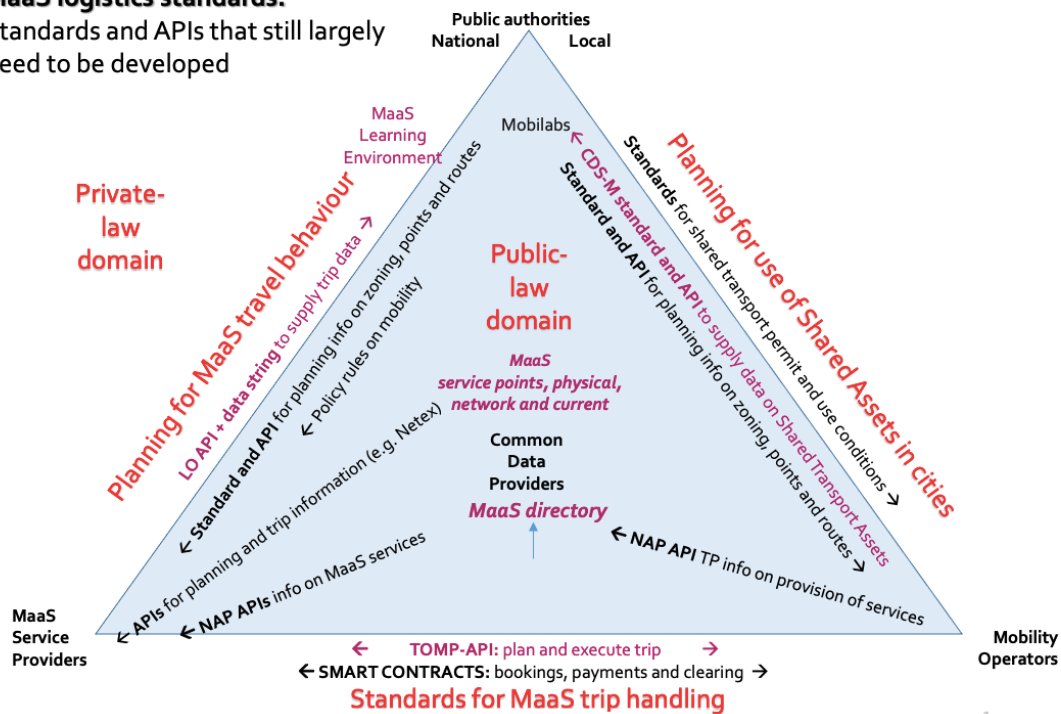
---

<sup>2</sup> Rotterdam, The Hague, Eindhoven, Amsterdam, Nijmegen, Arnhem and Utrecht

## 2. A data landscape vision (the MaaS ecosystem)

### MaaS logistics standards:

Standards and APIs that still largely need to be developed

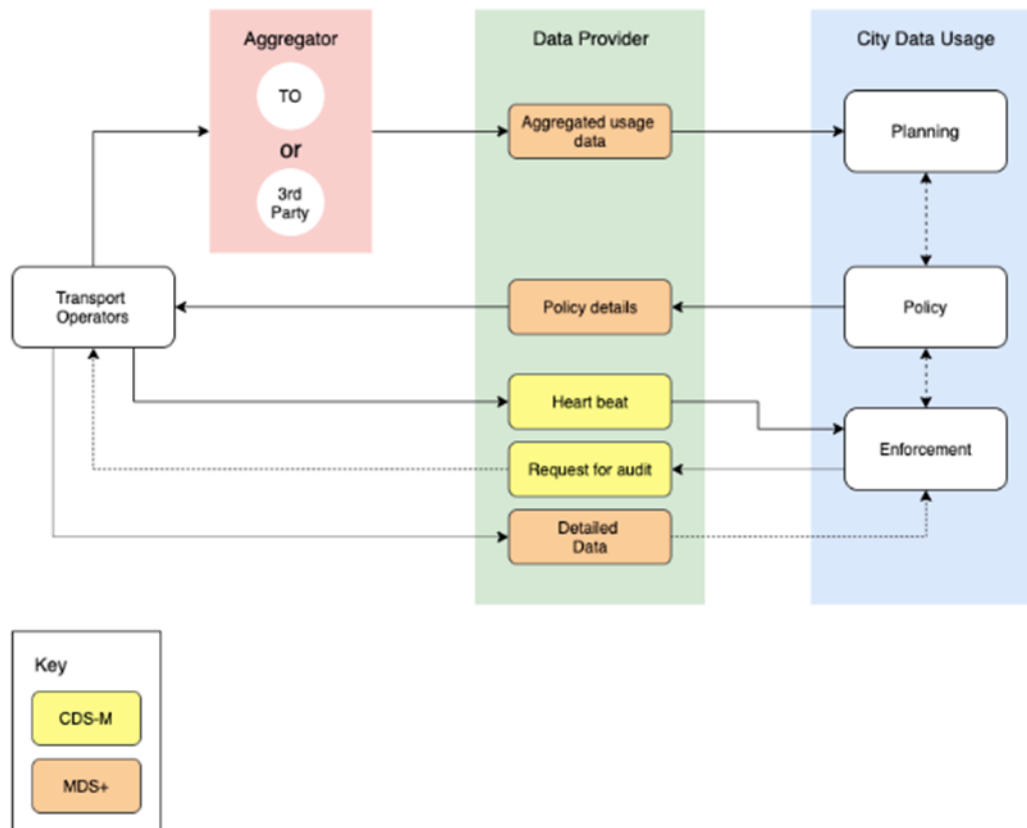


The Netherlands has a vision of the mobility data landscape (the MaaS ecosystem), which it is developing. The vision has been put forward by the Ministry of Infrastructure and Water Management in consultation with the organizations concerned and complies fully with the GDPR and other statutory requirements. Data sharing between shared mobility operators and public authorities is not a regular feature of the current data landscape. The CDS-M may eventually have a place in that data landscape. The chart below shows the Ministry's current vision of the mobility data landscape. It serves as one of the starting points in which the CDS-M has been given a proposed place. It shows that data can be collected and anonymized centrally, e.g. by a public authority or third party.

The CDS-M can thus become part of the data landscape. The diagram below shows what form a CDS-M data flow could take. The CDS-M Blueprint describes a data flow of this kind and develops it in relation to the data landscape and from a technical point of view.



## CDS-M Dataflow Diagram



### 3. The ingredients of a data request

Cities have differing requirements for data items to provide them with the desired information. The G5 cities have done a survey to find the common elements in those data requirements, which are set out in the 'Policy Needs Survey' document. This initial set of data requirements has been converted into the data items below and developed in the CDS-M API, dividing them into the following two categories of data, discussed in more detail below:

- Trip data
- Availability data

It is important to note that this is an initial selection, and new categories may well be added in future if they prove useful to both cities and shared mobility operators.

## **1. Trip data**

Trip data relate to users' travel movements, aggregated or otherwise. As these are individual users, their privacy has to be respected and the GDPR complied with at all times.

Trip data comprise:

1. Trip start location
2. Trip end location
3. Length of trip
4. Duration of trip
5. Number of unique users

This data can be requested at detailed level or aggregate level. The level of detail will depend on the ultimate purpose of the request. In principle, the CDS-M relates to aggregated requests, aggregated in terms of time intervals and areas. There are also use cases, especially in the area of enforcement, that require a higher level of detail: for example control of traffic in the city, digital enforcement and geofencing (these are referred to as 'policy APIs'). The CDS-M provides for a first step towards data sharing for the time being. If it works well and proves useful, other functionalities/APIs can be added to the standard.

## **2. Availability data**

Availability data relate to the availability and status of a transport operator's vehicles.

Availability data comprise:

1. Stationary vehicle location
2. General vehicle status (charging, not in use)
3. Availability and vehicle use status

The availability data is available in the TOMP-API and will therefore not be included in the CDS-M. It can be 'polled' from the TOMP-API, but this will require setting up a system that can become part of the data landscape. Another way of gaining information from availability data is to 'scrape' that data from the operators' apps, which can be done using an algorithm. A disadvantage of this approach is that it only provides data on vehicles in use, not those that are being charged or moved or have broken down.

## **3. The organization of data sharing**

For the data items to be exchanged, good understandings need to be reached between public authorities and between mobility operators, public authorities and data processors. To achieve good understandings, the following points will serve as an initial starting point for development and examination in specific use cases:

- The organization of the data infrastructure
  - Should a single national data infrastructure be used, as in the case of MaaS, where data can be received, or should cities receive data individually?
  - How does the data infrastructure contribute to the legal legitimacy of data sharing under the GDPR?
  - What are the technical principles used in the data infrastructure?
  - Who is responsible for creating it?
  - Who will maintain and manage it?
  - How is security organized in it?
- The data sharing process
  - Should a uniform blueprint (CDS-M API) be created for agreeing data sharing processes?
- Dashboard monitoring system
  - Should there be a national dashboard monitoring system (based on CROW's current Shared Bike system), or should cities create their own dashboards?
  - What data should the dashboard show?
  - Should the dashboard be open source and accessible to everyone?

#### 4. Legal legitimacy

A legitimate legal basis is needed for data from mobility operators to be accessed. Under the GDPR there are various roles and positions that enable data to be requested and received, thus guaranteeing legal legitimacy. These roles affect the level at which data can be requested and processed. Cities will need to enter into a dialogue to decide which role is most suitable.

It is important to determine the role correctly, based on the use and processing of the data, as it may be traceable personal data that needs to be handled with care in line with the GDPR.

What is personal data? Personal data is any information relating to an identifiable natural person, who can be identified directly or indirectly from identification data such as a name, an ID number, location data, an online identifier or other factors revealing that natural person's physical, psychological, genetic, mental, economic, cultural or social identity.

Once data on vehicle movements is shared it must not be traceable to the 'natural person' involved in the movement. It is important to ensure that anonymity is guaranteed at all times when sharing data.

For mobility data to be processed correctly, it will have to be decided what roles the parties concerned are to take on in relation to data processing. The following three data processing roles are identified:

- **Data processing** is defined as any operation or set of operations which is performed on personal data, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission,

dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

- The **processor** is a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.
- The **controller** is the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data (Article 4 of the GDPR). The controller is responsible for the lawful, GDPR-compliant processing of personal data. Taking into account the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for the rights and freedoms of natural persons, the controller shall implement appropriate technical and organizational measures to ensure and to be able to demonstrate that processing is performed in accordance with this Regulation. Those measures shall be reviewed and updated where necessary. Adherence to approved codes of conduct as referred to in Article 40 or approved certification mechanisms as referred to in Article 42 may be used as an element by which to demonstrate compliance with the obligations of the controller. Another way of doing this is to ensure data protection through design, by measures such as pseudonymization designed to implement data protection principles such as data minimization and storage limitation effectively and incorporate the necessary safeguards in processing to meet the requirements of the Regulation and protect the rights of those concerned (Article 25 of the GDPR).

## 1. The next steps

Developing a workable data standard requires proper policy and legal substantiation of what the data is needed for and how it is to be used. It is the role of the public authorities to indicate how they intend to use the data. A good application can be created, based on specific use cases, to enable the principles and problems to be further examined and developed.