Tampere University
Tampere University of Applied Sciences

BUSINESS TAMPERE

Lessons learned through project Älypysäkki (Smart TrafficHub)

European Union
European Regional Development Fund

Leverage from the EU 2014–2020

COUNCIL OF TAMPERE REGION

This document summarises the phases of **Älypysäkki -project** (Smart TrafficHub, 2.3.2020 - 16.4.2021) funded by Council of Tampere Region with European Regional Development Fund

The project was organised from Tampere University if close cooperation with Tampere University of Applied Sciences and Business Tampere. This report was written in cooperation with all stakeholders.

Tampere University
Riku Viri
Rodrigo Borges
Jyrki Nummenmaa
Kostas Stefanidis

Tampere University of Applied Sciences
Jukka Pellinen
Jari Ruokolainen
Ossi Nykänen

Business Tampere
Markku Niemi
Minna Kinnunen
Jari Ikonen

Cover image by Jonne Renvall, Tampere University

Contents

1. Background

At first, the identified present state regarding smart traffic hubs are discussed. In addition to present state, this chapter depicts the different stakeholders and other existing concepts of the topic.

1.1. Present state

A smart traffic stop is a place where people start or end their journeys, switch (public or private) transports, pay for trips and other services, wait, and while at it, access information about the journey and the nearby services. In Finland, simple traffic stops intuitively follow categorizations, ranging from terminal nodes and busy stops, to unused traffic stops, and offer various kinds of (informative) basic passenger services[1]. Indeed, in general, a traffic stop might be generalized into a traffic hub, a point integrating many related services and access points, and located in a central position in a city.

At a European level, local traffic services are part of a larger network services and social needs. In the bigger picture, the main challenges for the transport sector in the EU include creating a well-functioning Single European Transport Area, connecting Europe with modern, multi-modal and safe transport infrastructure networks, and shifting towards low emission mobility, which also involves reducing other negative externalities of transport. From a social perspective, affordability, reliability and accessibility of transport are key. However, this has not been achieved across the board. Addressing these challenges will help pursue sustainable growth in the EU[2].

Figure 1 **>** *Existing smart traffic stop in Paris, image from Smart Cities Dive*[3]

Figure 2 **>** *Existing smart traffic stop in Monaco, Europe, image from Extended Monaco*[4]

Figure 3 **>** *An existing smart bus stop in Jurong East, Singapore, image from Japheth Lim*[5]

Smart traffic stops have been extensively studied in the past. In brief, the existing examples provide information and services related to buying tickets, renting bikes or ordering taxis, and charging smart phones. More advanced bus stops might aim better integration with the neighborhood and enjoying ones' stay at the stop (buy a meal, listen to music, loan a book, acoustic design, etc.); see Figure 1, Figure 2 and Figure 3.

1 Airaksinen, S., Lehto, A., Oikarinen, I., & Kataja, A. (2015). Joukkoliikenteen palvelutason määrittely, Liikenneviraston ohjeita 31/2015, Liikennevirasto, Helsinki; Weiste, H., Mantila, A., & Seila, M. (2014). Valtakunnallinen pysäkkiselvitys-pysäkkiverkot ja pysäkkien palvelutaso. Liikenneviraston tutkimuksia ja selvityksiä: 43/2014

2 EC 2018; EC (2018). Transport in the European Union: Current Trends and Issues. Available online

3 What if bus stops were designed as if bus stops really mattered? at Smart Cities Dive

4 Smart bus shelters: a clever combination of infrastructure and digital technology at Extended Monaco

5 Jurong East Smart Bus Stop – Redefining Bus Stop and Community Space Making at Japheth Lim.com

Further, at least conceptually, there is an increasing trend in moving from traffic stops into social traffic hubs[6].

In the more abstract sense, this work might be seen as a part of studying smart traffic in general, as seen e.g. in Fiore et al. (2019)[7].

When introducing mobile services into the picture, traffic stops and the related services may become more virtual. Passenger information surveys indicate that people seek traffic information from numerous sources – notably at the traffic stops or from personal mobile devices. It is also possible to organize passengers into meaningful groups or profiles, such as visitors vs. locals, or young vs. elderly passengers, which then may indicate the main passenger behaviors and needs.

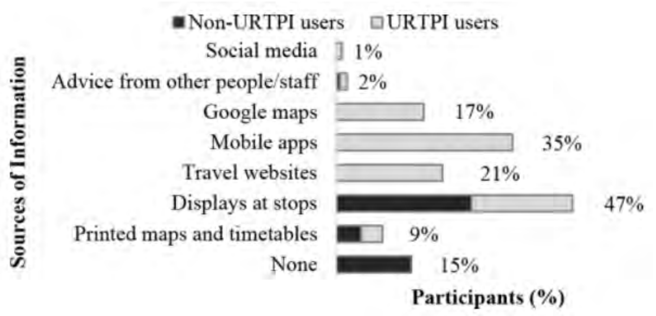

Figure 4 **>** *Use of different sources of passenger information (URTPI means "Ubiquitous real-time passenger information"); image from Md Faqhrul Islam et al. (2020)[8]*

It is important to note that de facto passenger information sources are rather fragmented. Figure 4 above depicts the use of different sources of passenger information, as reported in the survey Md Faqhrul Islam et al. (2020) considering Edinburgh area, UK[8].

By analyzing passenger information, useful statistics may be identified. According to Md Faqhrul Islam et al. (2020)[8], significant background variables indicating the use of information services included length of the trip, reason for travel, and familiarity of the trip. Considering passengers themselves, meaningful background variables included age, profession, and home city. Properties such as time of travel, alternative route, gender, or education did not have a significant effect.

When additional transport data becomes available, it also becomes possible to analyze large-scale passenger behavior and make statistical predictions about traffic volumes and even individual passenger destinations[9].

Ideally, bus transit boarding stop information using smart card transaction data should be available[10]. Unfortunately, this is not the case currently, e.g., in the Pirkanmaa region.

In general, it seems that traffic stop and traveling information design can have significant influence on passengers (by default people choose busiest line instead of the fastest line), but changes may take some time – for instance, in London, only 50% of the passengers have been reported to recognize new information screens after one year of operation. Transportation systems and the related information needs, however, are rather complicated, and a single stakeholder is typically unable to provide a uniform service to suit all the passenger needs.

6 c.f. Padrón Nápoles, V. M., Gachet Páez, D., Esteban Penelas, J. L., García Pérez, O., García Santacruz, M. J., & Martín de Pablos, F. (2020). Smart Bus Stops as Interconnected Public Spaces for Increasing Social Inclusiveness and Quality of Life of Elder Users. Smart Cities, 3(2), 430-443; What if bus stops were designed as if bus stops really mattered? at Smart Cities Dive; Jurong East Smart Bus Stop – Redefining Bus Stop and Community Space Making at Japheth Lim.com

7 Ugo Fiore, Adrian Florea, and Gilberto Pérez Lechuga, 2019. An Interdisciplinary Review of Smart Vehicular Traffic and Its Applications and Challenges. Journal of Sensor and Actuator Networks, 8, 13; doi:10.3390/jsan8010013

8 Islam, M.F., Fonzone, A., MacIver, A., & Dickinson, K. (2020). Use of ubiquitous real-time bus passenger information. IET Intelligent Transport Systems, 2020-03, Vol.14 (3), 139-147.

9 Silva, Ricardo, Soong Moon Kang, and Edoardo M. Airoldi. 2015. "Predicting Traffic Volumes and Estimating the Effects of Shocks in Massive Transportation Systems." Proceedings of the National Academy of Sciences 112 (18) (April 20): 5643–5648. doi:10.1073/pnas.1412908112.; Jaeyoung Jung, Keemin Sohn, 2017. Deep-learning architecture to forecast destinations of bus passengers from entry only smart-card data. ET Intelligent Transport Systems, ISSN 1751-956X, doi: 10.1049/iet-its.2016.0276

10 Zhen Chen, Wei Fan, 2018. Extracting bus transit boarding stop information using smart card transaction data. J. Mod. Transport. (2018) 26(3):209–219, doi:10.1007/s40534-018-0165-y

1.2. Stakeholders

When considering the context ITS services, the following stakeholders of may be identified (examples of the Smart TrafficHub stakeholders in the Tampere sub region provided as an example):

1 > Main public ITS customer (city of Tampere);
2 > Main ITS service providers (tram and bus operators);
3 > Target (new) ITS service providers (traffic stop operators with subcontractors);
4 > Auxiliary service providers (coffee shops, taxi providers, and electric bike or scooter service companies);
5 > End users (notably people using the main and the auxiliary services).

In a nutshell, the main public ITS customers decide the public transportation system, eventually operated by bus companies and other service providers. Due to its significancy, the transportation system not only directly serves the people living in the region as end users: The transportation system also interacts with the broader region development and city planning and may thus indirectly influence peoples' conceptions about certain city regions, and hence have an impact on, e.g., apartment pricing.

Of course, from the end users' point of view, the transportation landscape should ideally appear as a "single" entity, intuitively fulfilling the need of journey planning and getting from one place to another, while benefitting from various kinds of services along the way. From the city and the service providers perspective, however, the big picture is more complicated, perhaps without an explicit agreement of "ownership" of the various aspects or legs of a journey. This is notably the case when combining private and public services from different operators or providers (including the travellers' own cars, bicycles, etc.).

In the case of the Smart TraficHub project, the role of the traffic stop operators and new related services gets highlighted simply due to the focus area of the project. In more general, the distinction between the main, the target (new), and the auxiliary service providers is of course context dependent.

1.3. Lessons learned from current concepts

Due to its significance, the efforts for integrating smart transportation services have also been supported by the European Union. For instance, the project "European Smart Mobility Resource Manager"[11], or MyWay for short, investigated, developed and validated an integrated platform, the European Smart Mobility Resource Manager, including cloud-based services and facilities to support community supplied information collection and processing.

As a result, the MyWay project consortium identified the high-level conceptual architecture for MyWay and other services depicted in Figure 5. As such, the conceptual architecture is still valid today, and highlights the generic need for orchestrated coordination of different smart traffic activities – and a need for common interfaces, shared infrastructure and high-level objectives, with aligned business models.

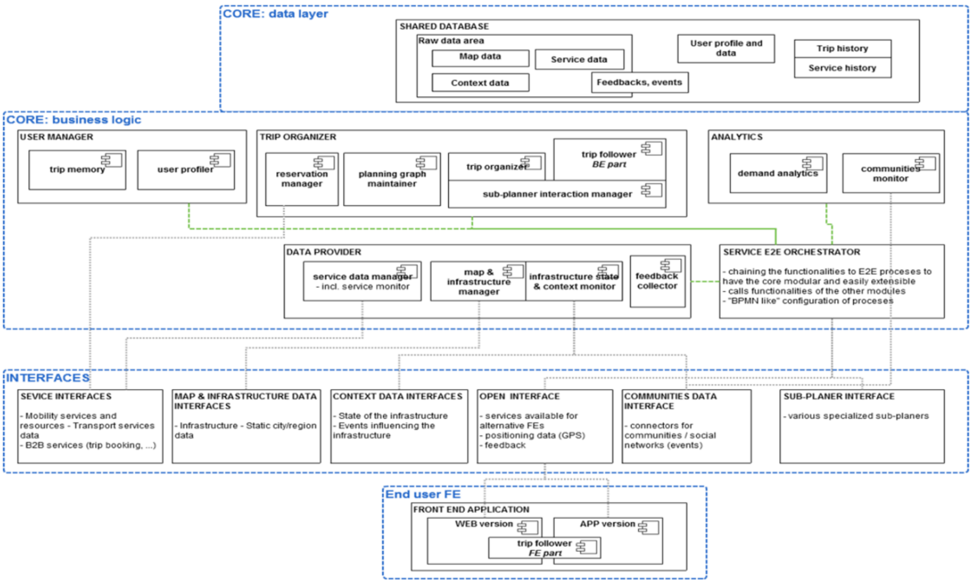

***Figure 5** > MyWay conceptual high-level architecture*[12]

11 European Smart Mobility Resource Manager, 2013-2016, FP7-SMARTCITIES-2013

12 Fernández, J. & Bures, M. (2013) D2.1 Requirement specification and analysis of MyWay: European Smart Mobility Resource Manager

2. Current data availability

The city of Tampere and the nearby stakeholders actively publish data and applications which can also be exploited in ITS development[13].

The data sources are organised in into the following five categories related to the city of Tampere area: Maps, activity data, position data, smart passenger card data, and application data. Next, these will be discussed each category in turn in the context of the Smart TraficHub project. In order to provide meaningful examples, first, the outline data sources from the perspective of existing end user applications is outlined.

Maps

The maps category includes map resources with several layers of information, including transportation and other service information. Aerial photography map views (etc.) are also available.

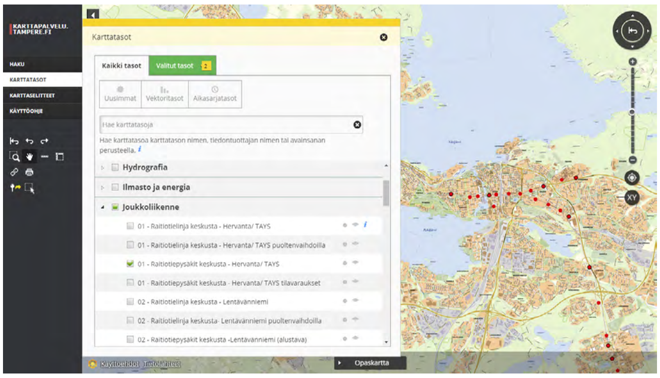

Figure 6 **>** *Screenshot from Tampere Oskari map service with some tramway stops highlighted*

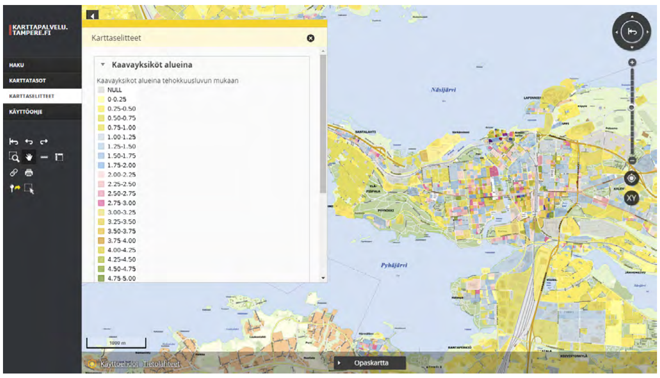

Figure 7 **>** *Screenshot from Tampere Oskari map service with building efficiency ratios by city block*

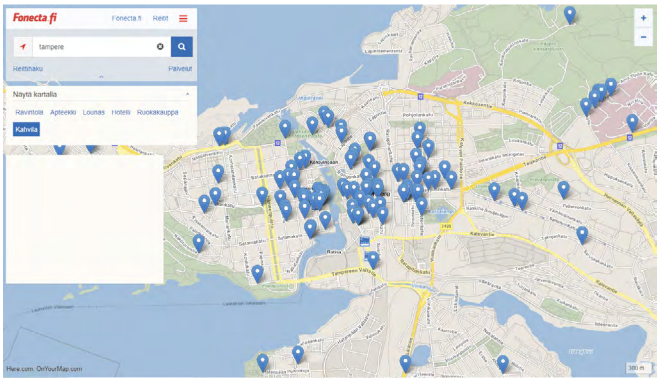

Figure 8 **>** *Screenshot from commercial Fonecta.fi service: Cafes on Tampere map*

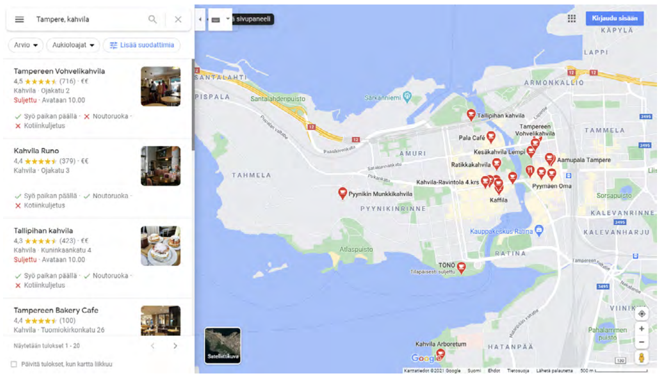

Figure 9 **>** *Screenshot from commercial Google Maps service: Cafes on Tampere map*

The city of Tampere provides map browsing and annotation service called Oskari, in accordance with the INSPIRE directive (see Figure 6). Oskari service provides access to a rich system of layered map information. The service includes (all) public transportation, speed limits, bicycle roads and bicycle parking areas, administrative regions, schools, public healthcare and sport services, and property and city planning information, to name a few. Official city planning information is also available, such as building efficiency ratios by city block (see Figure 7). In practice, the map information provided by the city of Tampere is very extensive, and the essential content is available in machine-readable format at the city datasets portal[13].

As suspected, the public information does not cover, e.g., most services offered by private companies. Such information may be available from commercial providers, such as Google or Fonecta.fi. Figure 8 and Figure 9 illustrate the (incomplete) list of cafes in the Tampere city center, as offered by Fonecta.fi and Google Maps. Similar queries are available, e.g., for grocery stores, (lunch) restaurants, etc.

13 For portal displaying open data from Tampere region, see data.tampere.fi

Activity data

The activity data category includes data, statistics, and projections compiled from or for transport services.

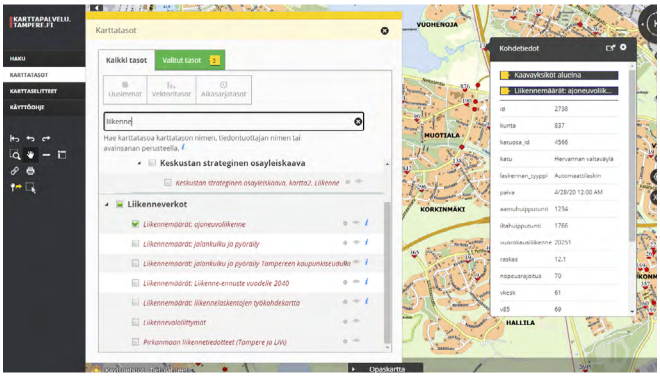

Figure 10 > *Screenshot from Tampere Oskari map service with vehicle volume data sensors and the Hervannan valtaväylä street sensor (id 2738) selected (some data archived and not real-time)*

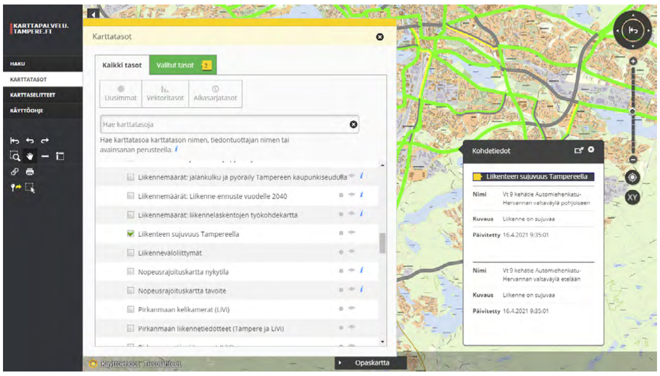

Figure 11 > *Screenshot from Tampere Oskari map service with traffic indicator information and the Route 9, Tampere Ring Road motor highway selected (with real-time slow traffic data)*

The aforementioned Oskari service from the Tampere, also provides activity and projection data. In the ITS context, relevant information includes, e.g., vehicle, light traffic, and road construction information. Figure 10 depicts location data of the various kinds of sensors measuring road vehicle volumes. The data includes, e.g., numbers and types of vehicles, speed limit, and average speed. Figure 11 depicts main road information, associated with normal or slow traffic.

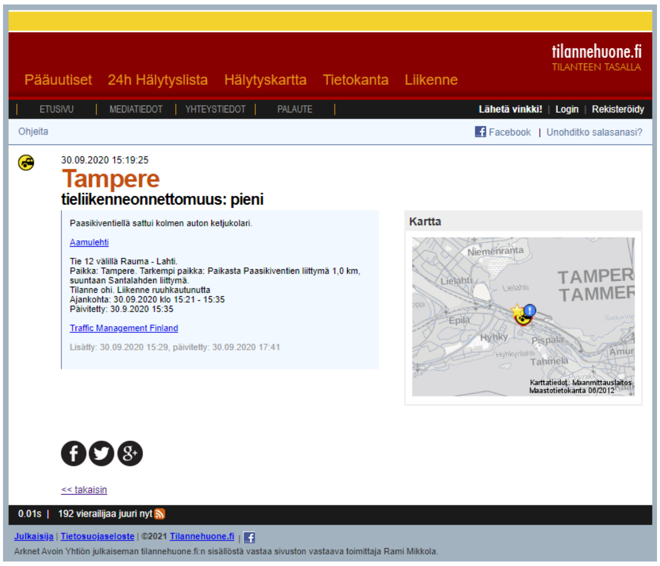

Figure 12 > *Screenshot from tilannehuone.fi showing a traffic accident on Tampere Paasikiventie street (information potentially updated in real-time)*

There are also several additional activity information sources, including traffic accidents, traffic jams, and special incident reports. Figure 12 demonstrates an example of a small, multiple collision accident information at Tampere, as provided by commercial tilannehuone.fi.

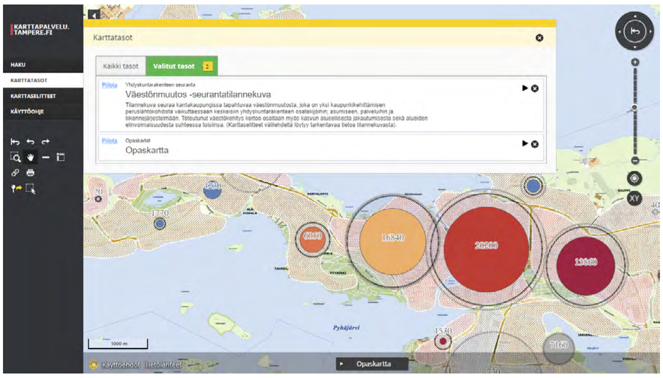

Figure 13 > *Screenshot from Tampere Oskari map service with demography projection, visualizing the change of people living in certain city parts (change of people vs. room for growth in terms of city planning)*

The city of Tampere also provides indirect, long-term activity data in terms of projections. Figure 13 depicts a visualisation of demography changes and planning potential in different parts of the city. Such information might be used in recognizing or planning city areas where additional services are required in the future. For instance, the visualization shows significant increase in people living in the city center (large filled circles with big numbers), and the fact that there is little room left for additional growth (the outer circle rims around each filled circle).

Position data

Some of the (public) transportation vehicles are equipped with location sensors, and provide real-time information about the positions[14]. The data consists of every active bus separated by line number, vehicle and direction and contains the latitude, longitude, timestamp and speed of the buses.

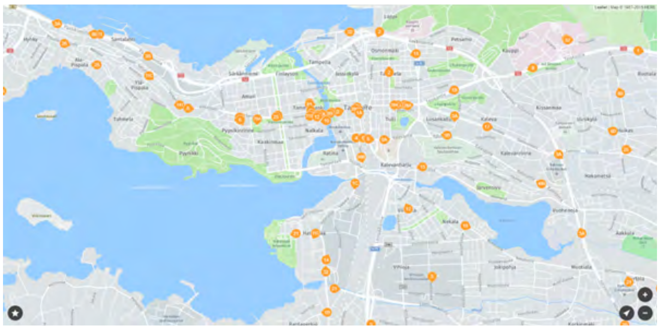

Figure 14 > *Screenshot from busse.fi showing real-time Tampere city bus position information projected on Tampere map*

As an example, Figure 14 provides a visualization of the real-time bus position information, as an aggregation service by Kimmo Brunfeldt, based on the public data provided by the city of Tampere. In addition to bus information, the city of Tampere also provides, e.g., position data of certain kinds of road maintenance vehicles, such as snowplows.

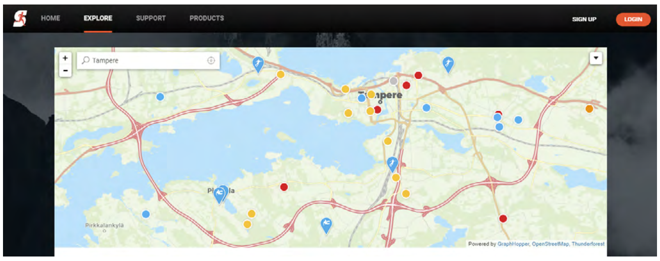

Figure 15 > *Screenshot from public Sports Tracker information around Tampere*

In addition to the publicly available information, other position information might collected as well. For instance, taxi companies and emergency vehicles include position sensors, but the related data is not by default open to the public. Individual people may also distribute their own data, e.g., related to sports activities, such as (cross-country) bicycling or orienteering; see, e.g., Figure 15 with Sports Tracker information.

Smart passenger card data

An important special case of activity data is established by collecting by bus trip payment or smart passenger card data. The city buses currently collect passenger hop-on information, but not hop-off information. The daily passenger activity data is collected by each bus and uploaded overnight to the central data repository. Passenger data is not publicly available. By definition, however, smart card activity data would potentially enable tracking the entire journeys taken by passengers, and not just individual parts of a journey. In addition, buses may collect all sorts of other diagnostic data, including energy consumption, which is of special interest with respect to the new electronic bus(es).

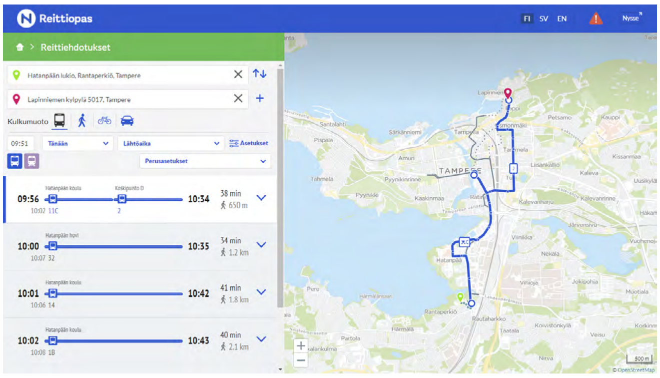

Figure 16 > *Screenshot from Tampere Reittiopas itinerary suggestions*

Application data. In some cases, data sources might be readily available to solve particular ITS application problems, notably journey planning (as an end-user application example, please see Figure 16 above).

The city of Tampere offers access right to journey planner interface, when application or service supports public transportation usage and transportation information availability. Public Transport Journey Planner TRE API offers route, timetable and stop information from Tampere Region[15]. At the time of writing, however, developer information was not accessible. A GTFS-feed of public transport is also available, supporting development where machine readable timetable and route data is needed[16]. The GTFS-feed itself also support some more static data that can support the development of routing, such as transfer connection abilities withing two stops, wheelchair boarding abilities and fare classes and fare calculation[17]. However, not all supported data, such as the wheelchair access, are available in Tampere region within GTFS-feed.

14 Journeys API, information available at ITS Factory Wiki

15 Public transport journey planner API for the Tampere region available at data.tampere.fi

16 Tampere Public Transport GTFS feed, information available at ITS Factory Wiki

17 GTFS dataset reference available at Google Transit APIs

We will not consider this application further in this article, but it is worth noticing that hybrid (AI) ITS applications might indeed directly access sophisticated ITS services in their design. Thus, instead of only accessing "raw" map, activity, and position and usage data, AI applications might also have access to full-fledged service components like journey planning.

Of course, having nice end-user applications (as above) does not suffice for making actual data-driven AI applications: In the abstract sense, the data needs to be

1 > accessible (technical read access and permissive license);
2 > machine-readable;
3 > understandable (at least for the human developer but ideally machine-understandable, i.e., equipped with machine-readable metadata or semantic descriptions);
4 > reliable;
5 > properly scoped (coverage and resolution);
6 > up to date.

3. Data based traffic forecasting test

To test the hypothesis that buses' activity presents patterns, both in space in time, and that if provided with enough historical data, it would be possible to predict future traffic conditions, a prototype was implemented. The prototype for predicting GPS coordinates of buses based on previous activities that can be described in three main phases: (I) collected and pre-processed data from the publicly available API, (ii) compiled data to build images containing the location and speed of buses, and (iii) applying these images as the input for Spatiotemporal model based on Convolution and Recurrent Neural Networks. Our prototype[18] was tested by predicting new images that were not used for training the model and it was detected that it can predict future traffic conditions. It was also noticed that Neural Networks models require great amounts of data, and that if provided with more data, the model can improve the performance.

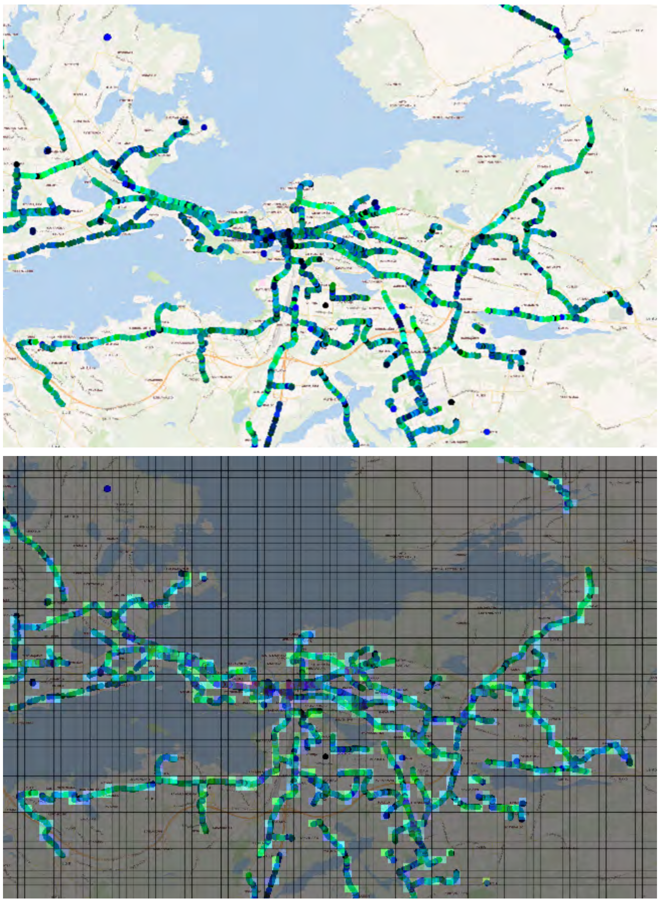

Figure 17 > *Data preparation phase: (Top) Tampere map with GPS coordinates accumulated for 5 minutes (bright colors indicate high speeds). (Bottom) the map is divided in a regular grid, and the average speed is calculated for each square*

The data was collected from Vehicle Activity, provided within the open public transport data[14] in regular intervals of 1 minute, and this data was collected for three consecutive days, resulting with 2,340,070 lines of activity with values for date, time, line, vehicle, direction, latitude, longitude, speed. These activities were grouped in 5 minute intervals and one image in which the colour of the points represents the speed registered for that specific point was produced. One image is generated for each 5 minutes as can be seen on the top of Figure 17.

Instead of considering the exact location of the GPS coordinate (latitude and longitude), a regular grid was defined, which was coloured each with the total speed measure for its internal points. One example is presented on the bottom side of Figure 17.

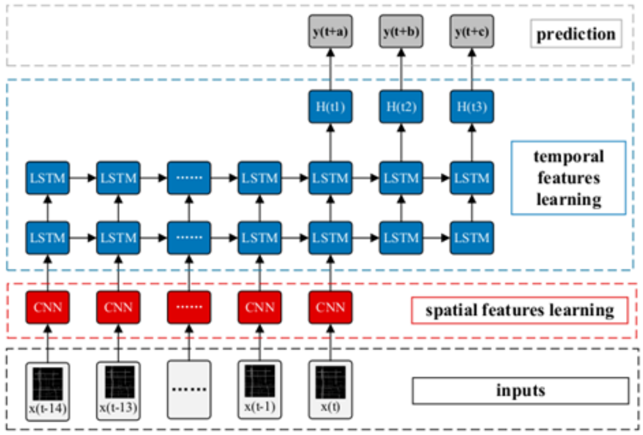

Figure 18 > *Sequential grids are used as inputs in a Spatiotemporal model capable of predicting future traffic conditions, figure and method from Yu et al. (2017)*[19]

A model based on Convolution Neural Networks (CNN) was then implemented. The methods are better described in chapter 4.4. Normally it is used for the computer vision domain for classifying images; and Long Short-Term Memory (LSTM) usually applied in temporal prediction tasks. The model is illustrated in Figure 18 together with its bibliographic reference[19]. The CNN is responsible for building a compact and meaningful representation of the image presented as the input, and the LSTM is responsible for learning temporal patterns in these representations. When trained together they can detect spatial and temporal patterns in buses' activities.

18 The source code of the prototype is available from Rodrigo Borges' GitHub repository

19 Yu, H., Wu, Z., Wang, S., Wang, Y., & Ma, X. (2017). Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks. Sensors, 17(7), 1501

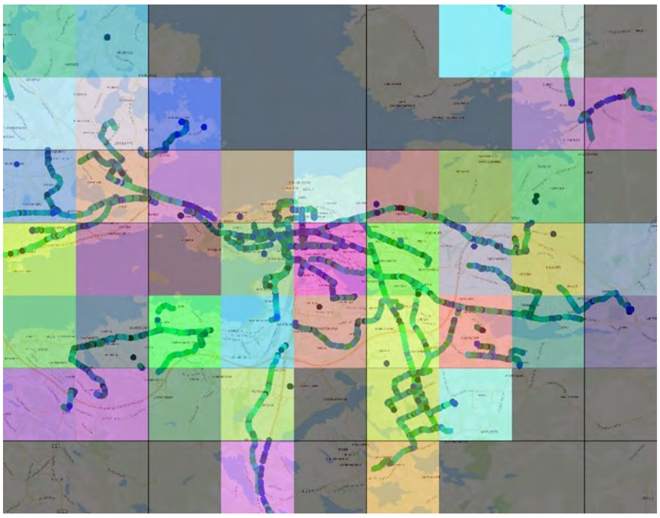

Figure 19 > *The average speed is predicted for each region on the map for the next 5 minutes*

The model was trained with the data collected previously and instead of presenting the same grid structure as the target for predictions, we decide for wider squares, as the one in Figure 19. The idea is that a higher resolution (smaller grid) is important for detecting nuances in the spatial patterns, but lower resolution (bigger grid) is more useful as target for predictions. In the end, the model is trained with consecutive images of speed and location of buses information, and after optimizations it is capable of predicting the future average speed in each area in the city.

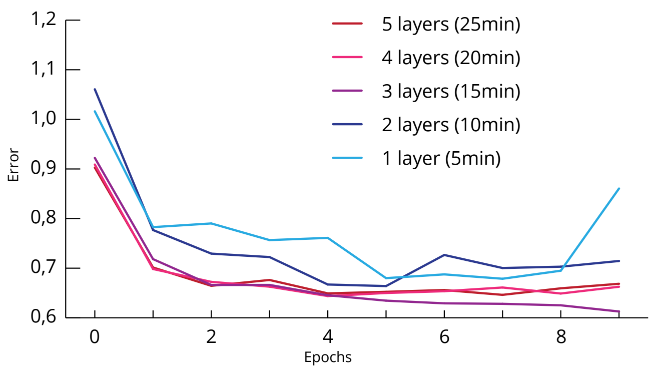

Figure 20 > *The prediction error when considering time intervals of 5, 10, 15, 20 and 25 minutes*

To test the method, 80% of the activities were used to train the model to predict the remaining 20% of values. After several rounds of experiments with measured error rate for different time intervals (5, 10, 15, 20 and 25 minutes) of activity combined in the images. These errors measured are presented in Figure 20. As one can see, some intervals provide more accurate prediction than others, and the lowest error rates were measured for the time interval of 15 minutes. The prediction results are expected to improve when more data is applied for training the model.

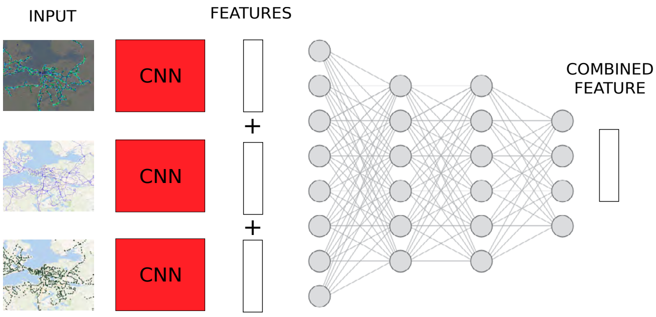

Figure 21 > *Combining several inputs into one single representation for increasing accuracy*

One possible extension of the method includes combining several features in the input in order to increase the accuracy of the prediction task. As an example, the same situation described previously can be imagined, but in addition to that also having an access to air pollution sensors and weather forecast. If collected in the same format as the bus speed data (latitude, longitude), and in the same time interval, they can be combined together in one feature, as in Figure 21. There is no limitation regarding the number of new inputs that can be combined, the only restriction refers to the fact that they must preserve the same format as the first one describing bus speeds. The new inputs are submitted to a similar CNN and the resulting feature is learned by a new MLP network that will adjust its weights depending on the relevance of each input in the current prediction task.

4. Concept co-development

In this section, the initial concept development of Smart TrafficHub is discussed. First, when specifying the concept of the Smart TrafficHub, it is important to note that the lessons learnt, e.g., already from the MyWay project are directly applicable: Implementing an intelligent transport system requires an active, well-specified ecosystem of collaborating stakeholders, and when actually developing services, a technical roadmap and an architecture with well-documented interfaces are eventually needed (c.f. Figure 5, MyWay conceptual high-level architecture).

The development of the Smart TrafficHub project should thus be considered as a coordinated effort of adding a specific layer of knowledge from the Pirkanmaa region to the bigger ITS picture, providing insight on the particular needs and particular stakeholders w.r.t. to smart traffic stops and hubs in this region. Having said this, the underlying "hard" requirements remain similar from one region to another: Organizing a critical ecosystem of capable stakeholders, aligning business objectives, agreeing which concrete smart services are provided to best serve customers and the business needs, establishing the necessary technical infrastructure with open interfaces and technical instructions, and providing operational base systems and applications for delivering, maintaining and continuously improving services. Therefore, the project consisted on workshops to gain stakeholder involvement from the design stage. The stakeholder events are described on the next chapter, followed by the summarized results of these events.

4.1. Stakeholder events

The project hosted two workshops and two steering group meetings during the study phase to gather the demand of intelligent transport stop points and the important aspects that should be further studied. The focus points of the two workshops are described in Appendix A: Workshops. This section describes shortly the discussions, the actual outcomes of all conversations are presented in chapter 4.2.

Main points of the conversation in 1st workshop were related continuous guidance around the city, the possibilities to measure or estimate the capacity of buses and trams and whether stops could also serve guidance in emergency needs. The stop operator's viewpoint was also discussed whether some services could also create information of use of stops to support operator's business cases. Also, entirely new set of concepts were also discussed, such as food or package delivery systems in stops or e-bike charging and transfer points.

The first steering group meeting was held after the workshop and they were presented with the summary and findings of the workshop for discussion. On the discussion, it was noted that there have been pilots regarding the capacity information on buses, but nothing continuous. However, trams will be equipped with system that can provide the information. Based on this, it was discussed whether available data, such as smart card data could be used to build an AI-model to estimate the capacity where no direct measurements are a possibility. Regarding stops, it was also discussed whether sensors in co-operation of AI-model could be used to track the status of different stops to allow predictive maintenance.

Main points of the discussion in the 2nd workshop hovered around several aspects, such as the responsibility of the stakeholder to create an environment to establish long term services to businesses involved, the different types of stops and how the services could differ, methods of providing information at stops to make it easy to use for everyone and how data is handled and error corrected, when needed. The second steering group was held after the 2nd workshop and on the discussion it was noted that there are several other projects and also new interests that would be interested of the results and background of the project, thus creating a strong need to not only report the results but also the background studies and lessons learned.

I addition to workshops and steering group meetings, within the project there has also been discussion with other projects within the area that focus partly on same aspects. Discussions between SmartRail-project[20] covered aspects on how different vulnerable to be excluded user groups should be notified as part of the guidance. The same aspects of easiness to use and understand the systems for users have also been discussed with KITE-project[21]. Finally, also different aspects of how both event and emergency guidance could be included in Smart TrafficHub concept were also discussed with SURE-project[22].

20 SmartRail - Development and piloting of accessible public transportation services for user with special needs.

21 KITE - Human-Centered AI Solutions for the Smart City

22 SURE - Smart Urban Security and Event Resilience

To support the dissemination of the project results, a 3rd workshop of the project aimed to briefly present the results and allow conversation of new ideas and further project needs that could be discussed after the project, including the ones listed at the end of this report on chapter 6.

4.2. Initial development

The initial concept is developed based on the conversations at both workshops, steering groups of the project, the survey results and conversations with different projects and stakeholders having same types of themes. This chapter summarizes the different viewpoints of the conversation.

Table 1 > *Summary of results from stakeholder events*

Area	First steps	Next steps
Organization	Long term contracts and data to provide value for stakeholders	Data is share among the stakeholders to allow new innovation
Guidance	Passenger guidance available continuously to all transport modes	Personalised guidance for user needs and habits
Data use	Open sources of data from different sensors and data providers	AI-based predictions and error correction based on new and already gathered date
Public transport	Real-time information for users to better use the network	Network can adapt to the situation based on number of passengers and different events
Emergency use	Emergency use is noted and possible to some extent	Frameworks and automation to allow fast reactions to places where needed

Smart TrafficHub organization

One of the key enablers of the Smart TrafficHubs is their organization and how different agreements are made. The contracts to provide the services should have longer timeframes to allow different companies to create new solutions and develop them with long term on the stops. As data for different solutions can be both open or produced at the stop, long term contracts are also important to allow other companies to develop solutions based on data provided by other companies operating at the stop, thus allowing new services to be created.

Funding of the stop constructions itself is an important aspect as well. As the stops are majorly funded by placed advertising, in general, these new solutions and services should not disallow the stop to work as an ad. However, different services hosted at stop could also help the companies running the stops if it allows some data gathering or profiling of the users to allow better to target the displayed ads. However, for this aspect GDPR and other rules relating personal data use should be inspected and user consent would be needed.

Passenger guidance

Large stops currently serve as a hubs for different modes and lines of public transport. Since mobility consists also from other modes transport, it would be important that all the guidance in stops should include these other forms, such as e-scooters, city bikes, taxis and other forms for true MaaS-support and the guidance would work across these modes. Since different screens and systems show guidance around the city, it should be noted that the guiding in stops should follow the same terms and visual identity as the other system to make it easier to understand.

One aspect of the guidance systems are also different large events, such as concerts or sport matches that can draw many tourists in the city. For different guidance screens around the city, these events should be made visible and in addition to regular timetables and stop name information, it should be clearly marked, where the event is happening and how to get there. For example people arriving to train station can have on-time information of which line to take to arrive to that location already at the station and the same type guidance throughout the trip could be given when leaving the event as well. This could be also used to vary the routes where people enter and exit the venue to better control the people flow.

In addition to guidance around the city, different user related aspects should be noted as well and the guidance should note if for example the user needs ramps or seating to complete the journey and guide them accordingly to both correct lines and directions. This is crucial for making the different trip chains to work for everyone and both the information and the stop spaces itself should support the easiness of line and modal changes. One key aspect is to also inform all the users regarding the interruptions within the network and how they affect the routing and connections.

Data collection and use

There are already many open data sources and existing frameworks that support the development of the different services within the stop environment. However, there are still different new aspects of data collection and use that would help generating new services.

Stops are important parts of the journeys and their looks and feels affect directly on people waiting. Different types of sensor could provide information of cleanliness and slipperiness of the stops to allow maintenance at times when it is needed. In addition to sensors, the different passenger involved apps and short queries could also allow the actual users to report if maintenance is needed. Based on the sensor data and passenger reports, weather data and forecasts could be used to create an AI solution to estimate on what conditions for example anti-slip measures are needed to allow predictive maintenance.

One main interest related to data needs is also the ability to inform and predict the capacity status of incoming bus or tram. Currently, there is no solution gather this type information but some pilots have tried different data gathering methods. The information system of public transport would already allow this type of information shared and the trams will include sensors to provide the capacity status in the future. However, as the buses in the public transport system do not have any technology to measure this, different solutions to allow their capacity tracking are seen interesting as it would allow to track and inform user on capacity status of the whole network. One possible solution to this in the future would be to use smart card data to create an estimation model for given timeframe and estimate the crowdedness based on previous data. This could be later enriched with data from the stops, if some services gather for example beacon based data of waiting times and number of people.

On more used and crowded stop points the buses don't necessarily have the space to stop at the specific location if other bus occupies the location directly front of the stop. This may be challenging of the users to know where the bus can be boarded and they may wait at the wrong locations. The already available GPS data form buses can be used to track down specific points where the buses tend to stop on a given area. This data in combination of GPS track of other lines using the same stops would allow to better track and estimate the conditions of when the stops is usually occupied by multiple buses and the locations differ to note this in the guidance. Real-time data and AI-models can also be used to estimate, whether the next bus arriving the location can stop at the correct place and inform the passenger in advance on where to board.

Data can also be used as measuring and informing user about modal choices and their effect. For example, stops containing park and ride -services could share information about the amount of parked cars not only to inform its capacity but also to show the emissions and modal share effects of these choices. Same modal choice emissions effect can be also informed as part of guidance to indicate the effects of different transport methods.

The data gathered would not only serve different users of the stops by allowing use of better real-time information, but it could also allow different dashboards for example to stop point maintenance and bus operators to allow them better track the status and plan the needed changes accordingly. The same data could also be later implemented as a part of digital twin and other solutions that provide real-time information to support development and planning in general.

For current public transport services, all the data could be used to better develop and track the use of the services as well as to inform the passengers for a new alternative routing based on the current status of network. Fur future network having different autonomous feeder systems, the information could also be used to allow real-time, AI-based decisions to be made based on the current capacity and passenger information. For example, on days with major evets, if a tram line towards Hervanta has a larger than usual number of passengers, more autonomous vehicles would be deployed to carry all those passengers to their destinations.

Stops as part of network

Stops are a key part of public transport. The main aspect to note is the difference of these stops. They serve different types of passengers, are some large stops are natural hubs (such as stops at city centre, major shopping malls or railway stations), some stops are mostly used as only exit nodes of the network. There is also a difference on how the stops are operated, as some of them are connected to electricity and network, whereas others may only have a electricity connection to have a light, but nothing else. This affects on what can be provided in these.

As already noted, the guidance within stops should follow the same guidelines in all different guiding systems around the city. The same key concerns affect the stop displays as well as any other displays; on which extent the data can be show using icons and numbers that are not dependent of the language, what languages are used when more information is show and how the change of language can be made easy, even when the display at that point shows a language that is not familiar for the user.

Also as not every stop should show the same information, but information that is important to the connections regarding that stop and services around that area, it should also be noted that different user value different information. One of the key questions therefore is also how the information can be customised and those different users can be identified. Same type of information could also served for user on different screens, whether they are at stops, at buses, or at mobile. Many of users already use their mobile to see this type of information and it also allows the easier personalisation of the information, but the same information would also be provided on displays on stops for users without a suitable phone, dead battery or too cold weather to use a phone. The same information from screen to mobile is also one key aspect to allow continuous guidance during the whole trip chain, which is seen important on all users, but especially of the groups having special needs, like difficulties of hearing or seeing.

It should be noted that people with limited eyesight typically struggle in 1) finding a proper stop, 2) stopping the correct bus, and 3) getting out in a proper stop. The first and third issues could be aided with virtual information, but for the second, in addition to virtual information, also existing physical bus markings should also be improved. Different guidelines[23] dictate the targets of final design.

Stops could also serve as a location for entirely new developed services for the users, such as food or package order collection points, to allow charging of e-bikes, e-scooters or shuttle buses. The new network would also allow new environment to run pilots that benefit from central data collection points and network having multiple mobility modes.

Emergency aspects should be noted when developing these new systems. In case of security cameras in the stops, can AI be used to track and alert personnel if slipperiness or dangerous items are detected. Also, as the stops host different screens to provide information and ads, these screens should be included when emergency guiding, or notices need to be given to the public. The screens and guidance can also help avoid emergency situations, as traffic can be guided beforehand to avoid areas to get too congested.

4.3. The concept

The main concept of Smart TrafficHub can be summarized and described in information and interaction at different parts of the journey; Planning, starting and during the trip and approaching destination. One key part of the concept is also the information that can be provided to other stakeholders. These parts and their interactions with the data needs are presented below at those different stages.

Planning the journey

A passenger wants to move from one location to another using public transport as main transport mode. At first, the person should be given information of the different routes and modes that could be taken towards the destination. This information should be accessible on different devices, so the planning would take place for example on-the-go with a mobile phone, well in advance at home with computer or at a capable stop having interface to use the planner. The general information regarding the trip is already largely available, also noting different accessibility challenges. However, a comprehensive system would also need to include all 3rd party provided mobility systems as well, thus allowing the passenger to book a e-scooter or taxi for a last-mile solution in case that is the most suitable option.

In addition. the guidance should include the estimated capacity forecast of the lines and segments the passenger is suggested to take. This could be also noted in the routing systems to guide people on less congested lines when possible. With actual capacity information and estimation, the information of both the availability and occupancy of special areas of the vehicles should be noted as well, such as available wheelchair, bicycle, or trolley spaces. Another added information would be event related guidance, allowing an event visitor to just select the event one's attending to, to see the same rich information and alternatives on how to go there, without any need to know the lines or stops used through the journey.

When the passenger is planning the route using the system on the stop itself, the passenger can also be asked some quick, optional questions regarding the status of the stop itself, such as the slipperiness of the stop floor or general cleanliness of the stop, which can later be used to develop when and where maintenance should be focused.

23 Accessible Bus Stop Design Guidance by TfL and Guide to design accessible public transport by Finnish Federation of Visually Impaired

Starting the journey

When the passenger has information on what stop and route to take, the system can inform the passenger on the capacity of the arriving vehicle and in case all-door boarding is possible, the information on what would be the best possible door to enter capacity-wise. When multiple buses arrive the stop at the same time, the system can also inform the passenger if the correct bus is arriving behind the first arriving bus, thus informing the passenger that the bus should be entered 15 meters to the guided direction. Also, when the system informs the passengers of arriving lines, the information can also show, it the arriving bus should be taken when going to a large event, thus allowing the event visitor to board the bus only knowing the event one's attending.

During the journey

In case the journey requires a transfer, the passenger is informed on what stop the transfer should be made and in case the connecting bus leaves from a different platform, the system indicates the walking directions to that new stop. This information is given already when the route is planned and in case a mobile system is used, the same information also follows the passenger during the trip. For larger events and main destinations, the same transfer information is also given in the vehicle, when it approaches the stop. The transfer information also notes the capacity of the upcoming vehicles as well as the accessibility challenges the passenger may have.

Approaching destination

When the journey nears the destination stop, in the vehicle is shown, how the large event or an important destination can be reached from that stop. The same information can also be presented on mobile, enriched with information of last-mile solutions provided at the destination stop and how they can be used to reach the destination. The system should also include information on how to use them and with a payment solution that can be combined in the public transport ticket from the viewpoint of the passenger.

Other stakeholders

During the operation, the system itself can collect and analyse lot of information that can be used to help multiple stakeholders and should be provided via a personalised dashboards.

The information collected from passengers on the stops can be enriched with different sensor data and AI-models to estimate and track the development of different metrics. In addition to helping stop operators to focus maintenance where needed, it could also allow estimated proactive maintenance, when the system, based on different sensors and forecasts, detects patterns that lead to slippery stop floor.

When capacity is collected and tracked, that information can be used to better develop the network, as it would allow better origin-destination tracking and therefore help decision making from the public transport designer viewpoint.

At the same time, if the system knows how long people wait generally at different stops and where they travel, the information can also be used by the stop operator to focus the advertising campaigns, thus creating a better revenue from the system. It should however be noted that while better tracking and identification of passengers help to target the ads, the system should be build privacy-first and it should be thoroughly valuated on what information can be used for this purposes and what kind of consent to use would be required.

4.4. Models and technologies

In this section, the perspective of the system designer is considered, and potential data formats are described depending on the methods applied in the task. This includes three categories of Deep Learning techniques: **Multi-Layer Perceptron (MLP)**, **Convolutional Neural Networks (CNN)** and **Recurrent Neural Networks (RNN)**.

Multi-Layer Perceptron (MLP)

These were the first neural networks proposed in the literature and have the simplest architecture possible[24]. MLPs consist of input, hidden and outputs layers, each one of them constituted of units called perceptrons (also referred to as neurons) as showed in Figure 22. The number of hidden layers as well as the amount of perceptrons per layer can vary depending on the necessity and from layer to layer, perceptrons are all connected to each other. Non-linear activation functions are usually applied between consecutive layers given them the power to approximate complex and non-linear functions if designed correctly. The training process is usually conducted with Gradient Descent when each perceptron has its weights adjusted according to the optimization cost function. These networks are especially useful in situations when the data available is composed of real numbers

24 Multilayer Perceptron at DeepAI

and the task is to predict a categorical variable. As an example, we can imagine the task of predicting the average number of travels a passenger makes per week (low, medium, high) when having demographic characteristics (age, gender, nationality, etc.) in the input.

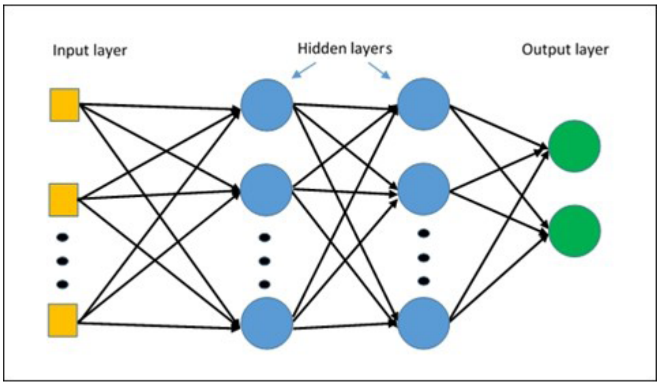

Figure 22 **>** *Multi-Layer Perceptron (MLP), figure from Tutorialspoint*[25]

Convolutional Neural Networks (CNN)

CNNs are very popular artificial neural networks architectures first proposed for classifying images in categories, as in the example of Figure 23, in which the input picture is classified as containing cat or not. These can be applied not only to images but to any data that can be represented as such, and within which there might exist spatial patterns that correlate with the task in hand. Yet in the previous example, we can think of cat nose or a tail as spatial patterns that be detected and used during the identification process. These networks share some features with the MLPs, layers can be stacked as much as necessary and are interleaved with non-linear activation function, but this time internal layers are responsible for learning convolving kernels that slide along the image and pooling layers are applied for reducing their dimensionality gradually. The training process is conducted the same way as before. CNNs can be particularly interesting, as mentioned before, when having 2-dimensional data in the input (as images are) and with potential patterns that can be detected. In this specific case, lets imagine an application where we have maps of the city with any kind of events drawn on the top of it according to geo coordinates, and the task is to classify map images according to the level of occurrence (low, medium, high) of traffic jams.

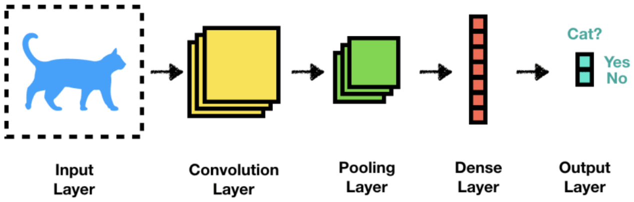

Figure 23 **>** *Convolutional Neural Networks (CNN) classification architecture, figure from towards data science*[26]

Recurrent Neural Networks (RNN)

And finally, RNNs are networks designed for estimating future events of a sequence, given its previous values. The main difference here, if compared to standard neural networks, is the RNN's ability to store the current state of each cell to be used in the following ones for detecting temporal patterns. In Figure 24 we see the cell states indicated as h, inputs are indicated as x and outputs as y. For each new value in the input the RNN can inform the potential next element, as well as its current state. The values of hidden states can propagate through the next elements, as much as necessary, depending on the phenomena under analysis. These architectures are especially useful in situations when temporal patterns are observed, for example for estimating the number of passengers inside one vehicle given the occupation of this same vehicle during the last 30 minutes.

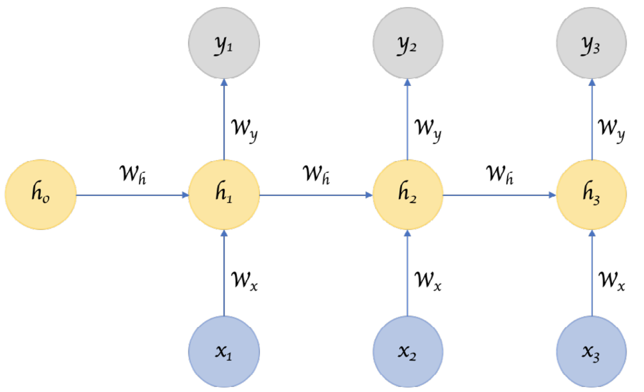

Figure 24 **>** *A Recurrent Neural Networks (RNN) typical architecture, figure from towards data science*[27]

25 TensorFlow - Multi-Layer Perceptron Learning at Tutorialspoint
26 Convolutional Neural Network: A Step By Step Guide, by Shashikant at towards data science.
27 Recurrent Neural Networks, by Mahendran Venkatachalam at towards data science

5. Next steps to pilots

Based on the developed concept and prior research on the subject, three different plans were created to describe how **capacity information**, **routing & guidance** and **stakeholder dashboards** could be piloted and what aspects should be covered. The results are divided between these three different cases. However, there are some internal links, where outputs from other phases could be used in other pilots as well. These are described within each topic.

5.1. Capacity information

How current capacity information could be tracked and how the capacity could be estimated for future journeys?

Stakeholders

Main stakeholders that should be included consists of public transport office and the provider(s) of information systems that collect and share the data needed to collect capacity information.

In case of camera and beacon based collection of data, stop operators and public transport operators would be important stakeholders as well as the sensor and system manufacturers.

For better estimation of the capacity, the stakeholders could also include different mobility related providers, such as long-haul buses and trains, parking facilities and other stakeholders that could indicate amounts of visitors that could use public transport.

Available data

Currently, no actual capacity data exist from the network in Tampere. The detailed smart card check-ins are collected and hosted centrally, however due to nature of the data, it is not open. When tram starts operating, it will have weight-sensors and vision based systems to estimate the current capacity and availability of free space for wheelchairs and trolleys.

New data needs

Main new data for building a capacity tracking and estimation model would be the actual smart card check-in data that would allow the pattern tracking. The data itself can be anonymized, but some unique identifiers for users are needed to allow tracking within different parts of the journey. In addition to the check-ins, static information about the fixed capacity of different bus types should be available as well as linked to the real-time data, so it is possible to know what type of vehicles are used in any of the routes.

In addition to data-based tracking, beacon collected information and camera feeds could also serve as an input for the calculation. This type of capacity prediction has also been done based on crowdsourced information[28].

Methods

Smart card data would allow an AI-model to estimate the possible destinations for a data-based know origin (validated check-in), based on the history of that unique passenger and one's other boarding locations and their timestamps. There are some studies available that discuss the possibility of this and state that for at least peak time travel, the routes are trackable[29]. This type of per person estimation would be possible with different multi-layer perceptron and recurrent neural network systems. When capacity information would be acquired, recurrent neural network systems could be used for time series prediction to estimate the capacity for future journeys based on historical evidence. The actual model and its estimation can be enriched with the measured capacity information (for example provided by the tram) and some other data layers that could include patterns to know if they are differences in amount of visitors (such as event visitor data) or any other location-based data source that could give indication on mobility needs. The system would also develop itself, as more and more data is gained from smart card check-ins.

In addition of the smart card data, camera feeds from buses and public transport stops could also be used to estimate the crowdedness based on computer vision and convolutional neural networks. A system by Counterest[30] has already been tested in a few buses in Tampere during 2018-2019 regarding this topic.

28 Crowdsourced Transit predictions at Transit Partners Help, Google Support

29 Trépanier, M., Tranchant, N., & Chapleau, R. (2007). Individual trip destination estimation in a transit smart card automated fare collection system. Journal of Intelligent Transportation Systems, 11(1), 1-14; Jánošíková, Ľ., Slavík, J., & Kohání, M. (2014). Estimation of a route choice model for urban public transport using smart card data. Transportation planning and technology, 37(7), 638-648.

30 Counterest - Automatic people counter and customer analytics

Beacons would also allow track to estimate where a passenger boards, on which vehicle and where the passenger departs. Beacon tracking system is in use in HSL-area[31]. This type of information could also be used to enrichen the capacity model, as it would provide information on amounts of passengers waiting on next stops.

Outputs

Camera and beacon-based solutions, as well as the tram sensor system provides real-time data form capacity as it currently is, with the location and timestamp. Smart card data could provide estimations for both present state and capacity state of the future journeys. In addition to only giving this type real-time added information, the capacity could also serve as an input to routing to allow different routing algorithms to consider whether the suggested routes are congested or not?

Limitations

The smart card data is not currently available and with the current state of the systems, it could not provide real-time information, as the data collection method adds one day delay for the data. However, in case the information system is further develop to allow real-time access on check-in data, the continuous information would allow some information of departing passengers as well. For example, when a passenger is validated for the transfer line, it is known that this passenger must have left the previous bus.

Also, if camera based systems are used, the ethical aspects should be noted, such as how the video feed is used, analysed, and stored in a way that the privacy of passengers are secured.

Open questions

There are some open questions related to the capacity calculation. Mainly, even though the developed results will be highly automated, both the data extraction to support this and the model creation will need resources. On top of these, resources needed to run the continuous analysis are needed as well. One main concern to be developed as well is how this information can be connected to other systems. Since tram will already give capacity information, a system would exist for this, however, a supporting system for communicating the estimations between different platforms would be required.

This all creates a question on who the actual stakeholder would be providing these estimations and how this operation could be funded as it is not clear, whether the capacity information and estimation could be provided as a sellable product or is there an actual demand for it.

Next steps

The first phase regarding this would be consistent capacity information based on both the metrics and the deep learning model given out with other public transport real-time information. When the forecast part of the model is further validated, this data can be used as a part of routing and guidance to make it possible to give directions that will avoid congested areas when it is possible by the network design. When further implemented, the capacity information could also be tied to autonomous feeder systems in public transport to allow demand-based offerings. This concept is better defined in chapter 6.

5.2. Routing & guidance

How a comprehensive routing & guidance platform could be achieved?

Stakeholders

For comprehensive routing & guidance, all different mobility service providers would be important to note when building the system. There are also some 3rd party service providers that are currently providing integrated search applications, information dashboards and mobile services based on the primary stakeholders' data, which should be included. To allow better guidance to visitors, different event organizations and event locations should be include as well to allow data shared of when events are taking place and how they should be accessed to include these in the system.

One major stakeholder to include as well would be the one running the guidance system, whether it is one of the providers or a completely 3rd party. In case the guidance is given on screens on stops and around the city, the stop point and screen operators would be natural stakeholders to include as well, depending on how broadly the same guidance systems would be used. Optionally, stakeholders could also include different logistic service providers that could be combining the mobility needs of both passengers and goods.

31 HSL Beacon API at Helsinki Region Infoshare

Available data

Lot of the data is already available. Different datasets allowing routing for different travel modes exist already, as well as essential information of all different transport service providers, since the are obliged to submit it in the national service catalog[32]. Continuous bus location data is also already available[14]. Different events also already display information on how to arrive and what would be the best connections, but this data may not be in a format that is easily accessible.

New data needs

As regarding transport service providers having to submit essential information, the information should be widely available. The system, however, does not necessarily have every new mobility provider or any API-access to those. Given that the system would want to give comprehensive directions to the passengers, it would also be crucial that all the stakeholders have open payment APIs to not only support the routing, but also allowing the passenger to book and pay for the complete journey, in case the passenger would want to do this.

In addition to access on different service providers, the capacity information and estimation of public transport would also be part of the comprehensive guidance system.

Methods

Given that the routes and timetables of all different service providers are accessible, the route calculation is fairly simple and many different solutions already support the combination of multiple services in the system.

For real-time information, if the system is wanted to show the predicted bus stopping location when it arrives to the stop, the previous collected bus stop location data can be analysed in relation to time and different other lines at the same stop environment. Convolutional and recurrent neural networks could be used to analyse these types of time series prediction. This would allow that given the current state (such as which buses are arriving and in what order) the system could estimate where each of the coming buses would stop. This information could be enriched with real time information of trams capacity sensors to display on which door would be best to board in case all-door boarding is a possibility.

Outputs

A simple results by the system would be requested route for the passenger, noting the choices that was made (like less walking) and the estimated capacity of the vehicles the passenger would take. In case all stakeholders are involved in the system, the same system could also a simple ticket purchase for the passenger.

One output for the stop environment would be the location where the bus would be expected to stop and where it would be recommended to move in the coming vehicle based on capacity status. Large events would also be noted throughout the system, thus allowing the passenger not to know what is correct stop or line to take to an event. Instead, the system should indicate when the bus is arriving, that the upcoming bus will go past a certain event and on the bus information screen should then be displayed that the upcoming stop would be the one to take for that certain event.

The system can also give the CO_2-effect of the routing choices compared to other choices that could have been used for that trip. In cases of the stop being in connection to a Park&Ride-area, also information about the amount of parked cars with the effect on CO_2-emissions could be given.

Limitations

As main limitation, the comprehensive system would need the information of all mobility service providers, but in case of new or changed supply, it will be challenging to have the system always up-to-date. In case payments are included, it would add a another layer, as also funding APIs of all different stakeholders should be available, useable and always up-to-date in the system.

Open questions

How a comprehensive routing could be funded as for routing only, something extra would be needed to allow also income from the product. In case also payments would be included in the system, it would give the operating stakeholder a possibility to charge a share per every trip sold on the platform. However, as the stakeholder is then providing and selling the trip, it would add complexity for the provider as it would be the ticket issuer. Therefore, that operator would also be the one to ensure that the system can be use as sold.

32 NAP - Finnish national open transport service catalog

Next steps

Personalised guidance is an important part of the experiment. If on first phase the guidance can only support a few different criteria, such as need for wheelchair access or poor eyesight, and change the routing and presentation accordingly, on later stages the system could indicate the actual user (either via mobile or smart card check-in) and tailor all the information and routing specifically to that person. This would include for example including the city bikes more heavily as a part of the route, if the person has been using the city bikes before. The guidance system can also adopt to external inputs, such as the capacity readings and estimates of different vehicles, when the model is proven to be reliable.

5.3. Stakeholder dashboards

New systems provide new information to passengers, but how other stakeholders could benefit? Since public transport data is already widely collected and visualised, this chapter focuses on stop point monitoring. It should still be noted that the capacity information presented could also be an important addition for public transport operators and planning.

Stakeholders

When dashboard-like views of the data are discussed, the most important stakeholders would be the users; stop operators and maintenance partners. In addition to these, one key role lies also in distribution of the data from different systems to the dashboards, as well as manufacturing different sensors, if they are used as a tool to gather data on the spot.

Available data

For maintenance use cases, current and past weather, as well as forecasts can be acquired as open data. There are also various more local measurement points around city, that could allow more local tracking of weather changes that would apply to a certain stop. In addition to this, some heavy vehicles can collect and share slipperiness data from road network, which could be applied to detect patterns that lead to slipperiness.

New data needs

For the actual circumstances at the stop, a weather sensor could be applied directly at the stop to collect the actual data of the correct position, including for example temperature, wind, humidity, air pollution and rain or snow accumulation. To detect the amount of people camera or beacon-based sensors could be used to track the crowdedness and time waited on the stops. Regarding crowdsourcing of the data, the routing and guidance system could also ask users at some interactions whether the stop is clean or whether maintenance or cleaning is needed, and also give an easy method to give feedback about the state of the stop. Similar methods have been used for example at Helsinki airport to track the cleanliness of toilets[33]. Same kind of feedback can also be collected from trash cans that can automatically indicate, whether there is need for emptying the bin[34].

Methods

When the local weather and environment variables are tracked, a time series prediction model can be build using convolutional and recurrent neural networks to estimate, what kinds of weather patterns lead to slippery floors or other non-desirable conditions at the stops. Local weather forecasts could the be used to estimate when those events are bound to occur, so a predictive plan in maintenance can be done to avoid the worst conditions. The crowdsourced data from the stops and open data from the vehicles can be used to better train and develop the model.

Outputs

Main output would be information of the status of different stops and their current maintenance needs. The system would also display information when and where new maintenance needs would occur. If the stops are equipped with cameras or beacons, the information can also include how long times and how much people are staying at the stop areas, which could be valuable information for the advertisement providers.

Limitations

Beacon-based systems could detect how long people are staying on the stop and when their leave and with what line, which would arise privacy concerns regarding use of the data. This is further discussed at chapter 6. Camera-based systems could also estimate the crowdedness, but allow only estimation of the amount of people, but it does not include information whether the people is the same within the shots.

33 Epäsiisti vessa? Lentoasemalla sotkuista voi ilmoittaa saman tien siivoojalle uuden teknologian avulla at Finavia
34 Lehtovuori FinbinCare allows direct notifications at Lehtovuori CityPack

Open questions

Depending of the consent of the data usage, the data collected could allow a lot of methods for advertisers, as the passengers can be exploited for identifying the ideal ads on the bus stops. The further the system can track the users, the more targeted ads could be given. This would also allow this information to be used for producing an investment plan for the bus stops and their nearby area. However, all this is directly linked to data privacy and how much of collected is shared between the stakeholders, which opens a series of more questions as discussed in chapter 6.

6. Conclusion and future topics

There are a couple clear results from the stakeholder discussion and development, as well as a couple of new larger targets where future research and pilots could be aimed to both support the development of the concept and Tampere region as one of the major testbeds globally.

First, capacity information and estimation is seen as interesting information from multiple different viewpoints, thus making it a clear next phase that should be further researched and tested in the environment, as it would also give inputs to our other proposed pilots. Secondly, when deploying these systems, a common, semantic and non-monopolized infrastructure is needed for the stakeholders.

Common, semantic infrastructure

When actually implementing and deploying smart transportation systems, a critical and concrete challenge lies in establishing a common information infrastructure for the ecosystem. This requires setting up a technical infrastructure, agreeing policies and APIs, and describing services and data semantics in machine-understandable way.

This observation quite naturally provides necessary steps also in the smart traffic stop area as well. Indeed, if such common information infrastructure with is not established, the common pitfall is either that ecosystem actors prefer working (only) in the data consumer role, or that data providers evolve into silos without commonly agreed interfaces – making the life of application developers both unnecessary laborious and uncertain.

A concrete model for tackling this problem, is to organize smart transportations systems into peer service networks which using common semantic interfaces, both not only provide access to their own data, also replicate and share data provided by others (intuitively in a fashion similar to common DNS services). This way developing services is both easier and less risky, and thanks to the replication-driven redundancy, a service break from some individual data provider does not cause any breaks for the (end-user) applications. At semantic interoperability level, such activities fall under the broad umbrella of (open) linked data[35].

Hervanta smart traffic test area to allow pilots and development

Already now, all infrastructure (roads, bus stops, traffic lights) will have many different integrated sensors that will collect a lot of data on the local conditions. Furthermore, smart vehicles, smart cars, smart and autonomous cars and maybe even smart trams will generate a lot of data during the operation. As well as the users generating a lot of data from cell phones and other the devices they have with them at all times.

One of the key questions in the future is that how this data from different systems can be used to build a traffic system that is tailored to individual needs, is more efficient, more sustainable and provides better services for the end users. At the same time, taking care of the data security, privacy and ethical aspects.

Tampere is building a smart city and smart traffic test area that will be located in Hervanta suburb. The concept will provide opportunities to test and use latest technologies, like for example, 5G and different C-ITS systems to develop new start traffic concepts. The test area will be open for universities, research institutes and companies. Thus, one interesting possibility would be to develop some of the key findings and ideas generated in this Smart TrafficHub project further, and implement and test the ideas in the smart city test area in the area in the future. This would make it possible to test how the smart bus stop would work as part of the smart traffic system.

Ethical aspects of data collection and use

Already, a lot of data regarding the mobility patterns of the passengers are stored in multiple different systems, which could allow a bit too thorough tracking of one's personal daily life. In this case, even though the data exist, it does not necessarily mean it is used for said personal profiling, which would be the case with the smart card data. However, is this data would be opened to a 3rd party to allow capacity calculation, what should be noted to allow the best possible prediction but without giving away too private patterns? Then, if this data would be combined also with camera and beacon-based tracking, how a passenger can be sure that his travel choices is not known by everyone and what would be the main privacy aspects when designing a system like this? If the capacity calculation is funded by selling some general data for advertisers, will it affect on what data could be collected?

35 See, e.g., W3C Data Activity, SmartOpenData and Share-PSI 2.0

Easier access to information for passengers: chatbot-based user interface

The recent evolution of Natural Language Processing (NLP) techniques gave birth to Chat Bots, which are artificial intelligences that can interact with users via a text-based chat in a question-and-answer fashion. To achieve the naturalness of a human conversation these bots require careful training processes with big datasets containing labelled examples of the adequate answers to many possible questions. These bots could be integrated to bus stops and could even operate behind a fictional character. Instead of checking long lists and big tables, the user would be able to ask questions in natural format as: “Is bus number 3 delayed?” or “How can I go to Hervanta departing from Pispala?”. While making gaining the information more natural, it could be easier to use within larger share of different types of user groups.

All-adaptive last-mile autonomous solutions for public transport

Autonomous shuttlebuses are one answer to last-mile mobility needs within suburbs, such as Hervanta in Tampere. The public transport network of the area is planned to be developed with autonomous feeder lines connected to the tram network. Pilots of this are already planned within the SHOW-project[36]. However, given that it would be possible to acquire new information from the public transport network, such as the capacity and other passenger metrics and there would be a system that can estimate the future demand based on the information, it would be possible to know, when demand from passengers meet the supply of feeder vehicles and when there’s a mismatch. The events can be easily forecasted, such as ice hockey game ending in the city centre, or something that is more challenging to forecast, like a sudden rush of tourists coming in the city. For further research, it would be interesting to look, whether it would be possible to scale the amount of autonomous feeder vehicles based on the projected and measured demand, to make the network adaptive to different situations.

36 SHOW - Shared Automation Operating Models for Worldwide Adoption

Appendix A: Workshops

1st workshop

The first workshop[37] was hosted on 7.10.2020 on Microsoft Teams. The workshop aimed to gather the different stakeholders and businesses interested to discuss about the actual needs, ideas and possibilities of the Smart TrafficHub to support the upcoming work on the project.

The workshop hosted four different presentations around the topic. The present status of different concepts related to intelligent bus stops and the prior research related to AI and data gathered from different transport sources were presented. Third presentation explained the guidance systems in the city of Tampere and how it could be linked to public transport guidance in the future and the last presentation showed the operator-viewpoint of running a tram stop and how different concepts would work in their environment. At the end of the workshop, time was left of conversation regarding the needs and ideas that could be further examined.

2nd workshop

The second workshop[38] was hosted on 11.2.2021 on Microsoft Teams. The main agenda was to have conversation and answers regarding the main concept of Smart TrafficHub: what should be showed to users, how it's gathered, who provides it and where it should be accessed? To gather this information, we also hosted a short survey to allow participants and other stakeholders to answer.

The workshop hosted three different presentations linked to the topic. At first, the accessibility of different vulnerable to exclude user groups in public transport were presented based on surveys and interviews of the SmartRail-project. Then, the basics of real-time information were presented, consisting of what is currently provided by the buses and what type of information is generated to for example the transport office or a transport operator. From results within the project, the workshop had a short demonstration of the first steps on how GPS-location and speed data could be used to estimate the traffic results and the first results of the forecasting model and its accuracy build on that data.

Based on background information already provided in the survey, the main points were presented to support the conversation phase of the workshop. The main workshop-part was then divided in three phases. At first, some questions and comments arise from the presentations. Secondly, the attendees were divided into three separate breakout rooms to freely start a short conversation regarding the aspects of the workshop. At the end of the workshop, all participants gathered ack to main room, where these conversations were summarized and presented to all attendees.

3rd workshop

The third workshop[39] was hosted on 15.4.2021 on Microsoft Teams. Its main agenda was to go through the project findings (Chapter 6 on this report) and raise discussion of the future research and piloting topics mentioned.

37 Ensimmäinen työpaja 7.10. (in Finnish) - schedule and materials at project web page

38 Toinen työpaja 11.2. (in Finnish) - schedule and materials at project web page

39 Lopputyöpaja 15.4. klo 9-10 (in Finnish) - schedule and materials at project web page