# Digital Twin for Mobility

Concept and baseline study
Working paper (9 September 2022)

Authors: Janne Rinne, Juho-Pekka Virtanen, Sami Sahala, Juho Kostiainen, Aaron Koskela

MOBILITY LAB HELSINKI

FORUM VIRIUM HELSINKI

Supported by Helsinki

**Table of Contents**

# Introduction

What is a digital twin for mobility? This document explains what is meant by the concept. It provides a brief introduction to the concept, explains the data it consists of, and outlines its potential applications. It provides a snapshot of where the topical development is going and what the opportunities and barriers for further development are, especially in the context of Helsinki.

There is no commonly agreed definition for the “digital twin for mobility.” This document acknowledges the complex nature of not only the topic itself but also the terminology around it. As the topic is constantly evolving, this document should be seen as a starting point for discussing the further development of the digital twin in the domain of traffic and mobility. This working paper is not a final conclusion on the topic, but rather a starting point for the inventory of data sources, use cases and development needs. It is meant to serve as inspiration and a starting point for discussions about common interests and ideas.

This document is part of the Mobility Lab Helsinki project, coordinated by the City of Helsinki and Forum Virium Helsinki, the City’s innovation company.

Updated versions of the working paper will be provided as the project proceeds. This first version of the document was published on 9 September 2022.

# Digital Twins

The concept of digital twin originates from the manufacturing industry, where digital twins have been used in designing components, machines and systems. By modeling the behaviour of the system, the idea is to design “digitally first” as far as possible so that the functions and processes can be simulated and adjusted before the actual physical components and systems are built, and this way optimise functions and save resources (Figure 1).

In the context of urban digital twins, the same “digitally first” principle applies. However, as cities are not mechanical systems, this does not only imply e.g. simulation use, but also includes participatory approaches and visualisation for supporting decision making. By combining several data sources and visualising data from different disciplines involved in urban development, better decisions can be made, problems identified and synergies reached.

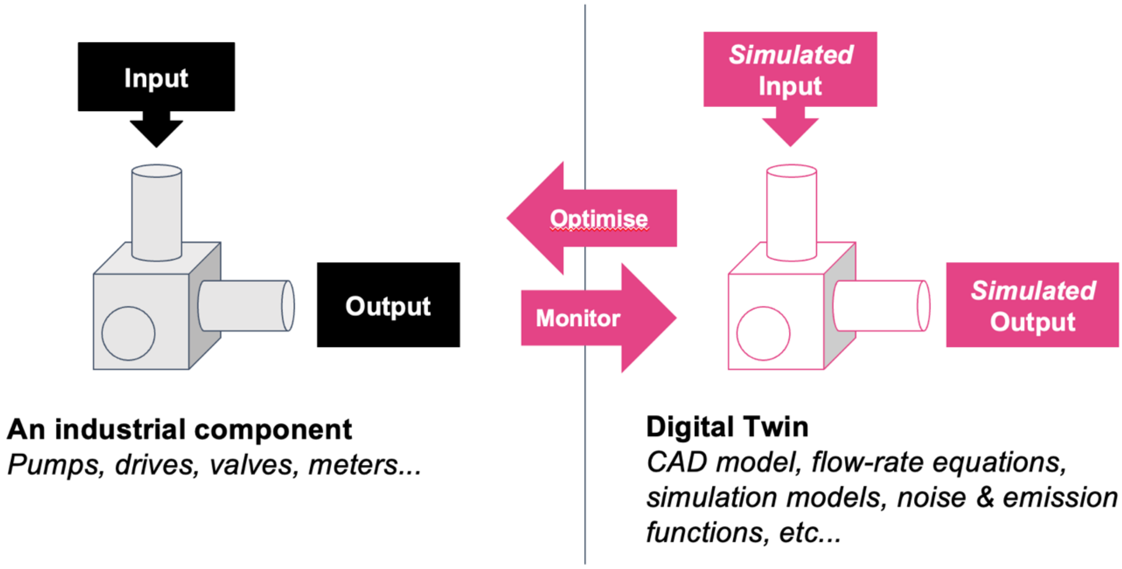


*Figure 1. Digital twin in industry.*

In architecture, engineering and construction (AEC), a similar aim is encountered when utilising building information models (BIMs) in design work. By combining the inputs from different design disciplines into a single model, “clashes” (non-compatible elements from different designers) can be detected beforehand. In addition, there is rising interest in utilising BIM models for facilities management and digital building permit processes.

The main difference between cities and BIMs or the manufacturing industry is that no complete design documentation is available for cities due to their complexity and long history. Thus, the urban digital twin has to be formed at least partially relying on mapping. The social fabric, human processes and people’s behaviour are critical issues to be addressed in an urban digital twin, in addition to the physical description of the urban environment.

City modeling is the prevalent starting point for forming an urban digital twin. In Helsinki, the 3D modeling activities have a long history, with the first experiments having been carried out in the 1980s. Currently, the City maintains two 3D city models: a city information model (in CityGML format) and a textured mesh model, based on photogrammetric reconstruction. These models and Helsinki digital twin aims are described online in the City's website. In addition to the regularly maintained models, various experimental models have been produced to illustrate the various possibilities for use, such as the city Minecraft model (Figure 2).


*Figure 2. Minecraft-Helsinki3D+ examples.* https://hri.fi/data/fi/showcase/minecraft-helsinki3d

Data integration is at the heart of a digital twin. In an urban digital twin, various dynamic data sources, such as IoT sensor data feeds, are combined to the digital model of the environment. As the city is a sociotechnical system comprising multiple organisations, processes and stakeholders, introducing the social dimension to the urban digital twin is a significant development direction. This is also required for the digital twin to become widely applicable.

While the digital twin of a city combines different data sources depicting the urban environment, it may not contain everything from a single thematic area. After all, in many sector-specific design and decision-making tasks, the datasets used can become very task-specific and perhaps even require dedicated tools. The digital twin does not become "one system to rule them all," but rather a complementary tool for cross-sectoral work. This implies that there may be partially overlapping, thematic digital twins in the city for meeting the needs of specific purposes.

For example, a city digital twin might contain the general description of the city road network and traffic flows, but not necessarily the detailed traffic counter statistics. The more detailed datasets and systems may, in the mobility context, form their own digital twin, oriented towards the mobility theme (Figure 3).

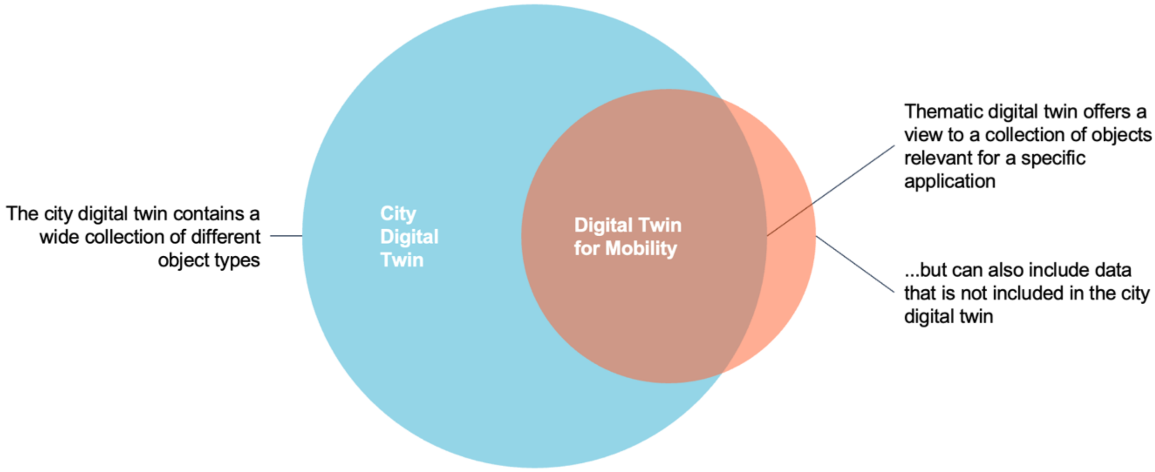


*Figure 3. Relation of a thematic digital twin and city's digital twin.*

# Digital Twin for Mobility

Recently, the digital twin approach has been introduced in the field of mobility and traffic. Digital twin is a means to combine information from different data sources, enabling better description, analysis and simulation of the urban environment and people's mobility as part of it. It enables development of better services, both public and private, for citizens.

The digital twin for mobility describes the traffic infrastructure and environment, the traffic itself, and related conditions and context. It comprises numerous data sources. It is not one system, but instead a constantly evolving combination of diverse data sources and information that can be applied for diverse use cases (Figures 4 and 5).

Many parties contribute in producing and updating the data. It can be studied, refined and combined using APIs and digital tools, creating added value to its users. With up-to-date data, changes in the real world are realised in the digital models.

One promising aspect of digital twin development is the transportation module of the CityGML 3.0 standard (https://docs.ogc.org/is/20-010/20-010.html#toc39). The module enables detailed description of the street environment, including street segments, lanes, pavements and intersections, for example. The new version of the standard was finalised recently, and the practical implementations of it are still scarce. Mobility Lab Helsinki will explore the potential of the CityGML 3.0 Transportation Model during 2022-24.

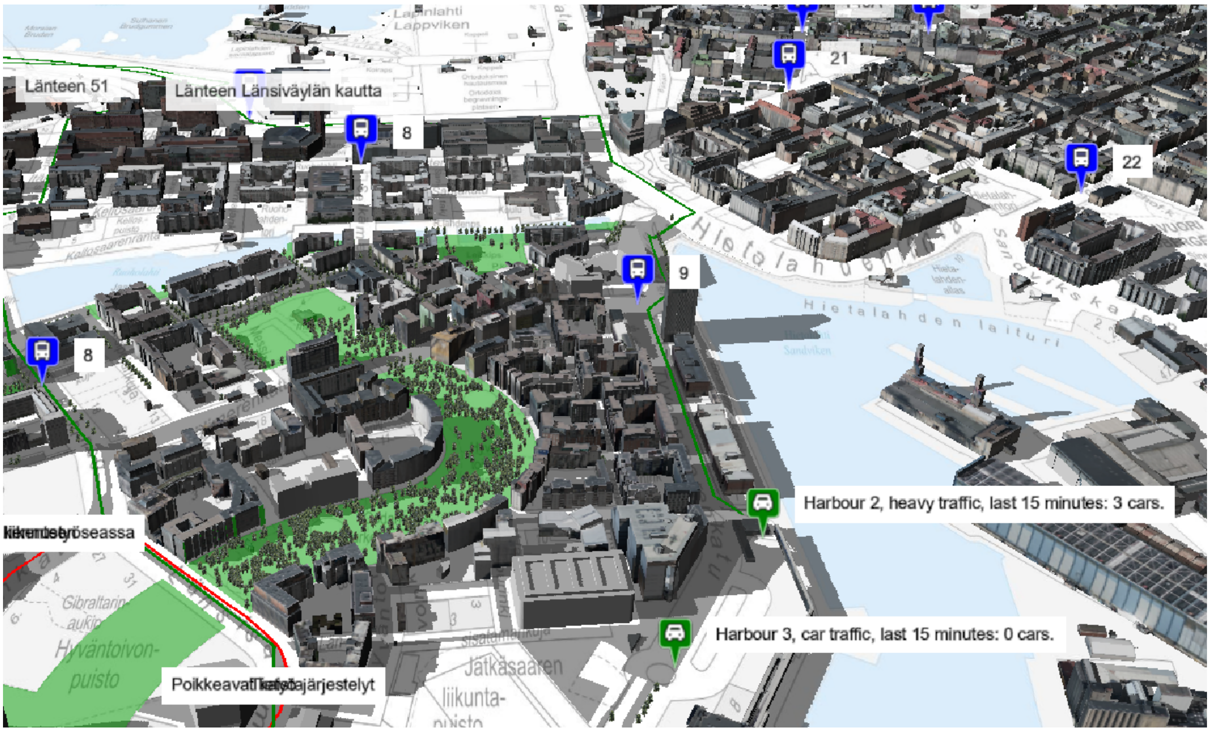


*Figure 4. Example of an interactive visualisation combining traffic counter data, real-time vehicle positions, suggested routes and the 3D description of the urban environment.*

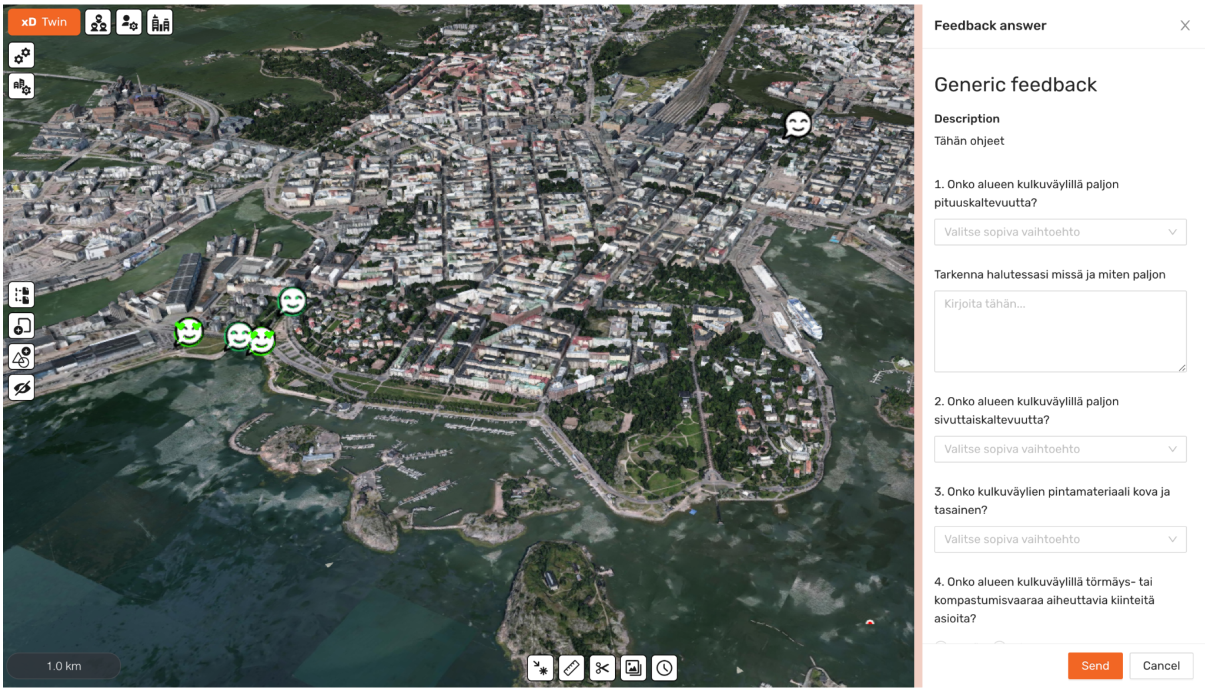


*Figure 5. Citizen feedback gathered via digital twin (xD Twin) can be transferred to the City's participatory democracy platform (Decidim). Source: DVECE project.*

***Why do we call it a "digital twin" - isn't it just a bunch of data and data models?***

Digital twin for mobility integrates mobility-related datasets to serve diverse use cases. It links the traffic to its infrastructure and context, which in many cases are presented in three dimensions (i.e. physical infrastructure). Although all use cases of the digital twin do not require 3D presentation, it is often present in the digital twin for its illustrative power. When discussing "non-3D" developments (i.e. new sensors for monitoring traffic flows) the digital twin is still a useful concept in emphasising the data integration aspect. The possibility of applying different data sources together, leading to innovative combinations of data, is highlighted in the digital twin.

# Data in digital twin for mobility

The data in digital twin for mobility can be divided into three categories (Figure 6):

1. Infrastructure and traffic environment (e.g. detailed street structure, traffic signs, accessibility)
2. Traffic (e.g. number of vehicles, cyclists, pedestrians)
3. Conditions and context (e.g. air quality, roadworks, maintenance needs, resident feedback)

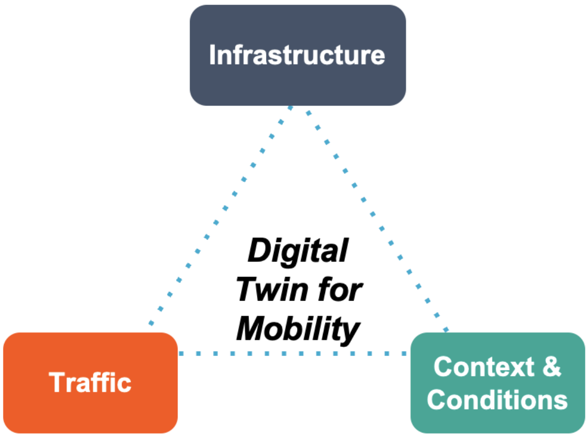


*Figure 6. Data types of digital twin for mobility.*

**Infrastructure and traffic environment data**

Conventionally, most infrastructure-related data about traffic environment is either point-type, line-type or administrative information about certain parts of infrastructure. Additionally, more and more data is now managed as area-type. While this information is typically kept and maintained as separate datasets in their dedicated systems, it can later be integrated with a city information model or other GIS datasets.

Point information refers to a single point in the environment that has a certain function or other relevant information. It can be an area or structure with little or no information about its dimensions, but it is referred to as a single point. Examples are traffic signs, bus stops and crosswalks.

Line format is used by breaking e.g. a road/street into smaller segments, with specific information about that segment. This information can be the width of the road, restrictions or other regulations valid in that segment (e.g. speed limits) or information about the structure, e.g. what kind of pavement it has.

Area information can be used to describe information that would be impossible or complex to describe as point or line information. Examples are different zones (e.g. for parking, low emission) or even the area a street or a public square occupies. Area information is traditionally used in geographic data, so the standards and tools exist.

Additional related information can be a status of the above objects. This could mean the condition of a pavement or a traffic sign, or maintenance work carried out on it, or the current availability of the area for a specific function.

Increasingly, the data describing the traffic environment can be integrated into the city information model, utilising the CityGML standard and its “Transportation” module (https://www.ogc.org/standards/citygml). Thus, the description of traffic infrastructure becomes a part of a larger depiction of the urban environment. This supports easier understanding of context and the interactions between traffic and other elements of the urban environment (Figure 7).

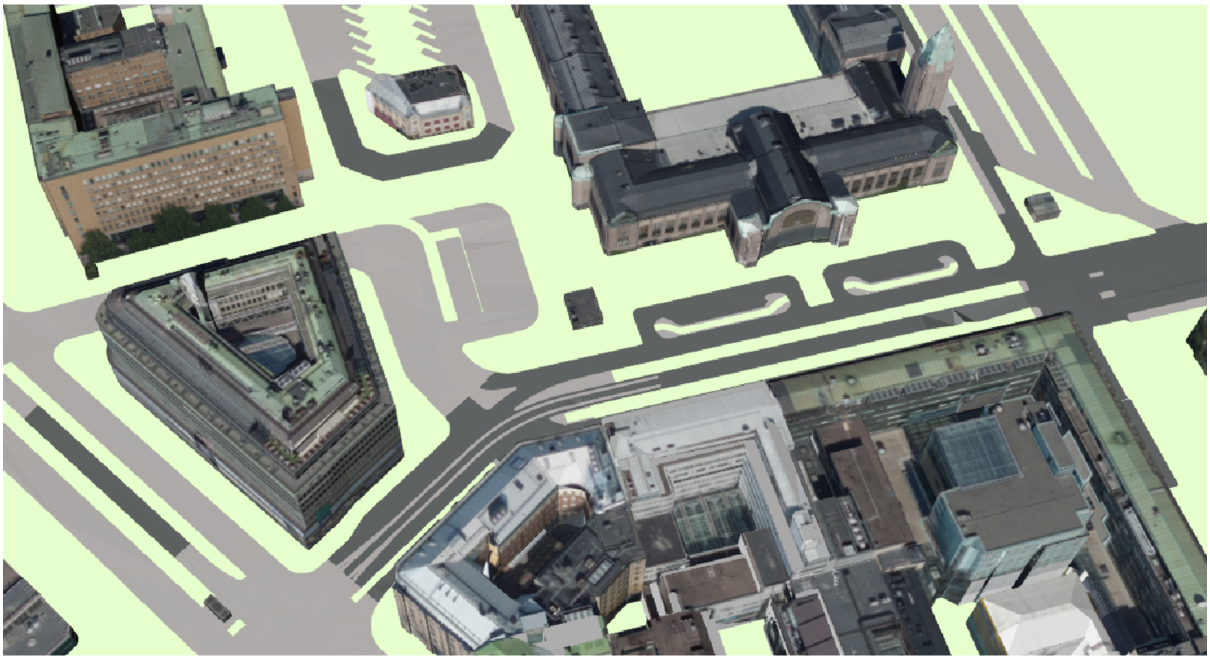

*Figure 7. Example of traffic infra shown together with city information model buildings.*

**Traffic data**

Traffic data (or mobility data) refers to monitored information about the movements of people and vehicles. These can be observations (i.e. “monitoring from street side”) or the vehicles themselves sending information about their movements and trajectories, which is increasingly available through internet connectivity and GNSS positioning.

Most common types of traffic data are traffic volumes (e.g. how many cars have passed a certain point in a given timeframe, what kind of vehicles), traffic speeds as either point information or travel-time between two points, directions and routes. Specifically, Origin/Destination matrices are one common, often sought-after type of data about movements.

Traffic data of public transport falls under the above categories, but has its own standards and use cases. Additionally public transport movement data is tightly linked to scheduled movements (e.g. timetables, routes).

Occupancy of vehicles is another measurement, which can also be linked to e.g. usage of separate lanes. In public transport, passenger counting is a common measurement.

New transport services, e.g. car-sharing, shared city bikes and e-scooters, can provide the above data.

**Conditions and context data**

Besides the data about infrastructure and traffic itself described above, there are different types of data that are about, or relevant to, mobility.

Environmental information about the weather, air quality, noise and snow are highly relevant to mobility. Close to these are measurements such as friction, i.e. slipperiness of pavement.

Information about traffic accidents or other special situations is a form of traffic data, but also linked to the infrastructure.

Maintenance work can be considered as status of infrastructure, but as one of the main functions of road authority it is often considered its own category. The information in the current datasets depicting the city infrastructure is not sufficient to provide enough details about the different aspects, progress and results of the maintenance work.

Other information linked to infrastructure data can be relevant. One example of this is last mile logistics. The information about infrastructure, including addresses, is the basis for a successful delivery, but it is not enough to provide instructions of how to make that delivery (route or door to use). Logistics is an area where Helsinki has experimented with new types of data, i.e. extra information about the actual delivery point (door) of a given address, accessibility and walking routes there.

## Data ecosystem

Traditionally, the data concerning urban traffic has mostly been acquired, used and distributed by the city administration, mostly for the purposes of traffic engineering and urban planning.

Today, new data sources describing urban mobility are challenging the traditional city administration centric approach to traffic data. Private vehicle fleets, light electric vehicles and MaaS services all accumulate mobility data but aren't controlled by the City (unlike traffic counters or public transportation). Robotic systems equipped with cameras and LiDar sensors have the potential to map the urban environment on a high level of detail. Mobile network operators and other technology companies gather and provide detailed data about people's mobility and the urban environment (Figure 8).

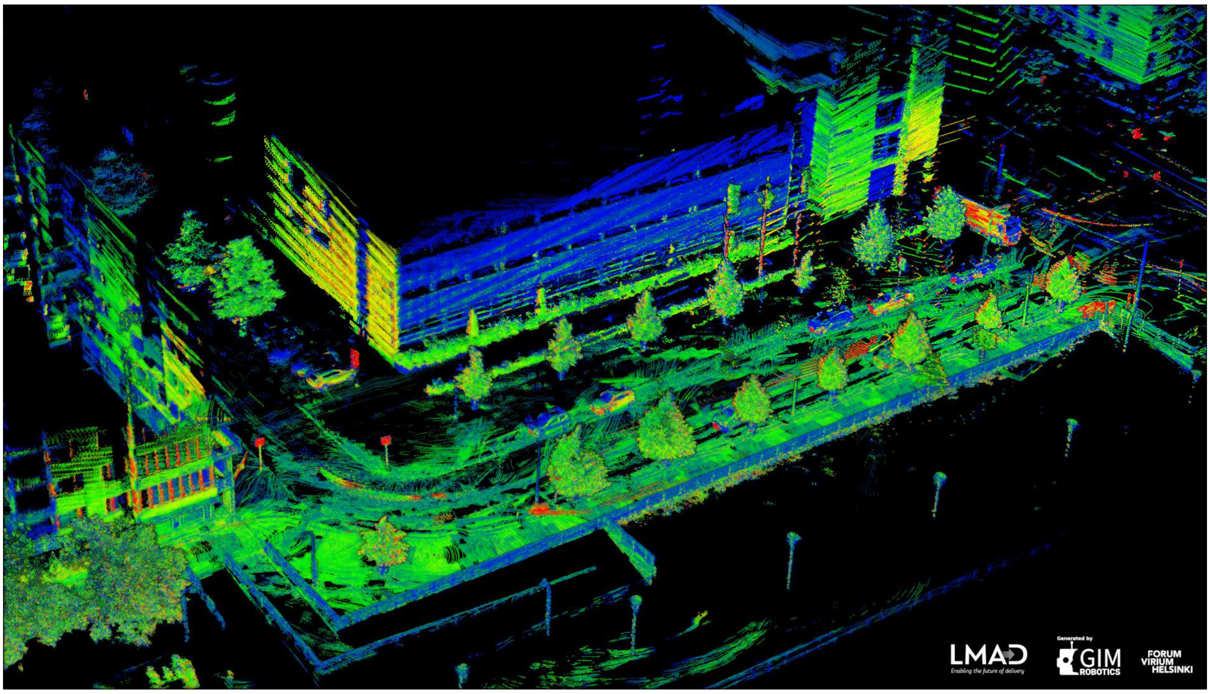


*Figure 8. Point cloud produced by LMAD delivery robot.*

New technologies and sensors open up opportunities for crowd-sourced data production, citizen participation and feedback. In the urban context, it becomes important to address the social component of a digital twin and tap its potential in supporting better planning and service development. In the future, the MyData approach may offer new ways for citizens to share and manage their own personal data, also regarding mobility.

The City and its contracted consultants are not the only users of urban mobility data anymore. There is increasing demand for cities' data in private companies for the development of new commercial services, solutions and innovations. For example, robotic systems require dense point clouds (HD maps) to successfully operate in the urban environment. Mobility service providers can benefit from data concerning events, services and existing traffic flows of the city. We do not even know all the potential data users, nor the potential ways of producing, combining and utilising the data in the future.

How should these new demands of providing and using data be addressed? It is not feasible for the city organisation to procure all relevant, or *potentially* relevant, data. Furthermore, if the data providers are engaged in selling the data, they most likely are not that excited about the open data policies enforced by the cities - if they stipulate the open release of their products.

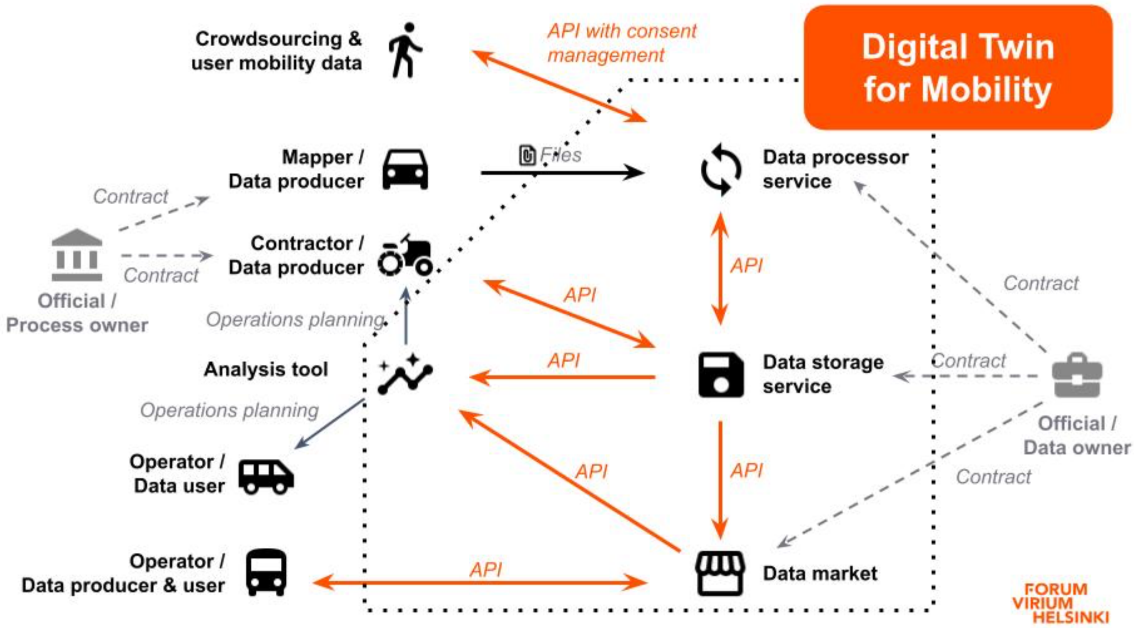


*Figure 9. Data ecosystem of Digital Twin for Mobility. Google Material Icons used under the Apache 2 licence 2.0.*

In general, there tends to be a shared understanding that an ecosystem-based approach towards the development of digital twins and mobility data is desirable (Figure 9). An ecosystem with many data producers and users, with APIs and integration tools enabling better use of datasets, benefits from the data sharing as a whole. However, sharing data in an ecosystemic way is not merely a technical challenge, but involves various issues that need to be solved in terms of using the data, property rights, pricing, and business models. New procedures for data sharing are necessary.

Many open questions remain. How should the urban mobility data exchange be organised, and who should facilitate it? What are the mechanisms that dictate data pricing, and what kind of contractual models are needed? There are several open questions in the first steps of progressing from city-centric data acquisition and distribution to a data ecosystem combining public and private actors, using both open and private data.

# Data providers and repositories

**HRI**

Helsinki Region Infoshare, HRI, is the main starting point for all open data about Helsinki. HRI is a “phonebook” of most openly available data in the Helsinki region - not just Helsinki but also the cities of Espoo, Vantaa and Kauniainen. It may be easiest and quickest to find links to required data through HRI (https://hri.fi/en_gb/).

**City of Helsinki**

The City of Helsinki is responsible for the official information about its infrastructure and environment. The main stakeholder, the Urban Environment Division, is responsible for planning, construction and maintenance, building supervision and environmental services in the Helsinki city environment.

The city’s geospatial data is maintained in a number of registries (such as the register of public areas, “YLRE”). From these registries, several geospatial data products are provided. Data can be viewed in the online map service (https://kartta.hel.fi/) that also offers download functions to some datasets. WMS and WFS interfaces provide API access to both open and closed geospatial data (with the closed map layers requiring authentication). Many of the registries are split amongst a number of map layers in the APIs and the map service. The open data portal Helsinki Region Infoshare (HRI) provides metadata for the WMS and WFS layers and offers download possibility to some of the datasets. Most of this data is available via WFS interface to be used e.g. in QGIS: Helsingin kaupungin avoimet paikkatietoaineistot (https://kartta.hel.fi/ws/geoserver/avoindata/wfs). There is also a guidebook for novices on Helsinki geographical data, although it is only available in Finnish: https://kartta.hel.fi/avoindata/dokumentit/2019-05-hki-aloittelijan-paikkatieto-opas.pdf.

Related to traffic data, the City of Helsinki has its own programme developing ITS in Helsinki: https://www.hel.fi/static/liitteet/kaupunkiymparisto/julkaisut/julkaisut/julkaisu-16-19-en.pdf. One of the main purposes of this programme is to develop “situational awareness,” i.e. real-time information about the traffic system. This will result in much more mobility data to be opened in the coming years.

**STARA**

Helsinki City Construction Services, STARA, provides services related to construction, maintenance and logistics in the urban environment for the City of Helsinki. The City of Helsinki has roughly 1,700 vehicles registered, of which 1,240 belong to STARA. Social and health care, the Urban Environment Division and other industries lease 450 of these vehicles.

STARA’s main focus is on the construction, care and maintenance of streets and parks, the City’s facilities – schools, day care centres and other premises – and Helsinki’s nature, city forests and coastal waters.

Through this work STARA collects and possesses a great amount of information about the city. This information includes smart asset and lifecycle management data from the infrastructure, as well as observations of the assets and environment conditions that are collected, monitored and reported by telemetry systems, mobile applications, drones and machine vision.

STARA operates multiple production control systems for construction, vehicles, machinery and equipment management to collect a wide range of information from the environment, construction sites, driving routes and lead times of work phases as part of the production's daily routines. Currently, vehicles are not tracked to produce floating-car data or other meaningful open traffic data for other than internal purposes.

The City of Helsinki and STARA are providing street cleaning plans and schedules (http://www.puhdistussuunnitelmat.fi/helsinki/en/) to add transparency and streamline the process. Communication with the citizens, feedback and condition information play a significant role in anticipating changes in street space and traffic.

All the data itself is being used as a part of internal production and work plans development. In light of Helsinki's digital twin development, opening some of STARA's data is being investigated, and the first results are expected in 2022 - 2023.

**HSL**

Helsinki Region Transport HSL is a joint local authority whose member municipalities include Helsinki, Espoo, Vantaa, Kauniainen, Kerava, Sipoo, Tuusula, Kirkkonummi and Siuntio. HSL plans and organises public transport in the region and procures all the public transport services. HSL is responsible for ticket sales, ticket inspections and passenger information.

Besides public transport, HSL is also responsible for the Helsinki Region Transport System Plan, which is the regional traffic system level plan.

HSL has been pioneering mobility open data (https://www.hsl.fi/en/hsl/open-data), with the Journey Planner API released already in 2008. HSL also has an active developer community reachable via https://twitter.com/hsldevcom and https://fi-fi.facebook.com/HSLdevcom/.

**Fintraffic - Digitraffic**

As a special assignment under Finland's Ministry of Transport and Communications, Fintraffic provides and develops traffic control and management services in all traffic modes; road, rail, maritime and air. They collect, administer and open data and create opportunities for new business in the market.

Fintraffic's goal is to create a permanent information platform for use by traffic ecosystem operators. Digitraffic (https://www.digitraffic.fi/) is the data marketplace that aims to collect, combine, share and invoice traffic data and thus promote ecosystem operations (https://www.fintraffic.fi/en/fintraffic/platform-and-information-services).

Digitransit (https://digitransit.fi/en/) is an open source journey planning solution that combines several open source components into a modern, highly available route planning service. Route planning algorithms and APIs are provided by Open Trip Planner (OTP). The Helsinki Journey Planner Reittiopas by HSL is also based on Digitransit. In fact, it's the original version of Digitransit.

Fintraffic also operates the Finnish National Access Point NAP (https://finap.fi/#/).

**Väylä – Digiroad**

The Finnish Transport Infrastructure Agency (FTIA), in Finnish "Väylä," is responsible for developing and maintaining the state-owned road network, the railways and the waterways.

Väylä maintains several datasets which belong to the field of application of the European Union's (EU) INSPIRE directive. They have open WMS, WFS and REST services, and a downloading and viewing service for their data (https://vayla.fi/en/transport-network/data/open-data/api).

Digiroad, hosted by Väylä, is a national database that contains the geometry of the Finnish road and street network featured with the most important road attribute data. It is in use for anyone free of charge as open data (https://vayla.fi/en/transport-network/data/digiroad).

The Digiroad Database provides a comprehensive and up-to-date description of the Finnish road and street network online. The data enables and supports the development and commercialisation of services and applications for e.g. route planning, navigation, tourism, and intelligent transportation systems (ITS) purposes. The latest documentation of the data is available here: https://vayla.fi/en/transport-network/data/digiroad/documents.

Digiroad has been available for users since 2004. The Finnish Transport Infrastructure Agency (FTIA) Väylä administers the Digiroad service. Municipalities and the Regional Centres for Economic Development, Transport and Environment (ELY) take care of the data administration in co-operation with the Finnish Transport Agency. Hence, much of the infrastructure data about Helsinki in Digiroad is also available from Helsinki's own sources, but may be easier to use via Digiroad.

**HSY**

Helsinki Region Environmental Services HSY has plenty of open data, with the most relevant to mobility being air quality and emissions data, for which HSY is the responsible authority for (https://www.hsy.fi/en/environmental-information/open-data/).

**FMI**

The Finnish Meteorological Institute FMI provides loads of open data and some open source software available at https://en.ilmatieteenlaitos.fi/open-data. FMI also has a long history in developing "road weather," a pinpoint weather forecast for road authorities that they also sell and consult worldwide.

**NLS**

The National Land Survey NLS (Maanmittauslaitos) hosts Paikkatietoikkuna (https://kartta.paikkatietoikkuna.fi/?lang=en) which is the national geoportal presenting spatial data and related services and their benefits. The website contains over 2,000 map layers from more than 60 organisations. Paikkatietoikkuna is intended for both spatial data professionals and everyone who is interested in maps and spatial data, and it’s an important part of Finland's implementation of the INSPIRE directive.
Much of the information is also available via other sources mentioned in this document, but you may find some useful data not easily reachable elsewhere. It does have an easy user interface, and by using the service you can look at several superimposed map layers, embed a map on your own website, make statistical thematic maps and perform simple spatial data analyses.

# Data

Digital twin for mobility includes diverse data about traffic infrastructure, traffic, and conditions and context (Table 1). These aspects include both real-time and historical data. As such, the digital twin for mobility is an umbrella term covering all relevant data for researching or creating, implementing, piloting and operating mobility services in Helsinki.

Mobility Lab Helsinki and other related projects have identified some key data sources, which are described below.

*Table 1. Data types relevant to digital twin for mobility.*

| Traffic infrastructure | Traffic | Conditions and context |
|---|---|---|
| • Base information about traffic infrastructure<br>• Traffic guidance<br>• Temporary infrastructure changes<br>• Public transport infrastructure<br>• Other sources of traffic infrastructure data | • Traffic volumes<br>• Travel times<br>• Routes and Origin/Destination matrices<br>• Public Transport<br>• Digitraffic<br>• Other traffic data | • Accidents and incidents<br>• Maintenance and roadworks<br>• Weather and road conditions<br>• Air quality and noise<br>• Public transport interphases<br>• Other condition and context data |

## Traffic infrastructure and environment

### Base information about traffic infrastructure

Out of the open data by the City of Helsinki, the main source for relevant traffic infrastructure data is the Register of Public Areas in the City of Helsinki (Hgin kaupungin yleisten alueiden rekisteri, YLRE). It contains data about the **city's "street and green areas," namely street network as polygons**, i.e. the area the street, road or a path occupies, with additional administrative information, such as classification and maintenance responsibilities. YLRE is available in WFS format (https://kartta.hel.fi/ws/geoserver/avoindata/wfs). Much of the data in YLRE is also available at https://kartta.hel.fi/.

NOTE: Official centre line geometry, which is often used as base data for linked street networks, has not been updated in Helsinki since 2014, but there are recent plans to do that again. The old material is available at https://kartta.hel.fi/ in .dgn-format (Streets and Parks → Traffic Routes).

Potentially noteworthy is the dataset for The **Prioritised Winter Maintenance Network for Pedestrians and Cyclists** in the City of Helsinki ('Helsingin jalankulun ja pyöräilyn talvihoidon priorisoitu reitistö'). The information is from the 2016-2017 winter season and is being updated until 2019. Available at https://kartta.hel.fi/, under Walking and CyclingEnhanced Winter Maintenance Network.

Finnish government agencies, especially Fintraffic and Väylä, have **various datasets about Helsinki's traffic infrastructure**. The main source is Digiroad (https://vayla.fi/en/transport-network/data/digiroad), the National Road and Street Database that contains the geometry of the Finnish road and street network featured with the most important road attribute data. It is in use free of charge as open data under Creative Commons by 4.0 licence. The latest documentation of the data is available at https://vayla.fi/en/transport-network/data/digiroad/documents.

Note that much of the data about Helsinki in Digiroad is submitted by the City of Helsinki, and available in Helsinki's own sources as well. Still, accessing this data via Digiroad may in some occasions be more straightforward and easier.

As a part of repeated urban surveying, the City of Helsinki acquires and provides **airborne Lidar point clouds** covering the entire city administrative area (excluding some of the remote isles). The classified point clouds are available as 500 x 500 m tiles in LAZ format from the https://kartta.hel.fi/# service. Currently, the years 2015, 2017 and 2021 are provided. Additional airborne Lidar data is available from the National Land Survey of Finland (https://www.maanmittauslaitos.fi/laserkeilausaineistot).

The current **3D city models for the city of Helsinki** are available at https://hri.fi/data/en_GB/dataset/helsingin-3d-kaupunkimalli. Their production and contents are described at https://www.hel.fi/helsinki/en/administration/information/general/3d.

## Traffic guidance

There are specific datasets about **intersections with traffic lights, traffic signs and parking places** (including residential parking zones, payment zones and special areas for tourist busses), all available at https://kartta.hel.fi/. Please note that this data is unofficial and contains errors, hence to be used as informative. All official parameters of parking places are to be checked from the installed traffic signs.

The City of Helsinki has started a project to develop a new system that would act as the single place to find all data about the above elements. The timeline of the project is not fixed, but during the development there will likely be some temporary experimental datasets released.

## Temporary infrastructure changes

Helsinki is - especially in the summer season - riddled with numerous **road works and other construction projects** causing temporary changes in the traffic environment. Information about these projects is not available in a comprehensive manner in any system. A special task force set by Helsinki's mayor in 2018-2021 came to the conclusion that there may not even be a legal basis for the City of Helsinki to get all information about road works.

There are several datasets that may be applicable depending on the use case, available at https://kartta.hel.fi/:

- Under City and Traffic Planning, e.g. '**Traffic Plans of the City of Helsink**i' ('Helsingin liikennesuunnitelmat')
- Under Streets and Parks → Public Area Register are several datasets about **events and temporary traffic arrangements**, e.g. Event and Land Use Permits in Public Areas of the City of Helsinki ('Helsingin kaupungin yleisten alueiden tapahtuma- ja maankäyttöluvat') can also be useful, as it contains the granted permissions for use of public space, including streets and public parking lots.

Users are advised to note that the above and many more datasets mostly contain information about plans. As plans change and e.g. construction is rescheduled, that information often does not exist in the proper manner. Additionally, there are numerous construction or other activities in Helsinki by many stakeholders, especially in the summer season, which are not available in any database.

## Public transport infrastructure

Helsinki Region Transport (HSL) is the regional public transport authority. HSL offers open data in the form of API services and data packages (https://www.hsl.fi/en/hsl/open-data). Geographical information is available via their GIS portal (https://public-transport-hslhrt.opendata.arcgis.com/, in Finnish only).

The most relevant infrastructure data HSL has are the **bus and tram stops, and city bike stations.** More detailed information about e.g. the stops' **accessibility** exists in HSL registers (Joukkoliikennerekisteri, JORE). It is not openly available, but can be requested from HSL.

Despite not being infrastructure data as such, information like routes and fare zones are geographical data and available via the above links.

Note that HSL uses OpenStreetMap in their Journey Planner and contributes to it by updating walking paths, for example (https://wiki.openstreetmap.org/wiki/Helsinki_Region_Transport).

## Other sources for traffic infrastructure data

Smart mobility pilots and experiments of Forum Virium Helsinki have produced **experimental datasets**, which will be listed under the Mobility Lab Helsinki website during 2022 (please contact Mobility Lab Helsinki for further details).

As an example, a dataset related to traffic infrastructure was created in 2020-2022 to map **routes to buildings and entrances**. This detailed data, mainly for last mile logistics purposes, helps delivery drivers understand which precise door is meant for deliveries to a given address, as it can often be "around the corner" in another street or in a courtyard. This data is stored and available at OpenStreetMap, and for demonstration purposes an example application (**app.gatesolve.com)** has been created to visualise entrances and routes of the last 50 meters to them.

# Traffic

## Traffic volumes

Helsinki has **traffic volume data** from "motorised transport" (i.e. from motorcycles and scooters to cars of all sizes), cycling and pedestrians. At the moment this information and documentation is available almost solely in Finnish.

**Motorised transport, its datasets, statistics and reports** are explained here: https://www.hel.fi/helsinki/fi/kartat-ja-liikenne/kadut-ja-liikennesuunnHelsingin liikenneittelu/tutkimus-ja-tilastot/moottoriajoneuvoliikenteen-maarat/.
A nice dashboard (incl. the English version) for this data is available at https://helsinki-public.azurewebsites.net/.

**Bicycle volumes** are produced by 22 automatic counters and annual manual countings (https://www.hel.fi/helsinki/fi/kartat-ja-liikenne/kadut-ja-liikennesuunnittelu/tutkimus-ja-tilastot/pyoraliikenteen-maarat/). The English dashboard of the automatic counters is at https://data.eco-counter.com/ParcPublic/?id=5589.

**Pedestrian volumes** are counted manually during summertime and constantly with eight automatic counters, of which six are in the city centre, one in Malmi and one in Itäkeskus (https://www.hel.fi/helsinki/fi/kartat-ja-liikenne/kadut-ja-liikennesuunnittelu/tutkimus-ja-tilastot/jalankulun-maarat/).

**Share of transport modes** is monitored annually on a few main directions entering/exiting the city: https://www.hel.fi/helsinki/fi/kartat-ja-liikenne/kadut-ja-liikennesuunnittelu/tutkimus-ja-tilastot/henkiloliikenne-kulkutavoittain/.

Statistics of **the influence of COVID-19 on travel volumes** (vehicles, bicycles, pedestrians, public transport) is available at https://www.hel.fi/helsinki/fi/kartat-ja-liikenne/kadut-ja-liikennesuunnittelu/tutkimus-ja-tilastot/liikennemaarat-korona-aikana/.

Usage of city bikes can be calculated from the dataset for their OD matrices,

As Jätkäsaari district is the main location for Helsinki's Mobility Lab (https://mobilitylab.hel.fi/), there is some **experimental and/or area-specific data** available from different projects and pilots. One such project has been monitoring the traffic in and out of the Port in Jätkäsaari. Specifically, the traffic unloading from ferries to the street network is available also online: https://helsinki.liikennenyt.fi/opendata/.

Commercial providers have **traffic volume data**. The City of Helsinki has not benchmarked or evaluated that data, but some experiments regarding data validation is planned for autumn 2022, please contact Mobility Lab Helsinki.

## Travel times

Travel times are not available openly. The City of Helsinki uses commercial providers for this data apart from some occasional pilots with e.g. WiFi/BT beacons. A report investigating

**traffic fluency** in 2010-2017 is available in Finnish: https://www.hel.fi/static/liitteet/kaupunkiymparisto/julkaisut/julkaisut/julkaisu-07-18.pdf.

The University of Helsinki's Digital Geography Lab has produced a **Helsinki Region Travel Time Matrix** in 2018 for walking, cycling, public transport and cars. The data may be used freely: https://blogs.helsinki.fi/accessibility/helsinki-region-travel-time-matrix-2018/.

## Routes and OD matrices

OD matrices are not available openly, except for city bikes which are available from HSL: https://www.hsl.fi/en/hsl/open-data#journeys-made-by-city-bikes. Also see this blog post about usage of the data: https://towardsdatascience.com/helsinki-city-bikes-exploratory-data-analysis-e241ce5096db.

On the regional level, HSL conducts travel surveys regularly to create **OD matrices of passenger flows**. The City of Helsinki does not currently collect OD matrix data on vehicle or any other traffic

Related to OD matrices, there is a recent study (2022) by the University of Helsinki on population distribution in the Helsinki Metropolitan Area, with the dataset available (https://blogs.helsinki.fi/digital-geography/2022/02/08/open-spatial-data-reveals-24-hour-population-dynamics-of-people-in-helsinki-metropolitan-area/).

## Public transport

HSL has very good information about both **the scheduled and executed public transport**, described and available via https://www.hsl.fi/en/hsl/open-data.

**Routes and timetables** are available in GTFS format and via the Journey Planner API. Real-time data about alerts and changes in the service is available via an API, as well as real-time location and movements of all vehicles.

The statistics and **anonymised data on movements about the city bikes** (incl. O/D matrices) are available as monthly and annual packages in csv format.

## Digitraffic

National-level traffic information is managed by the government agency Fintraffic, who also host the National Access Point NAP (https://finap.fi/#/).

Most traffic data is available via DigiTraffic (https://www.digitraffic.fi/en/).

Note that while Digitraffic focuses mainly on the **road network** that is owned and maintained by the national government (not cities or municipalities), in Helsinki that includes the ring roads as well as the main arteries leading into Helsinki. These play a significant role in Helsinki's traffic system, and data collected in them is highly relevant.

Digitraffic has run a pilot incorporating data from two Finnish cities. Helsinki's data will likely be available through Digitraffic eventually, but no agreements or a roadmap have been made yet.

### Other traffic data

**Commercial providers,** such as TomTom, HERE, Google/Waze and V-Traffic/Flitsmeister, have good quality data about traffic in Helsinki. City of Helsinki traffic planners use their datasets in their work as well, but due to licensing that data cannot be opened.

Different smart mobility **pilots and projects produce experimental data** about the traffic. These will be listed under the Mobility Lab's website (https://mobilitylab.hel.fi/materials/).

There are a number of data that is of interest, but which do not exist currently. An example of this would be the length of queues (e.g. in congested intersections or main highways leading into the city).

## Conditions and context

### Accident and incident information

**Accident statistics**, with links to further data, all in Finnish, are gathered in https://www.hel.fi/helsinki/fi/kartat-ja-liikenne/kadut-ja-liikennesuunnittelu/tutkimus-ja-tilastot/liikenneonnettomuudet/. Some data is also available at https://kartta.hel.fi/#.

The City of Helsinki Traffic Accidents Register ('Helsingin liikenneonnettomuusrekisteri', TARE) contains further **data about accidents** that the police has investigated. This data is not public but is available per request in a limited version due to GDPR issues.

Digiroad's host organisation Väylä also provides statistics about **road accidents** in mainland Finland. This information is only available in Finnish (https://www.avoindata.fi/data/fi/dataset/tieliikenneonnettomuudet).

### Maintenance and road works

There is some information about **road maintenance** in Digiroad, although it is not very comprehensively updated by Helsinki. For the national roads (i.e. in Helsinki these are ring roads, main arteries), however, this information can be useful.

There is an **experimental interface related to incidents and roadworks** around Jätkäsaari district with Datex2-based API available at https://helsinki.liikennenyt.fi/opendata/.

STARA has been experimenting with telemetry as part of the fleet (e.g. tractors and multi-purpose maintenance vehicles) and has fairly comprehensive telemetry installed. The first open data from this telemetry was an open API for snow-ploughing activities in winter (https://dev.hel.fi/apis/snowplows), which has become very popular via an external website (https://auratkartalla.com/). The goal is to supplement the content of the service with the entire urban maintenance environment equipment during the coming winter seasons 2023 - 2024.

More notably, there is an intention that STARA's vehicle fleet could act as a platform for measurements, i.e. sensor installation. As this is a work-in-progress, the current possibilities are available per request.

## Weather and road conditions

Digitraffic (https://www.digitraffic.fi/en/) as the main national source for traffic information also provides data about the **road conditions**. Namely, road weather cameras and weather station data are among the most used in Finland, and there are plenty of them around Helsinki and in the main highways leading into Helsinki. Incident information is nationwide but applicable to Helsinki as well.

The Finnish Meteorological Institute FMI has a long history of leading the field of traffic weather, with their **Road Weather Model and Pedestrian Weather Model** being of interest worldwide. They provide extensive open data (https://en.ilmatieteenlaitos.fi/open-data).

Besides the more standard weather and air quality information, it is worth highlighting the Finnish Meteorological Institute's **weather warnings for pedestrians** that shows conditions of walkways, i.e. warnings if they are turning slippery.

## Air quality and noise

Helsinki provides **noise mapping** of the City of Helsinki ('Helsingin melutiedot'). It has data about the noise measurements carried out every five years. The latest one is from 2017, and an update should be done during 2022.

Helsinki Region Environmental Services HSY provides a good amount of interesting open data and APIs: https://www.hsy.fi/en/environmental-information/open-data/.
The most relevant to mobility are **GHG and traffic emissions** and **Air Quality information**. They also provide some **travel time and accessibility data**.

The University of Helsinki's **MegaSense programme** enables solutions by gathering and fusing spatially variable **gas and particulate measurements** from high-end scientific instruments, commercial air quality transmitters, dense low-cost sensor arrays, and consumer wearables. You can find further information about the specific datasets and APIs **here**.

## Public transport interfaces

HSL's OpenMaaS API offers a **sales interface for both single and season tickets** available to all operators interested in transport services. More information: https://sales-api.hsl.fi/.

HSL operates the **Park&Ride places** in and around Helsinki. Their information system is also online with an APi and documentation at https://p.hsl.fi/docs/index.html.

HSL has **bluetooth beacons** in most bus and tram stops as well as vehicles. Their signal is free to use for any purposes (https://hsldevcom.github.io/beacon_api/).

The Journey Planner APIs also contain i**nformation about cancelled services**, for example the routing API's disruption-info and the real-time API's service-alerts.

### Other context and condition data

Helsinki Region Infoshare provides information about public street events (https://hri.fi/data/dataset//helsingin-kaupungin-yleisten-alueiden-tapahtuma-ja-maankayttolupajarjestelma).

Public web cameras are provided by the Port of Helsinki (https://www.portofhelsinki.fi/en/port-helsinki/web-cameras) and Finntraffic (https://liikennetilanne.fintraffic.fi/listanakyma?list=cameraPictures).

Smart mobility pilots and experiments of Forum Virium Helsinki and the City of Helsinki occasionally result in **pilot data and datasets** that may be relevant for mobility development. These will be listed under the Mobility Lab Helsinki website (https://mobilitylab.hel.fi/materials/).

## Data gaps, preconditions and barriers of development

Combining experiences from the ongoing smart mobility and digital twin projects of Forum Virium Helsinki (i.e. LiiDi2[1] and Mobility Lab Helsinki), prior projects and discussions with city officials and companies, some gaps and barriers for developing the digital twin for mobility have been identified.

**Regional data boundaries**

The datasets concerning traffic and mobility are split across a number of services following organisational boundaries between e.g. the cities of Helsinki and Espoo and national actors, such as Fintraffic. Between different providers, there are varying systems and formats used. Inside the city, the documentation of different layers available over the APIs is not always available in English. The Helsinki Region Infoshare (HRI) open data portal provides open data, but a similar metadata service is missing for the closed assets of the City, or those from commercial providers.

**Commercial data and privacy**

The procedures of operating with private data providers in an ecosystemic fashion are still very immature, and the role of public administration in this ecosystem is not yet fully clear.

[1] LiiDi2 project website: https://forumvirium.fi/en/projects/liidi2-the-mobility-digital-twin/

This complicates using closed data sources in City operations and b2b actions involving data trade. The City has a stated aim of supporting companies, but for data trade, this is yet to be articulated as clear actions.

Apart from restrictions related to business interests and commercialised data products, privacy issues are a major concern for what data can be shared. Traffic and mobility related data is often personal information by nature, often involving e.g. the users' mobile phone or vehicle along with location information. The consent procedures for ecosystemic data use are not well developed. The use cases and justification for using the data need to be defined and appropriate anonymisation and aggregation methods applied.

**Awareness of what is available**

Awareness of the types of data available and their potential uses can help define the cases and forms of sharing and collaboration also for the commercial datasets. Examples, pilot cases and demonstrations are a good way to communicate and illustrate possibilities, increase awareness of what is out there and spark new ideas.

Finding more users and use cases for data can help justify its production. More versatile use of data means more value out of it, covering the acquisition costs related to sensors or data procurement.

**Standards**

Concerning the data itself, the set of exchange formats that should be offered to support most applications is not yet well known. From a standardisation point of view, the phenomena related to mobility can be approached from geospatial data standards (e.g. OGC CityGML), those originating from the mobility sector (e.g. Datex II), or those coming from sensor technology (e.g. MQTT). There are also more context-specific data formats such as GTFS or OpenDrive.

**Interfaces between platforms and services**

A significant amount of technical development is also needed. Private data producers tend to provide their data in their own dashboards, which do not always support integration or combination with other data sources. Platforms able to combine the different data types concerning mobility are needed. Simulation methods that support understanding a multimodal system are required to turn data into useful predictions. Interfaces for using and providing mobility services are needed for supporting the growth of the mobility ecosystem.

**From data to information to services to interoperability and impact**

While data itself is a necessary building block for different services, it needs to be useful and usable as well. More refined and practical, commonly needed information accessible by simple APIs reduces barriers for developers. If integrating a new data source for a different location, for example, requires lots of manual effort, it may well be seen as inefficient and not worthwhile to scale a service. Besides data, more functional APIs for e.g. booking or managing services are important. In other words, going from PDFs to CSVs to APIs are significant steps forward in making data not just digital but useful.

Similarly, from a user point of view, all the bits and pieces in the value network need to be in place from data production to integration, analysis, service development and end user interfaces. Developing an impactful service with a valid business case around data therefore needs to include all the relevant actors, ensuring there are no gaps along the way.

**Quality and expiration dates**

Different purposes have different temporal and spatial requirements for data. For traffic planning, historical data is likely more useful than accurate real-time data, which in turn is more suitable for dynamic traffic management. For understanding and navigating the urban environment, up-to-date information about the infrastructure or its changes as well as real-time disruptions and available services are important. The type of information, and how accurate either to the scale (centimetre/district) or time (second/year) it needs to be, varies by the use case.

## Potential applications

Digital twin for mobility enables diverse applications and use cases (Figure 10). In fact, it is hard to imagine a mobility innovation which would not make use of data about traffic, traffic infrastructure, or its conditions and context.

New data utilising services have potential in facilitating smooth, safe and sustainable mobility, and better use of urban space. Data and digital twins support evidence-based planning and regulation, leading to fluent mobility, a vital city centre and a pleasant urban space. They also open up new ways of citizen participation and crowd-sourced data production.

Importantly, data helps private companies develop new services, optimise their operations and choose locations of their businesses, for example. So the benefits of developing the digital twin extend beyond the traditional traffic sector, e.g. in retail, events, tourism and robotics.

Mobility Lab Helsinki aims to support the development of new innovative solutions, utilising digital twins and data. Due to the diverse and complex nature of the digital twin, the approach selected is to develop and utilise the digital twin through concrete use cases. Instead of developing massive ICT systems, clearly defined use cases are sought, which can be elaborated into concrete pilots and experiments.

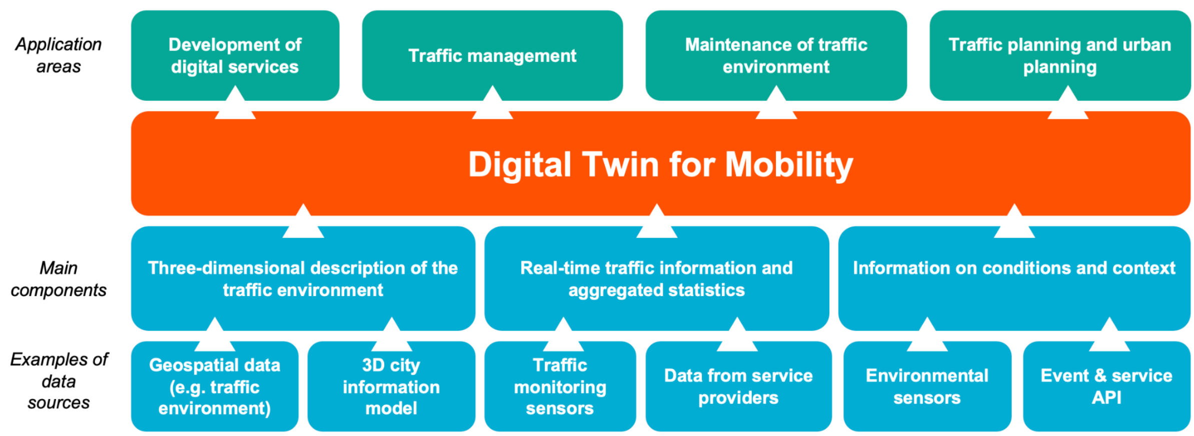


*Figure 10. Conceptual overview of data and applications in digital twin for mobility.*

## Examples of use cases

Forum Virium Helsinki and its projects (Mobility Lab Helsinki and LiiDi2) have identified potential use cases in which the digital twin enables the development of new services, solutions and innovations (Table 2). These have been identified in workshops and discussions with companies, city representatives and other experts.

The table is not comprehensive. It is a list of examples that have been addressed in the many discussions with different stakeholders, illustrating the relevance of the different data types. The table is mainly focused on data needs of the use cases, not specifying functional APIs such as ticketing interfaces or booking of parking spots that are needed to provide and manage the different services. That is to say, while data can be enough to provide a demonstration of a proof-of-concept of a service, cooperation with relevant stakeholders is often required to go beyond delivering just information services.

It is worth highlighting that not all potential use cases are even known yet. Mobility Lab Helsinki encourages contacting us with any idea of utilising data or digital twins for mobility innovations.

The data examples in the use cases (right column) are not limited to or indicative of how well they are available at the moment. Rather, they serve as a starting point for discussions on common topics of interest, i.e. if someone has a solution or need related to a use case, Mobility Lab Helsinki will support in refining that further, identifying the current availability of the suitable datasets, as well as the gaps that need filling.

*Table 2. List of example use cases identified with the stakeholders in the LiiDi2 and Mobility Lab Helsinki projects.*

| Potential use cases | Examples of data |
|---|---|
| **City logistics**<br>- Improved efficiency and flow of deliveries through accurate and detailed information on traffic, | **Infrastructure**: availability and accessibility of loading/unloading locations, detailed locations and features of entrances, gates and stairways, parking locations and |

| | |
|---|---|
| loading/unloading locations, and entrances | restrictions, traffic signs<br>**Traffic**: real-time traffic situation, real-time and predicted availability of loading places<br>**Conditions and context**: weather, winter maintenance |
| **Urban transport system modelling, planning and impact assessment**<br>- Modelling and impact assessment of new constructions, interventions and policies (e.g. land use planning, traffic planning, low-emission zones, road use and congestion charges, tunnels, construction works, new services)<br>- Analysis and simulations of traffic, parking, businesses, vitality of city centre, emissions, energy use, etc. | **Infrastructure**: road network, buildings, infrastructure, accessibility for people and businesses, capacity<br>**Traffic**: traffic flow per mode and type, people flow<br>**Conditions and context**: demographic and socio-economic data, car ownership, travel habits (travel times and demand for businesses, schools, shopping, logistics), winter maintenance (e.g. room for snow), road works and alternative routes |
| **Mobility and transport services**<br>- Optimised operations and service provision (on-demand, others) | **Infrastructure:** parking spots, stations, accessibility, routes<br>**Traffic:** traffic situation, congestion, people flow<br>**Conditions and context:** expected demand (e.g. events, festivals, ferry arrivals), points-of-interest (hotels, etc.), complementary services |
| **Infrastructure maintenance**<br>- Optimised street maintenance (e.g. winter maintenance, up-to-date traffic signs, road condition monitoring and maintenance) | **Infrastructure:** roads, pavements, traffic signs<br>**Traffic:** traffic flow, capacity and alternative routes<br>**Conditions and context:** weather, citizen feedback (i.e. perceived quality of routes and maintenance) |
| **Dynamic use of curb space**<br>- Dynamic curb space management based on time, need, priorities, etc. (e.g. for allocating or booking space for parking, logistics, walking or managing dynamic demand-driving pricing) | **Infrastructure:** parking and loading spot information including times, prices, accessibility<br>**Traffic:** traffic flow, people flow, curb use status<br>**Conditions and context:** disruptions, up-to-date rules (e.g. curb space use), demand data (logistics, parking, micromobility) |
| **Navigation and wayfinding**<br>- Information, routing and navigation tools for citizens with specific requirements and functional limitations (e.g. the visually impaired, wheelchair users, baby | **Infrastructure**: accessibility (on the street, stations, buildings, elevators), accessible routes, detailed structure of pavements, steps, doors and curbs<br>**Traffic**: real-time and schedule information for transport services, accessibility of vehicle types<br>**Conditions and context**: transport service |

| | |
|---|---|
| strollers, language barriers, cognitive challenges, etc.) | accommodations and features (e.g. accessibility, service dog, t-coil availability), disruptions, roadworks |
| **Traffic management and service operations**<br>- Dynamic traffic management: alleviating congestion, prioritising active and green mobility | **Infrastructure:** routes, lanes, traffic light status<br>**Traffic:** traffic flow, capacity and alternative routes, congestion<br>**Conditions and context:** weather, citizen feedback (i.e. perceived quality of routes and maintenance), roadworks, disturbances, demand for mobility services |
| **Traffic light optimisation**<br>- Improving traffic flow based on situational awareness through active traffic light optimisation according to different priorities (e.g. active modes, public transport, emissions, congestion) | **Infrastructure**: road network (cars, bikes, walking), pedestrian crossings, traffic light status<br>**Traffic**: real-time and historical traffic flow data (i.e. speeds, routes, all modes)<br>**Conditions and context**: weather, time of day/week/year and holidays, disruptions and road works, expected demand (e.g. events and festivals, ferry arrivals) |
| **Micro-mobility management**<br>- Support to planning and regulation<br>- Development of multimodal travel chains and mobility hubs | **Infrastructure:** cycling lanes, road network, parking areas<br>**Traffic:** traffic and people flow data, number of e-scooters, locations, O/D matrices, routes, curb clutter<br>**Conditions and context:** traffic rules, condition of bike lanes (e.g. winter maintenance), public transport use, service maps, accident data, geofenced micromobility regulation (e.g. speed limits, no-park zones) |
| **Urban air space and drones**<br>- Urban air management and operations | **Infrastructure:** landing areas, no-fly zones, buildings and line-of-sight information, flight zones and corridors<br>**Traffic:** air traffic, traffic and people flow data (i.e. for risk analysis)<br>**Conditions and context:** local rules and regulations, weather, 5G coverage |
| **EV charging**<br>- Planning and modelling charging infrastructure needs and requirements to match the demand in the near-, mid- and long-term<br>- Current and predicted availability information on charging stations | **Infrastructure**: energy grid data, energy demand, road network, parking (on-street, housing, businesses)<br>**Traffic**: parking data, traffic patterns,<br>use of taxi stands,<br>**Conditions and context**: car ownership, charger use, demand for charging (private, public, business use cases), socio-economic data |

| | |
|---|---|
| **Automated driving and robot logistics**<br>- Operating automated vehicles in the urban environment<br>- Operating automated robots for package delivery, street sweeping, park maintenance, etc. | **Infrastructure:** Accurate road information, traffic signs and lights, speed bumps, curbs, parking areas, accessibility, building entrances, elevators, green areas<br>**Traffic:** real-time traffic and vehicle data, people flow<br>**Conditions and context:** disruptions, road works, alternative routes, local rules and regulations |

## Mobility Lab Helsinki supports piloting and experimentation

Mobility Lab Helsinki invites all parties interested in digital twins and mobility data to contact us in order to support concrete pilots and experiments. Whether it is a preliminary idea or a more developed innovative technology or service, Mobility Lab will support bringing it forward towards a concrete pilot or experiment. We have an open approach to proposals and always find the best way forward with each solution, case by case.

The Lab supports piloting in many ways: facilitating discussions with stakeholders and potential end-users, helping out with permits, and identifying and providing already available data. We will help you with the data gaps related to piloting needs. Open calls for solutions both for building the base of and utilising digital twins will be launched, starting in autumn 2022. A data catalogue will be built, offering links to relevant data sources to companies.

Another project developing the digital twin and supporting its use for developing new innovations is *LiiDi2 - Digital Twins enabling mobility innovation.* The project is conducted by Forum Virium Helsinki and the City's building and construction services Stara. The project is focusing mainly on the infrastructure aspects of the digital twin, and includes open calls for data collection pilots, innovative procurements, hackathon, development of IoT and data platforms, and stakeholder workshops on digital twin development. The experiences, learnings and outputs of the LiiDi2 project have contributed in the content of this document.