Data Economy - Winter Semester 2019/20

Smart Urban Mobility

How do cities use data to understand and manage transportation patterns in order to ensure efficient and sustainable urban mobility?

Lukas Groeninger[a], Hsin-Yu Ku[b], Lukas Schmid[c], Shaurya Dev Singh[d], Eva Tesch[e]

October 16, 2020
Supervisor: Prof. Dr. Paul Drews

[a] *Matr. No. 3038589, lukas.groeninger@gmail.com*
[b] *Matr. No. 3038591, hsinyu.jade.ku@gmail.com*
[c] *Matr. No. 3038594, lukas.schmid@stud.leuphana.de*
[d] *Matr. No. 3038547, shaurya.d.singh@stud.leuphana.de*
[e] *Matr. No. 3038587, eva.tesch@stud.leuphana.de*

Contents

Glossary

API	Application Programming Interface
GPS	Global Positioning System
IoT	Internet of Things
LADOT	Los Angeles Department of Transport
MaaS	Mobility as a Service
MDS	Mobility Data Specification
MSP	Mobility Service Provider
ROI	Return on Investment
SAV	Shared Autonomous Vehicle
WBCSD	World Business Council for Sustainable Development
WHO	World Health Organisation

List of Figures

List of Tables

1 Introduction[1]

The usage of Big Data has changed the industry of transportation and will continue to do so (Zhu et al. 2019). One specific part of transportation is the passenger transport in cities, which leads to our research topic: Smart Urban Mobility. This term refers to modern solutions which allow moving around the city to be fast, comfortable and sustainable (HERE Mobility 2020). Transformation in this field is necessary and relevant due to fast population growth and urbanization, making daily commuting in cities a social, economical and environmental challenge (Cloutier et al. 2017). By 2030, urban population is expected to reach 5.2 billion people, which will be over 60% of the total population (Ritchie and Roser 2018). Hence, the United Nations has addressed this issue within the Sustainable Development Goal #11.2:

> "By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, notably by expanding public transport, with special attention to the needs of those in vulnerable situations [...]." (United Nations 2015, p. 26)

Through preliminary literature research, five recurring challenges were identified to be major issues or problems for transportation in cities. These served as subareas to break down the overall topic:

1. Segregation and Social Exclusion:
   How to create a transport system accessible to all citizens?
2. Road Safety and Accidents:
   How to improve safety on the roads and reduce deaths caused by traffic?
3. Air and Noise Pollution:
   How to reduce pollution and resulting health problems stemming from traffic?
4. Connectedness of Data:
   How to make intermodal transportation beneficial and easy to use?
5. Traffic Overload:
   How to optimize traffic flow in cities and reduce congestions?

An increasingly important approach to tackle the described challenges that cities are facing is the application of Data Science Methods or Big Data Analytics. The extent of recent literature about this field shows how much is going on at the moment in order to change the industry. Traffic data is told to contain hidden values (Neilson et al. 2019, p. 1) and unlock an array of opportunities to transform the mobility ecosystem overall (WBCSD 2020, p. 3). For instance, it could reduce and decarbonize transportation through shared mobility, autonomous and electric vehicles as well as smart public transit (Rolnick et al. 2019, p. 19). Data could furthermore help to understand crash risks for

[1] written by Eva Tesch

improving road safety (ITF 2019, p. 7) and also reveals valuable information about social exclusion caused by transportation disadvantage (Lucas 2012, p. 110).

Throughout this report, a framework will be presented and applied to classify, select and analyze a range of case studies in order to answer our research question: How do cities use data to understand and manage transportation patterns in order to ensure efficient and sustainable urban mobility? Putting it all together, the discussion part will show similarities and differences that appeared across subareas and cases. A final conclusion sums up major findings with regards to our research question. Eventually, we will reflect on the project work.

2 Approach and Framework Development[2]

A possible approach would be to use the Business Model Canvas (as described in, for example, Osterwalder and Pigneur 2010). However, the focus of this paper lies on cities as the actors to tackle the described problems, not on private companies. There is some debate as to whether the concept of business models is applicable to the public sector at all. Some authors advocate its introduction (Kaplan 2011; Ranerup et al. 2016), but the examples found in a literature research that used business models in the public sector either try to understand the role of the public sector to influence private companies' environment (Poel et al. 2007), or focus specifically on e-government (Janssen et al. 2008). The use of the business model was therefore discarded, mainly for two reasons: Firstly, the business model is centered around the idea of capturing value, i.e. creating services for which costumers are willing to pay (Osterwalder, Pigneur, and Tucci 2005, p. 10, p. 17; Osterwalder and Pigneur 2010, p. 14f). In contrast, the public sector does not rely on revenue streams to finance its operations. Secondly, the scope of actions the public sector may or have to engage in is often legally restricted (Peters 2015, p. 20). Based on a preliminary literature research, we therefore develop framework to analyse and compare cases. It is loosely based on Tao et al. (2014, p. 90) and consists of five questions: What objective does a case try to reach? What information does it gather? What data does it use to attain the information, and which methods does it use to do so? And what transport modes does the case apply to? An overview of all the analysed cases can be found in the appendix.

3 Results

3.1 Social Exclusion[3]

3.1.1 Relevance of the Problem

Segregation, social exclusion and lack of participation through an unequal access to the public transport systems are issues which cities of today are facing. But why is social exclusion or the lack of social cohesion a problem and how is it related to the (public)

[2]written by Lukas Schmid

[3]written and researched by Lukas Groeninger

transport system? And how could data science methods address and/or analyze these problems? These questions will be answered by looking at specific cases of how different cities can address these issues.

First there is the question of what segregation means and what implications it has for cities. Why should it be seen as bad or as a problem that needs to be fixed? Segregation is an ambivalent, contradictory and a highly complex phenomenon. There are scholars who argue for and against it and it can be seen as an balancing act for politicians and city planners. On the one hand segregation serves a purpose for people who are coming new into a city and or a society. Immigrant neighbourhoods form bridgeheads, provide information, practical help and also social support. Only on the basis of such a secure identity immigrants can look forward to a productive confrontation with the surrounding foreign society (Häußermann and Siebel 2001, p. 73). The ethnic economy and the own cultural and social infrastructure not only make everyday life easier for immigrants, they are often attractive to locals and thus offer places and occasions for communication between cultures (ibid., p. 73). You can find many examples in immigrant towns like New York when walking from Chinatown to little Italy. In this regard segregation is a functional condition for the successful integration of migrants. People choose a specific area or city quarter for a reason. In the 1980s the German alternative or punk scene in Kreuzberg / West-Berlin was also highly segregated. But there is a difference between the ethnic or cultural (functional and mostly voluntary) segregation and a more structural segregation on the basis of socio-economic constraints (ibid., p. 75). The segregation, for example, of members of the upper class in an exclusive neighborhood or even gated community ensures rising property prices. Often the spatial inequalities are the result of social dynamics of inequalities.

Segregation can therefore be seen as “the projection of a social structure onto space” (Cassiers and Kesteloot 2012, p. 1912). In that case the social structure is defined through several inequalities for example with regard to socioeconomic positions, education, housing or political representation. Many of these dimensions are connected and result in the social division of spaces in the city. These inequalities can be considered as strong structural drivers of segregation. That is why segregation should not be reduced to the spatial separation of different ethnic groups, but regarded with a broader view that is also linked to general societal exclusion mechanisms (ibid., p. 1912).

How can we promote the ability to move easily and freely around the city and find places to live and work? An important aspect in this respect is the (public) transport system of a city. It can ensure the accessibility to these spaces for all citizens so that they can meet and participate in societal life. Because what is a city when not “a human settlement in which strangers are likely to meet” (Sennett 1977)? If this is not the case and citizens are constrained in their possibilities of traveling around, the danger of social exclusion arises. Therefore, as socio-spatial inequalities increase, social exclusion has become a relevant category for the policy making of cities (Schönfelder and Axhausen 2003, p. 273). Social Exclusion does not have to mean poverty or social isolation, but these criteria often appear together. It rather relates to “a regular physical and social exclusion from the resources of a dignified life” (ibid., p. 273). This might not be a specific or strong definition, but a dignified life consists amongst other things of the “integration in the wider networks of civic life” (Schönfelder and Axhausen 2003, p. 273). This is where the chances of an accessible public transport system become relevant. Transport and land use policies have the power to decrease perceived costs of travel or reinforce the risk of social exclusion. When people have their home location far away from transport services and must walk large distances to stops or stations this can increase these costs. The “psychologically weighted sum of travel times, out of pocket costs and comfort” (ibid., p. 273) play an important role in the question whether people engage in travel and whether they participate fully in society. Obviously low costs of travel cannot guarantee social inclusion but definitely make it more likely.

Karen Lucas (Lucas 2012) has introduced a framework where she brings together the concepts of social disadvantage and transport disadvantage where the intersection is called transport poverty. This transport poverty “leads to inaccessibility to essential goods and services, as well as ‘lock-out’ from planning and decision-making processes, which can result in social exclusion outcomes”(ibid., p. 106). This social exclusion is reinforcing the transport and social disadvantage again. Transport policies have the possibility to intervene in this vicious circle. Figure 1 illustrates her framework.

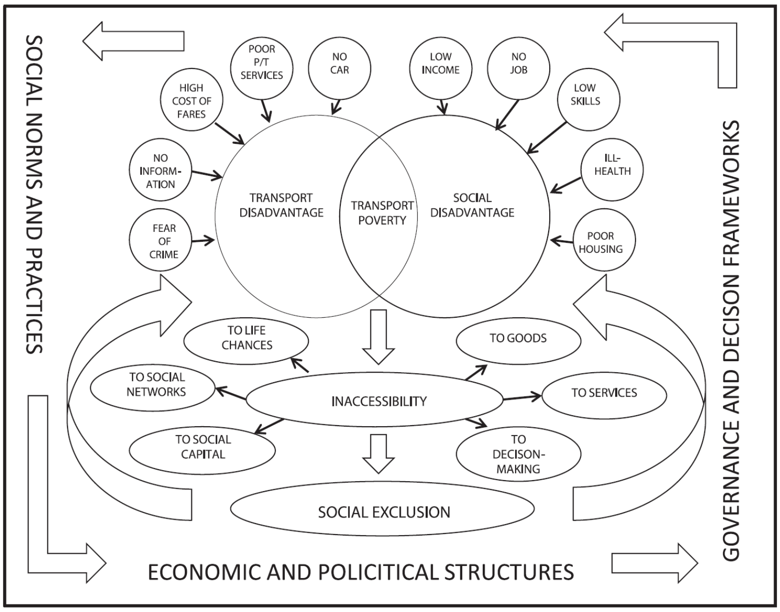

Figure 1: Model of social exclusion as a result of transport poverty (Lucas 2012)

3.1.2 Case Selection Procedure

Now I will introduce three different cases where segregation and accessibility of transport systems are analyzed to help city planners tackle the issues of social exclusion. Three scientific articles from the field of study of urban transport were chosen. The articles were published in the Journal of Urban Affairs and the Urban Studies Journal. The nature of the cases are therefore more conceptional and can be understood and used as a basis for urban planning in these cities.

The cities of interest are Madrid, Stockholm and Jerusalem which can be identified as European capital cities where the city network of today's streets and places developed over centuries. All of these three cities pursue a sustainable urban mobility plan and acknowledge social exclusion and accessibility as an important issue. That means the selected cases are similar in these dimensions, which means as well that I decided not to inspect cases with most differences. The comparison therefore cannot be applied to other regions or continents. The cases will be compared with respect to the possibilities of addressing the before mentioned issues of social exclusion and accessibility, which (data science) methods were applied and what implications this has for the future transport and city planning.

3.1.3 Case Description

Geographies of ethnic segregation - Stockholm (Sweden) The study wants to analyze if the public transport infrastructure represents a contributory factor in maintaining separation between foreign-born and ethnic Swedes. The transport mode which is focused on is the metro system of the city of Stockholm. It consists of seven lines which are divided in three subcategories (the blue, the red and the green lines). Additionally Rokem and Vaughan (Rokem and Vaughan 2019) used data of housing allocation policies and calculated prominent ethnic clusters in the city. The data which has been used for their analysis consists mainly of demographic data from the database managed by the Stockholm County Council in combination with data about housing allocation and the metro system. Besides quantitative data, the researchers also conducted several qualitative interviews with municipal planners. The methods that were applied can be summarized under the areas of spatial network analysis and the model of mobility flows. Space syntax analysis of street network configuration is a method for urban analysis that uses graph-mathematical measures of the relative accessibility of the street network to model the potential for movement across urban systems. Thereby they want to analyze the opportunities of interaction and the level of social segregation of economic immigrants of past decades and of newly arriving refugees of recent years.

Multimodal transport and potential encounters with social difference - Madrid (Spain) The author of the article Carpio-Pinedo cites the sociologist Pierre Bourdieu who said that the repeated experience of the spatial distance affirms social distance (Carpio-Pinedo 2019, p. 2). Carpio-Pinedo analyses transport infrastructures as facilitators of encounters with social difference. At the same time he promotes Multi-accessibility as a relevant and complementary policy avenue to enhance social cohesion through transport planning and land use policies in the future. He demonstrates two planning scenarios for the case of the Madrid metropolitan area and illustrates the application, utility and interpretability of the instrument. The data which is used can be divided into the public transit network, the automobile network and the street/pedestrian network. The data for the automobile network stems from the private company TomTom. The data for the street and transit network comes from public sources of the city of Madrid.

Segregation, mobility and encounters - Jerusalem (Israel) Urban segregation can be formed but also transformed due to the mobility flows and the public transport connections. The case of Jerusalem is special as the cities street network has been developed for more than two thousand years. Besides that the city is not only socio-economically but also ethno-nationally divided with a history of tensions and conflicts between groups in the last decades. These differences can also be seen in different levels of funding of planning and construction projects between Arab and Jewish areas. The case study aims to analyze the cities street network and its intersection with public transport provision as well as the socio-economic and ethno-national settlement patterns. The transport modes which are focused on are Bus and Tram lines of the public transit network and walking for the street network. There are separate Bus systems for the Jewish and the Arab area of which the Jewish system is far more developed. The data for the street network was obtained from openstreet.org and the demographic information from the Statistical Yearbook of Jerusalem 2015. The method which has been applied was spatial network analysis.

3.1.4 Comparison and Summary

All of the three selected cases want to address the issues of social exclusion with regard to the urban (public) transport system. The authors see the reason for the risk of social exclusion mostly in spatially segregated communities and in an inaccessible transport system. This risk arises when citizens of different social backgrounds don't interact with each other or even do not encounter each other in the places of their everyday life.

Carpio-Pinedo (2019) makes a case for the practical implementation of his concept of multi-accessibility as planning tool for cities transport and land use policies. He wants to focus on the role of multimodal networks for social diversity as well as for possible policy strategies. For him it is clear that this strategy should emphasize on "the increase of multimodally-balanced areas and the reduction of unimodally-accessible segregating spaces that foster the gathering of homogeneous crowds" (ibid., p. 25). The researchers argue for Stockholm that urban and public transport policy plays an important role in maintaining social segregation. The policies of recent years especially with respect to housing allocation could result in barriers to overcoming this segregation through mobility and co-presence in the city. This stands in contrast with the image of Sweden as a model of a tolerant, egalitarian and multicultural welfare state. For the case of the city of Jerusalem the authors raise the importance of "connectivity as an issue of spatial and social justice, rather than a function of urban economic development" (Rokem and Vaughan 2018, p. 3468).

The urban mobility system with its network of streets, buses, metro stations etc. represents on the one hand places of social encounters and on the other hand has the potential to change the mobility flows and which people could potentially interact with each other. It has been demonstrated that the transport systems of a city have a huge power of either facilitating or limiting the possibilities of different groups of people participating equally in societal life. That is the reason why the building of new pathways, bike lanes, streets, tram stations etc. has such a big impact on the life of people in the city. The studies revealed that the housing and land use policies of cities play a similarly important role. Transport and Housing policies are linked and have to be thought together. The use of data science methods can play a major role in informing policy makers in cities and in accompanying the process of implementing and evaluating policy changes. Methods like spatial network analysis and the modeling of mobility flows in cities can detect previously overlooked or neglected patterns of social inequalities or inaccessibilities. There seems to be a high potential for data science methods of making the spaces in our cities more accessible, to reduce the risk of social exclusion and for helping governments to make better and more informed decisions.

3.2 Road Safety and Accidents[4]

3.2.1 Relevance of the Problem

Safety problems are closely coupled to road mobility. Road safety is critical when heading towards sustainable urban mobility (Engels et al. 2019). It is a complex challenge, where urban density combined with private and public vehicles sharing crowded streets with vulnerable road users (vehicles, pedestrians, cyclists and motorcyclists) (Glasco 2018). According to the World Health Organisation (WHO), the number of road traffic deaths continues to rise (WHO 2018).

With a view to preventing fatalities and serious injuries, 'Vision Zero' started in Sweden in 1997 as a strategy to eliminate all traffic fatalities and injuries. (Vision Zero Network 2020b). Vision Zero is a concept calling for focuses on traffic systems design and its management for government. Under Vision Zero, European countries showed a successful 55% decrease in road fatalities since 2001 (WHO 2018). In contrast, the United States lag behind in road safety with no down trend in overall fatalities until 2018. Nevertheless, US cities have set out to form a Vision Zero Network started in 2012[5] and commit to develop and share practices toward an American prototype of Vision Zero (Shahum 2016).

Our research showed that, during the past 20 years, an increasing emphasis is being placed on the role of data to achieve Vision Zero. "Data-driven" is recognized as one of the nine components of a strong Vision Zero commitment (Vision Zero Network 2020a). It emphasizes the importance that "city stakeholders commit to gather, analyze, utilize, and share reliable data to understand traffic safety issues." The role of data is often highlighted for proactive transport safety interventions (PTSI), since "connected vehicles, smartphone apps, ubiquitous sensors, data sharing and machine learning make PTSI possible and prevent crashes before they happen." (ITF 2019, p. 81). Focusing on that connection, our research gave rise to the interest in how cities in the US improve urban road safety using big data or data science methods.

[4] written and researched by Hsin-Yu Ku

[5] https://www.citylab.com/transportation/2019/11/vision-zero-data-traffic-deaths-pedestrians-cyclist-safety/601831/

3.2.2 Case Selection Procedure

To develop evidence-driven road safety policies, it first requires data systems with reliable data collected (Peden et al. 2004). Thus, we first looked into cases regarding data collection or integrated systems, for example in San Francisco with its transportation-related injury surveillance system (San Francisco Department of Public Health 2020) and a data-integrating system in New Jersey (Numetric 2020).

Further research provoked interests into data-informed proactive network management. This approach focuses on designing road systems and transportation policy to avoid accidents proactively. Traditional approaches for road safety rely on crash data to identify the need for interventions. In contrast, proactive network management and PTSI predict risk levels and intervene before serious crashes happen by conducting analyses of risk factors and incidents (ITF 2019). Cases of different cities were examined to see how proactive approaches are practiced using big data that were not scalable in the past without machine learning capability. Deeper research revealed similar PTSI among cities, some of which with little information to be inspiring, or too simple to be analysed by our framework. After evaluation, New York was chosen because it offered the opportunity to analyze the approaches of different sectors to Vision Zero in one city.

In terms of case source, various types of source are chosen for better coverage of information. For further case analysis, we used sources from a combination of articles, news, websites, white paper and webinars.

3.2.3 Case Description

Telematics in city fleets as a Vision Zero Tool - New York City (USA) The selected case is about the cooperation between Geotab, a private telematics company, and departments of the city of New York. Telematics is a method of tracking vehicle using the global positioning system (GPS) to capture data — from speed, braking, fuel use of vehicles, to road conditions — wireless and in real-time (Michael 2018). As part of the city's Vision Zero initiative, the city implemented technology by Geotab to monitor its municipal vehicle fleets (police, fire, transportation fleets and more) and analyse the generated data in real-time. This way, about 35,000 vehicles were equipped with telematics devices (Vision Zero Network 2019) in which SIM cards and modem enabling the communication on the cellular network (Michael 2018). The case looks at how telematics fits for Vision Zero by its data-gathering and analysis capabilities. Various information collected for implied PTSI can be seen in table 1).

In an example for collision prevention, from the data collected, reports of detected collision with visualizations are sent to fleet operators for understanding certain dangerous driving behaviors. Method-wise, technologies like embedded accelerometers is used, together with machine learning and data analytics (detailed methodology not described).

With telematics, New York's government developed an operations and incident management system, tying real-time data from city fleet vehicles to city's other safety initiatives into one easy-to-use system. Thus, besides fleet managers, municipalities can also take proactive action (like speed limitation), with the information revealed by aggregated

Focus	Information sought-after	Implied PTSI
Fleets	driving behaviors (ex: speed, breaking, acceleration patterns etc.)	safety training, real-time coaching for fleet drivers
Fleets	collisions factors and locations	new fleet technology*
Safety factors	roadway conditions (ex: rough surfaces, potholes etc.)	city's Street Improvements Project

*ex : Automatic Emergency Braking systems / Forward Collision Warning

Table 1: Information gathered for possible solutions to PTSI

data (such as data that highlights hazardous intersections).

Datakind: Predicting Crash Locations - New York City (USA) This case is conducted by DataKind, a data science nonprofit, aiming at showing how data science can help the Vision Zero movement.

To develop a traffic fatalities and injuries prediction framework, DataKind leveraged open data and both public and private open geospatial data. Testing on predicting power of statistical models and machine learning algorithms, it came up with three final models (DataKind 2020). One of them uses random forests to predicts traffic volumes, the others use random forests and logistic regression to identify locations with many accidents. This information in turn was used to understand the effectiveness of interventions into the New York's road infrastructure. Besides, to understand the relative safety for different types of road users, four separate instances of each models are tested for different transportation 'modes' – auto, bike, pedestrians and all.

Overall, models are built to find likely crash locations to offer the government a basis to build upon for broader urban planning in the future.

Dash: App-Generated Data for Individual Safety Behaviour - New York City (USA) To advocate value of real-time user-generated data for transportation policy making regarding road safety, the developer of Dash, a telematics platform app, engages an academic team from NYU graduate school, to make analysis by integrating data from Dash with open public data.

The app Dash collects data in two ways. First, it leverages user's cell phone sensors when activated for data like GPS position Second, it collects data (such as fuel usage, engine and oil temperature etc.) from the vehicle's on-board sensors via Bluetooth connection. (Daud et al. 2020) For the analysis in this case, the academic team utilized eleven data fields out of 42 data fields collected by the app. It aims to demonstrate potential usage and challenges with integrated data to address road safety issues in transportation policy with four parts. The first part shows limitation of only monitored data, leading to potentials of real-time data from platforms like Dash. The second part reveals detailed methodology regarding gathering and processing Dash data along with public open data. It aimed to aggregate information in finding areas of speeding and hard braking, both as proxies for road safety. During the process, Geographic Information System, some algorithm (unknown) and packages in R were used. The third part shows data analysis results to identify locations of speeding and braking as well as Street Improvements Project evaluation with visualization. See Figure 2 for an example. Finally, policy recommendations are given (such as automating vehicle inspection process, designing usage-based insurance product) to further promote real-time data usage in transportation policy making for and beyond road safety.

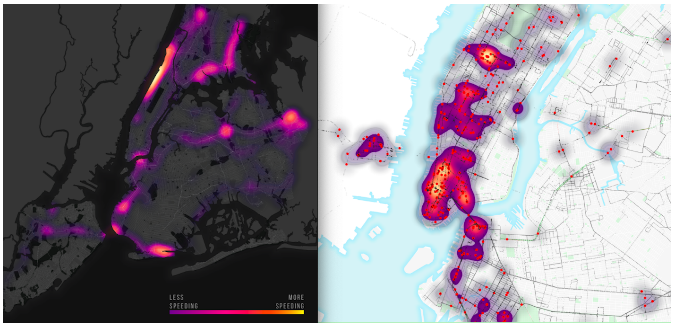

Figure 2: Locations of speeding (left) and hard braking (right) (Daud et al. 2020, p. 20)

3.2.4 Comparison and Summary

Even though all the cases come from New York, they pertain distinctive aspects and approaches. A reason for this is that each case involves different sectors. Different sectors target for safe mobility with each distinct angles of views, and thus results in interesting comparisons between cases. The city fleets case describes how government uses technology from a private company to manage and make use of data from city fleets. The DataKind case shows a non-profit helping government in Vision Zero initiatives with data science approaches. The case of real-data from Dash reals a private company advocating its data with a third academic party for governmental policy-making. The sectors involved for each cases can be seen in Figure3.

All three cases touch upon all modes of road transport. While the case of telematics involves monitoring public city fleets, the other two cases give direct attention to private road users. All cases use both first-hand and second-hand data, but partially focus on one. Case two uses mostly open data, while the other cases emphasize data collection through telematics devices.

All cases look for sets of distinctive information, among which are two shared groups of information for the same objective towards Vision Zero. First is identifying crash-related factors and places. Second is evaluating the effectiveness of safety interventions, aiming for effective PTSI. When looked closely, two cases with telematics both search for areas where speeding and hard braking occurs as it can be a proxy for "near accidents."

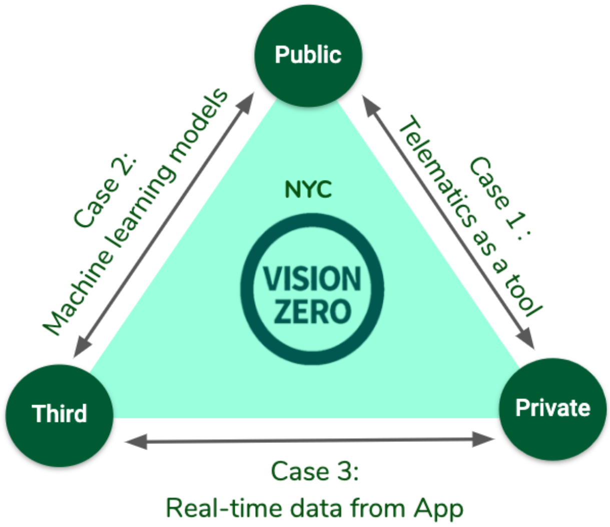

Figure 3: Sectors involved that conducted the cases

This demonstrates that instead of relying on traffic fatalities and injuries data, other data can be examined as a mean of PTSI for transportation policy making.

All cases incorporate usages of big data and data science, yet each laid different emphasis, which due clearly to the nature of sectors involved as mentioned earlier. Information of the city fleets mostly comes from the private company marketing its products in government application. Thus, it mentions no details about applied techniques, simply referring to usage of GPS and real time reports by some unknown machine learning methods. In comparison, The other two cases dive deeper into methodologies (machine learning algorithm and data analysis procedures respectively).

Also due to the nature of different sectors, case about city fleets has no mentions in potential limitations, while the others do. DataKind points out that models show little information about the impact of interventions into infrastructure due to a lack of data that was available when the models were created. (DataKind 2020, p. 2) The case of Dash additionally reveals limitations about data gathered by the app and phone sensors. By the app, data is limited by the number of app users and by the frequency of app usage. Phone sensors also have its limits, due to the nature of telematics such as GPS inaccuracy. Also, the data possesses strong selection bias since it is recorded not from randomly users but only from voluntary users that gives consent to share data (Daud et al. 2020, p. 17).

Both case of city fleets and Dash reveal the importance of aggregated or integrated data collected from telematics. However, only the latter points out the issue of data anonymity since it collects data from private vehicle users. It stresses that data used in the study was anonymized and not traceable back to individual user (ibid., p. 17).

Though the three cases are placed in different sectors, they ultimately appeal to the public sector to implement improvements to road safety. Moreover, all cases mention the possibility of broader applications in transportation problems as well as city planning. Case of telematics (Evan Mancini 2019) suggests that aggregated data also allows cities to map air quality or traffic flow DataKind indicates the traffic volumn model can be applied for broader urban planning questions like land use. Case of Dash highlights the usage of real-time data can redesign parking spaces to minimize congestion and improve air quality.

To sum up, Vision Zero appears to promote cross-sector collaborations in New York. During the process, data keeps its prominent role. "Vision Zero uses data shared across agencies and departments to identify problems, find solutions and ultimately, protect the lives of New Yorkers on the street" (New York City Mayor's Office of Operations 2017, p. 17). Result-wise, efforts towards Vision Zero in New York presented some success, with 28 percent decline in traffic fatalities since the launch of Vision Zero in 2014 (New York City Mayor's Office of Operations 2018). The chosen cases support the view that road safety is a multisectoral responsibility. This is hightlighted since more than ten years ago, while safe mobility continues to be an evolving issue that nowadays still requires commitment and informed decision-making by government, industry, nongovernmental organizations and international agencies all together (WHO 2004).

3.3 Air and Noise Pollution[6]

3.3.1 Relevance of the Problem

In view of our research question, we have chosen the sub-area concerning air and noise pollution, caused by road traffic in particular. Air pollution is considered to be an extremely relevant problem in today's times and poses significant health hazards to people. As air quality declines, it is known to cause chronic respiratory problems, cardiovascular diseases and cancer (WHO 2020a). Moreover, according to WHO's Regional Office for Europe, "[t]ransport is one of the main sources of air pollution, for which evidence on direct effects on mortality as well as on respiratory and cardiovascular disease is firmly established" and "[s]ome 40 million people in the 115 largest cities in the European Union are exposed to air exceeding WHO air quality guideline values for at least one pollutant" (WHO Regional Office for Europe 2020a). In addition to the harmful air quality effects originating from road transport, noise pollution is also a significant negative externality to be considered. Excessive levels of noise are detrimental to human health as they can lead to sleep disorders, affect social behaviour negatively as a result of reduced annoyance tolerances and cause cognitive impairment among children (WHO 2020b). The WHO's Regional Office for Europe remarks (WHO Regional Office for Europe 2020b):

> "Traffic noise alone is harmful to the health of almost every third person in the WHO European Region. One in five Europeans is regularly exposed to sound levels at night that could significantly damage health."

To narrow down the frame of analysis, we will focus on one particular segment of transportation – road transport. We chose this segment as road transport contributes

[6] written and researched by Shaurya Dev Singh

to 30 percent of particulate emissions in Europe and 50 percent of particulate emissions in OECD countries, according to the WHO (WHO 2020a). In addition, road traffic is also the leading cause of community noise in most cities, and noise levels are positively correlated with traffic volumes and speeds (WHO 2020b).

3.3.2 Case Selection Procedure

In the following sections, we will examine cases where Big Data approaches are used to tackle the problems of traffic induced air pollution and noise pollution. Our focus will be on the collection, organisation, analysis and presentation of air quality and noise data, which is a crucial requirement for the formulation of policies and sustainable urban planning. The three cases chosen for analysis and critical discussion are as follows:

1. Case Study published by GSMA: Telefónica Brazil, Predicting air pollution levels 24 to 48 hours in advance - São Paulo (Brazil) (GSMA 2018)

2. Big Data Analytics in Smart Mobility: Modeling and Analysis of the Aarhus Smart City Data set - Aarhus (Denmark) (Zenkert et al. 2018)

3. Smart city: A system for measuring noise pollution - Belgrade (Serbia) (Jezdović et al. 2018)

We chose these three cases from our relevant case pool because[7] all of them have a common thread of utilising Big Data approaches. The Brazil and Denmark cases put emphasis on data collection and in-depth data analysis using Inferential Statistics and Machine Learning respectively. Whereas, the Belgrade case puts a higher emphasis on designing a mechanism for collection, organising, storing and displaying data using an Internet of Things (IoT) Big Data architecture. Moreover, Visual tools like statistically generated graphics, maps, Application Programming Interfaces (APIs) and complex flow diagrams were employed in all three cases. These visual tools foster a more coherent understanding of processes and outcomes.

In addition to the commonalities that help to establish a frame of reference for comparison, the cases are also exclusive in various segments. The Brazil case is a case study conducted by a trade body representing mobile network operators globally, called the GSMA. This case contains an implemented system. On the other hand, the remaining two cases are scientific papers which discuss big data management system architectures and data analysis with scientific rigour and detail. It is to be noted that due to the reason that the two cases are scientific papers, they are tested prototypes and not implemented on a large scale. Secondly, the Brazil and Denmark cases pertain to air pollution and the Serbia case tackles noise pollution. We focused on cases that encompassed multiple aspects related to big data such as collection, storage, organisation, analysis and visualisation of outcomes. However, our bibliography contains cases that focus on particular

[7] Cases that are filtered out due to the reasons provided: Sharma et al. 2018; Rybarczyk and Zalakeviciute 2017; Rana et al. 2009; Filipponi et al. 2008; Wang et al. 2008; Hasenfratz et al. 2012; Kaur et al. 2017.

data analysis techniques as well. In the scientific paper by Kaur et al, several machine learning approaches to predict air quality are compared, such as Artificial Neural Networks, Radial Basis functions, Support Vector Machines, etc (Kaur et al. 2017). But, we did not select this case and similar algorithm oriented scientific papers for further exploration as they did not fit well into the overall frame of analysis - where we tried to look at particular use cases and gauge how big data was utilised in them. Lastly, the Brazil and Denmark cases are based on a city wide level tackling poor air quality in the cities of São Paolo and Aarhus respectively, whereas the Serbia case pertains to one crowded section of the city of Belgrade, tackling noise pollution.

We believe that due to the above mentioned factors, the narrowed case set contains a diverse combination of exclusive cases, while also sharing some important commonalities for better insights.

3.3.3 Case Description

Mobile Network Data for Air Pollution - Sao Pãolo (Brazil) The first case that we selected is a case study conducted by the GSM Association, which is a trade body representing mobile network operators across the world. The best practices that they brought out were of Telefonica, Brazil. Telefonica has developed and implemented a technical system in the city of São Paulo, Brazil, that can efficiently model and predict air pollution levels 24 to 48 hours in advance. The system is based on Machine Learning and Predictive Analytics approaches. However, the working of the algorithms were not discussed in detail as it is not a scientific paper. Essentially, the organisation uses mobile data (eg. Global Positioning System (GPS) positioning, devices synchronising with cell towers, messages and calls) as a proxy measure for traffic flows. This data is collected and eventually scaled up to be representative of the entire city. Final air quality predictions and maps are generated by combining the mobile traffic data with historical air quality data and weather data, using Machine Learning models. Telefonica Brazil claims that their index for traffic is highly correlated with actual observed traffic measurements, by up to 94 percent (GSMA 2018).

Healthiest Route Visualisation - Aarhus (Denmark) The second case that we have chosen is where researchers from the University of Siegen, Germany, analysed an open data set for the city of Aarhus, Denmark. The data set contains sensor data regarding traffic, pollution levels, weather and parking. The objective of the study was to calculate the healthiest road route between two points. The researchers accomplished this by undertaking parallel processing of sensor data through Apache Hadoop and the MapReduce algorithm, and developed a prototype Graphical User Interface system using the Shiny R package. The results were displayed on a map, facilitating ease of use. The web application allows users to input data, obtain the healthiest route and also observe other parameters such as pollutant information and dominant transport types in different road sectors. For the healthiest route calculation, the researchers used traffic and pollution data, recommending that weather data and parking data can be incorporated into the design as well. The researchers also opened up possibilities for further analysis with

Machine Learning, by providing key statistical measures for the variables (Zenkert et al. 2018).

Tackling Noise Pollution using IoT - Belgrade (Serbia) The final case that we selected is where researchers from the University of Belgrade, Serbia mapped noise pollution for one of the busiest parts of the city. They used two types of stations to collect noise data – fixed stations and mobile phones using an android application. The collected data was then unified and organised in a cloud based system using the MongoDB Non-Relational Database and ultimately visualised and presented on a map. An API was used to ensure proper functionality of the framework and provide the users information based on time of day, location and also samples of recordings. Noise maps were generated, and location and time based data was displayed in the web application. Using their prototype system, the researchers observed that noise levels were above the permissible limit, requiring government intervention (Jezdović et al. 2018).

3.3.4 Comparison and Summary

On evaluating the Brazil case, a unique mechanism in which mobile network data (which is streaming in nature) is used as a proxy measure for volume and flows of traffic was observed. We believe that this is a positive step towards cost-efficient data collection as it is very expensive to collect sensor data for every single intersection in the city. Moreover, the costs are further inflated as there is always a need to connect the sensor points through a cloud based network, so that the data can be merged. According to the case study, the proxy index generally is shown to be highly correlated with the actual observed traffic, by up to 94 percent. Furthermore, according to a survey done by an analytics company Newzoo, smartphone users will go up from 3.5 billion in 2020 to 3.8 billion in 2021 (Statista 2020). These numbers signify a growing trend in the number of smartphone users, making such projects feasible for further investment purposes. Bringing down data collection costs is essential to ensure that data is continually generated and policy makers are always updated with the latest facts and figures before taking any decisions. In addition, air quality sensors often suffer from device faults, battery issues and communication problems, negatively impacting the quality and accuracy of data that they collect (Kaur et al. 2017). Thus, another positive outcome of the system implemented by Telefonca, Brazil is that it reduces the dependence on such sensors.

The Denmark case presents an excellent use case for the analysis of open data sets. The study shows that open data can be used for public good and improving access to open data can act as a driver of innovation, ensuring transparency in decision making and improved information dissemination. In addition to the benefits originating from making open data available, another key insight is that the parallel processing algorithm and framework is developed in R and Apache Hadoop – both of which are open source software. This reflects that the technical infrastructure for complex data analysis exists and is accessible to everyone, without investing large sums of money. We believe that in an age of growing data literacy and popularity of open source software, the benefits stemming from their combination will grow exponentially over the years.

In the Serbia case, a possible improvement can be considered, which is developed in a study by Hasenfratz et al. The idea is to ensure mobile sensor calibration through the internet with static reference stations, maintained by public authorities (Hasenfratz et al. 2012). This would ensure that accuracy loss during the collection of noise pollution data is reduced for the mobile devices, through a cohesive data collection mechanism. Now, a key best practice that we observed across the three cases is the usage of visual tools such as APIs and the generation of maps. The pollution mapping systems in the Brazil and Denmark cases have the potential to be constructive catalysts for information dissemination. Such visual representations are important as they enhance the gravity of the message being conveyed. Moreover, the healthiest route visualisation concept used in the Denmark case breaks down complex chunks of information and puts it in a format where people can use it to take everyday decisions (Zenkert et al. 2018). If the system was improved to display real time data, it would act as an incentive mechanism coercing people to take steps to combat air pollution and bring more attention to the issue. In the Denmark case, the prototype web application provides users the option to calculate the healthiest route between multiple areas according to the time of day. The application also displays specific pollution information for all road sections of the chosen journey (ibid.). In the Noise pollution interface prototype discussed in the last case, noise signals can be tracked by location and time (Jezdović et al. 2018). The interesting addition in this interface is the availability of audio tracks for the chosen location. This type of information can prove to be useful for the authorities in pinning down the sources of noise pollution.

Lastly, we examined a critical limitation of the Brazil case, regarding data privacy. Telefonica claims that it ensures data privacy in the process of its model generation by anonymizing the data. They claim that the data is anonymized by adding noise, removing key identification numbers, normalisation and extrapolating out of their sample (GSMA 2018). Despite the fact that Telefonica claims that it uses these procedures, they were not discussed in detail in the case study as the case study was conducted by the GSMA - which is an organisation that serves to protect the interests of mobile network operators. Mobile data contains sensitive information that must be strongly secured and protected. Public authorities need to play a role in safeguarding citizens against the misuse of this data. The anonymization measures also need to be regulated by an appropriate body and the regulatory authority must ensure that non-anonymized sensitive data does not get into the hands of private corporations. Additionally, A suggestion is also made in the case study to provide the anonymized mobility data to third parties as it can prove to be highly beneficial for companies with regard to advertising and other commercial platforms (ibid.). However, this would raise the need for regulation and data protection even more.

3.4 Connectedness of Data[8]

3.4.1 Relevance of the Problem

This section covers cases related to the sharing of data between different sectors. While this topic has been mentioned in other cases (section 3.2), the cases covered here aim to show how shared data, especially between the public and the private sector, can enable, but also endanger new approaches in urban transport policies.

In the last ten years, new forms of mobility services have established themselves in cities around the world. Ride-hailing services like Uber blurred the boundary between private car trips and taxis, and micromobility sharing systems offered vehicles to rent for only a few minutes.

These changes create regulatory challenges for cities. First of all, innovative mobility service providers may become dominant enough to develop substantial bargaining power vis-à-vis cities. Even now, they often shape the way transport works in cities. Secondly, in spite of the industry's claims, their operations can increase emissions (Suatmadi et al. 2019), worsen accessibility issues (Golub et al. 2019), generally increase traffic volumes (International Transport Forum 2015, p. 20) and weaken existing, more efficient forms of public transport (Huber et al. 2019, p. 10). And thirdly, the diversification of the transport service landscape means that cities may have to bargain with multiple actors where earlier they could implement simple top-down regulation[9]. This means that old forms of regulation are not likely to be successful.

The changes associated with new forms of mobility also create opportunities. Different authors have shown that thse changes can make traffic more efficient (Suatmadi et al. 2019, p. 228; Golub et al. 2019, p. 697; Gilibert and Ribas 2019, p. 406). If cities want to alleviate the potentially detrimental effects of new forms of transport while enjoying the positive ones, they have to regulate mobility service providers (International Transport Forum 2020, p. 33; Huber et al. 2019, p. 42-48).

3.4.2 Case Selection Procedure

To do justice to the variety of cases, we developed three categories which describe the relationship of the public and the private sector in projects where data is shared. They closely match classifications in the literature (Huber et al. 2019, p. 76; Meurs et al. 2020, p. 179). From each of the three sketched relationships, one example will be discussed. The selection of examples was largely a convenience sample, i.e. those examples were chosen for which good information was available in databases and the web in general.

The first category encompasses top-down approaches in which cities compel private mobility service providers to share data with the city, but not with each other or the public. In effect, the data in this approach is generated by private companies, then the public sector is granted access to it. Examples to this approach are the "open transport partnership", launched by the together with three mobility providers (World

[8] written and researched by Lukas Schmid

[9] Of course, cities will probably either face one strong, or a multitude of smaller providers. Our argument here is that both constellations lead to new regulatory challenges.

Bank 2016), principles “for sharing, protecting, and managing data” (NACTO 2019) developed by the US-based National Association of City Transportation Officials. On a more technical level, Ciociola et al. (2017) develop a platform to collect data generated by mobility service providers. In a simliar vein, the Los Angeles Department of Transport (LADOT) developed an application programming interface (API) and compels mobility providers to use it (LADOT 2018b). The city of Portland uses a similar standard for slightly different purposes (Move Forward 2019b; Move Forward 2019a), and some other US cities have followed (Bailey 2019). In the past year, two German hackers spread that idea to German cities (radforschung 2019b).

The second category sees public actors on a level playing field with existing private actors, with the public actors integrating private mobility services into their existing transport portfolio. In this relationship, the public partner is regularly presented by transport agencies who seek cooperation with private companies to extend their services. Successful partnerships lead to public transport agencies allowing customers to additionally book private rides or vehicles through their portal. Examples are the city of Kochi (India) with “Kochi One” (Singh 2020), different cities in Belgium (Storme et al. 2020, p. 201), and Taiwan (Chang et al. 2019) as well as the USA[10], Berlin (Germany)[11] and Vienna (Austria)[12]. Some private companies have already specialised in cooperating with cities, offering solutions to implement MaaS (see Turnn [13] and TransDev with ”Breng Flex”[14] in the Netherlands as well as iomob [15] and Kyyti [16]). How data is shared depends on the specific circumstance, but as the selected case of Pinellas County illustrates, too little shared data can threaten a project’s success (SUMC 2019).

In the third category the public sector takes on the role of the facilitator by making its data publicly available. Private companies are then allowed to step in and provide MaaS-solutions based on public-sector data. The main difference to the other categories is that ultimately, it is private companies who offer services to customes. An existing practice is the API of publicly owned Deutsche Bahn (Germany), which enables access to timetables and other data (Deutsche Bahn 2020). Digitransit is a Finnish open-source project that aims at exactly that kind of multi-modal integration via APIs[17]. In Luxembourg a platform exists that combines a wide range of transport services and offers routing for cars (Frost 2019). Another project was Beeline in Singapore (Goodall et al. 2017, p. 121). It provided public transport data which companies could use to develop private services, like mini busses. Hereinafter, we will discuss the case of Whim in Helsinki (Finland), because it best illustrates how cities can actively create an environment in which MaaS-solutions are developed.

[10] Dallas, `https://www.dart.org/news/news.asp?ID=1179`, Atlanta `https://www.ajc.com/blog/commuting/marta-pilots-free-buses-and-uber-partnership/73IoThyBdDvbJkf6JncalK/`

[11] `https://www.berlkoenig.de/die-idee`

[12] https://www.ots.at/presseaussendung/OTS_20170608_OTS0065/wienmobil-mit-neuer-linien-app-die-stadt-im-griff

[13] `https://www.turnn.nl/over-turnn/`

[14] `https://www.tudelft.nl/en/2017/transport-institute/breng-flex-a-succes/`

[15] `https://www.iomob.net/`

[16] `https://www.kyyti.com/`

[17] `https://digitransit.fi/en/`

3.4.3 Case Description

Gathering Data to Regulate Mobility Service Provides - Los Angeles (USA) In 2018, LADOT published a strategy to “ensure that [...] autonomous fleets [...] in the sky and on the ground [...] are first and foremost safe” (LADOT 2018a). At the core of that strategy was the development of an API to facilitate data exchange between mobility service providers and the city. During its development phase, MDS had been intended for various transport services and modes, but it saw its first roll-out in connection with e-scooters (Zipper 2019). On the conceptual level, providers have to supply real-time information about their vehicles in the form of number of vehicles in operation, their location and their condition (LADOT 2018b, p. 1). The city can access that data and thus monitor the mobility service providers. Additionally, it can communicate back to the providers and, for example, define no-parking zones for vehicles (Zipper 2019; radforschung 2019a). When a provider updates its database with the API, these no-parking zones come into effect automatically for all customers of that service. In theory, cities could even regulate pricing models of scooters (Zipper 2019). Technically, the MDS is an open-source API currently maintained by the Open Mobility Foundation. It is currently available under a Creative Commons Attribution 4.0 International Public License on GitHub[18] and consists of two endpoints. One of the endpoints allows for the real-time provision of data from the mobility service providers to the regulatory agency; the providers have to make sure they grant access to the API. For this endpoint, two variants are implemented, on of which (the Agency API) allows for more fine-grained analysis, but is more complex to set up than the other (Provider API). The third endpoint, called Policy, allows the agency to regulate the MSP (Open Mobility Foundation on GitHub 2020). At every point, the agency maintains full control over the database.

Lack of Oversight due to Lack of Data - Pinellas County (USA) Facing budgetary constraints and political hostilities, the Pinellas Suncoast Transit Authority (PSTA) developed the idea of a partnership with private mobility service providers (SUMC 2019, p. 6f). At the core of the project called ”Direct Connect” stands a subsidy to users to cover the last or first mile of a journey, i.e. to reach the closest bus stop from home or the other way around. While the details changed over different phases of the project, the subsidy was eligible for every trip that had one end near a bus stop, and the other one inside a zone associated to that bus stop (ibid., p. 20). This mechanism intends to avoid people replacing the bus with other mobility services and guarantees subsidies only to trips that are likely to continue via bus. The pilot continues to operate and currently has contracts until 2021 with three private mobility service providers, among them Uber and a local taxi service (ibid., p. 20). On a customer-level, no further integration of the services seems to be planned, even though Uber customers directly see subsidised prices at booking instead of regular prices. While the project is multi-modal, it only extends to busses and ride-hailing (with no other public transport system in place in the county). The PSTA was given only very limited access to the data of its partners,

[18]https://github.com/openmobilityfoundation/mobility-data-specification

especially Uber, which only grants access to data via monthly invoices that state the number of subsidised trips. This lead to problems when the PSTA noticed (rather by chance) that Uber had charged more rides than would have been eligible for subsidy, due to a fault in the geofencing (SUMC 2019, p. 21). Even though the PSTA did not have to pay for those trips, this mistake “led to vastly inflated performance reports for nearly a year following the Phase III service design” (ibid., p. 25). One report concludes that the project might not have been continued, had it not been for these false figures (ibid., p. 25).

Providing Data to the Private Sector - Helsinki (Finland) The city of Helsinki aims at integrating all modes of urban transport at the level of the individual user to implement Mobility as a Service (MaaS) (Heikkilä 2014). The start was made by a Finnish legislation requiring all providers of transportation to enable access to their ticketing system via open APIs (Hirschhorn et al. 2019, p. 184). This created an opportunity for MaaS Global, a private company founded in Helsinki, to act as a broker between different mobility services and customers. It integrates different services like public transport, but also taxis, bikes, and rental cars and offers its customers an option to pay a monthly flat fee to use a certain number of these services. MaaS Global sees its business potential in saving its customers the need to pay for an own car (Jaffe 2019) and offering them an equally flexible, but cheaper mobility experience using various mobility services (Hirschhorn et al. 2019, p. 185). Essentially, Whim allows users to either pay for a plan that then contains a certain amount of free rides in included modes of transport, or to pay for single trips in-app.

3.4.4 Comparison and Summary

The cases in this section rely on data generated by transport services and shared across sectors. The data can be comprised of geospatial trip-level data or can include timetables in the case of public transport. What information can be extracted from the data depends on the level of detail in which these data are actually shared, which differs significantly among the cases. The downsides of sharing only aggregated data is especially highlighted by the case of Pinellas County, where lack of trip-level data led the programme to overestimate its impact. The reluctance of, in this case, Uber to share data is not limited to the private sector; in Helsinki, the city’s public transport agency has not granted Whim full access to its services (Zipper 2018) and intends to build its own MaaS platform (Hirschhorn et al. 2019, p. 188). Accodring to the emphasis placed on data sharing, methods used are limited to the implementation of data sharing technologies. The PSTA case is again an exception, since it relies exclusively on traditional methods, without realising the potentials of digitally generated data.

The fact that all of the cases aim at integrating formerly separated forms of urban mobility into one platform shows that multi-modal solutions are on the rise. MaaS Global places most emphasis on this, but the PSTA case shows how, even outside of start-up culture, very traditional players have come to realise the potentials of multi-modal transport solutions. Due to this integration, new forms of urban transport will inevitably encompass both private and public actors, and many private companies already offer services to the public sector in implementing MaaS solutions. Subsequently, the goals of implementing such solutions ranges from far-fetching goals like replacing all cars in Helsinki, to providing pragmatic answers to regulation challenges in Los Angeles, and finding a solution to budgetary and political constraints in the PSTA case.

The cases also show that the discussion around who should provide transport services is not purely technical, but informed by ideological alignments. The PSTA's pilot only became necessary after local initiatives had lobbied against the PSTA (SUMC 2019, p. 7), and Los Angeles' MDS has seen significant protests by companies like Uber (Hawkins 2019) and other groups (Center for Democracy & Technology 2018). On the other hand, the MDS is often portrayed as a move by cities to regain legitimate control about the public space (Zipper 2017).

3.5 Traffic Overload[19]

3.5.1 Relevance of the Problem

This last subtopic will deal with traffic overload due to the high number of vehicles on the roads, as a further challenge cities are facing. Especially at rush hour, many people are commuting at the same time which leads to a lot of traffic and congestion but also a resulting concentrated air pollution (see section 3.3).

One traffic measure is "congestion level", which is the amount of extra travel time needed in congestion compared to free traffic flow (TomTom International BV 2020). Results show congestion levels of as high as 71% in Bengaluru (India) and Manila (Phillipines) and, as most congested German city, Hamburg at 34% (ibid.). This means requiring 30 minutes on uncongested roads, the average actual travel time is approx. 50 minutes in Bengaluru. These figures tell how relevant the topic of traffic in cities is. Not only from a social perspective, losing precious time while in traffic, but also from an environmental point of view do congestion levels of this extent counteract undertaken green urban planning efforts. Driving in heavy traffic congestion has been proven to significantly increase greenhouse gas emissions and fuel consumption (Veurman et al. 2002). Additionally, time spent in traffic has a negative economic effect, since time spent in a vehicle often is lost working hours.

Private car-usage is one of the principal contributors to climate change. Cars account by far for the highest contribution to energy consumption in the passenger transport sector (Chapman 2007). The high consumption levels stem not only from fuel usage and its supply, but also from the manufacturing process of a car. In the meantime decreases the interest of consumers for private car ownership (Botsman and Rogers 2010).

Given above aspects, major improvements in the urban transportation have to and will continue to occur. We conducted an extensive literature and case-study research about initiatives to improve urban traffic flows, with special focus on Data Science related projects.

[19] written and researched by Eva Tesch

3.5.2 Case Selection Procedure

The literature research has shown how traffic and congestion issues are discussed differently by scientists and companies. While technology providers tend to publish their innovations online with focus on beneficial outcomes for advertising purposes, scientific articles go much more into detail about mathematical methodologies and also contain a critical evaluation of results, but often lack a simulation or even implementation in a city. Therefore, at a preselection stage, some case studies have already been sorted out due to lack of details or (even conceptual) implementation.

We would like to give a diverse insight into approaches towards optimizing traffic flow. In the following, several aspects are explained that led to the selection of four final cases for analysis. The problem of traffic overload in cities can be approached through various starting points with two main goals in mind:

- General traffic flow optimization - by road infrastructure, traffic lights, efficient parking, policies (e.g. road space rationing or reversible traffic lanes)
- Reduction of number of vehicles - by improving and incentivising public transport, ride-sharing or climate-neutral options as walking and cycling

Additionally, researches and especially their results and findings can address either of two main target groups: public entities such as the government, their transportation department and related corporations on the one hand, or private entities including citizens and companies on the other hand.

To cover all aforementioned aspects of classification (nature of case, goal, target group), following diverse case portfolio has been selected for further analysis, with case one and four being scientific papers, case two a blog article and case three a conference talk:

	Optimize current traffic flow	**Reduce number of vehicles**
Public	1. Traffic Prediction for Optimal Traffic Signal Timing	3. Traffic Lights Prioritizing Cyclists and Public Transport
Private	2. Smart Curbside Parking	4. Shared Autonomous Vehicles

Table 2: Systematic overview of selected "Traffic Overload"-cases

3.5.3 Case Description

Above selected cases will now be described, all having one superordinate objective: to ameliorate and optimize urban traffic flow caused by passenger transport. However, the approaches differ from one another.

Traffic Prediction for Optimal Traffic Signal Timing - Cranberry Township (USA) The first case investigates how to monitor and predict traffic and hence reduce congestion. The site is Cranberry Township (USA), a city with high levels of traffic, mainly caused by commuting to Pittsburgh (Yao and Qian 2019, p. 1). Traffic coordination is further complicated by a nearby intersection of several major roads. Researchers try to predict incident-included congestion at least 30 minutes in advance, which means that accidents or other incidents are anticipated. This shall help to propose optimal traffic signal timing, serving as decision basis for the Traffic Management Center of the city (ibid., p. 1). Whether they have actually implemented the findings is not clear from the article.

Data inputs are traffic speed data from sensors (no details given), updated every five minutes, and historical incident records in an area of 30-minute driving distance, as well as weather information (ibid., p. 2). An artificial recurrent neural network uses this data to predict traffic levels. Another algorithm uses the raw data inputs as well as the predictions to create a signal recommendation system as support for the Traffic Management Center with different contingency plans, which are action plans reacting to possible traffic conditions.

Training and testing of the model prove that the prediction model together with the signal recommendation system work well and accurately in proposing solutions at least 30 minutes before the congestion happening. Error rates are much lower compared to previously used predictor models (ibid., p. 11). Exemplary, relatively precise predictions compared to real traffic can be seen in figure 4.

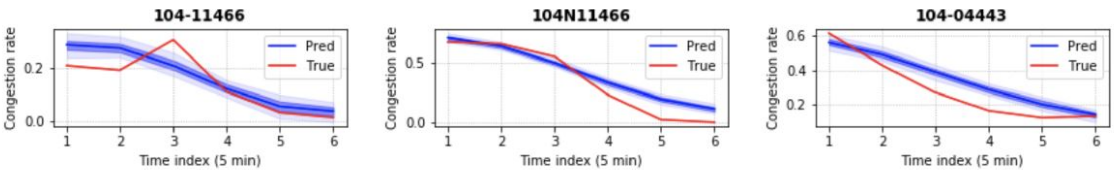

Figure 4: Excerpt of results showing model predictions versus actual traffic at a certain road segment after an incident had occured (Yao and Qian 2019)

Smart Curbside Parking - Shanghai (China) Since around 30% of urban traffic is caused by vehicles driving around looking after a parking spot (Krieg et al. 2018), quick parking can reduce congestion and fuel consumption. This, so far, is better manageable with parking garages than curbside parking, due to clear entrances and exits (Huawei Technologies 2020). Easier curbside parking within cities has been facilitated through a joint collaboration of Huawei, China Unicom Shanghai and TransInfo, which have created a new technology (ibid.). Narrowband Internet of Things is used coupled with magnetic vehicle detectors installed at parking spots. The collected sensor data is processed to detect free parking spots which are communicated to car drivers via a phone application. A trial has been implemented in Minhang district of Shanghai. The technology is said to have several benefits in terms of quality and durability, but especially results have been good (Huawei Technologies 2020). However no concrete numbers are given, and as this statement comes from the company itself, it should be viewed critically.

Traffic Lights Prioritizing Cyclists and Public Transport - Copenhagen (Denmark) Among several efforts towards a car-free Copenhagen is one project by the city's municipality in cooperation with a transport technology provider (Wandall 2016). This project aims at promoting the usage of bicycles and public transport while disincentivizing private car usage (Kabell 2017). Valby, a part of Copenhagen city with heavy traffic, served as trial area for the implementation of a smart traffic light system. Cameras were installed at intersections, detecting shapes of objects and classifying them into whether they are cars, lorries, bicycles or pedestrians. That information is transmitted to a computer that is functioning as a controller of the intersection. If busses or groups of cyclists are recognized, the computer can decide to leave lights green to enable these commuters to cross in the current green-phase. The longer car waiting times are justified by the lower road use efficiency of cars as compared to busses and cyclists. The result was satisfying, as busses could run 40% faster than before (ibid.). This led to a roll-out of 380 smart traffic lights (Vestergaard Andersen 2016). One major problem constitutes data privacy concerns when cameras are used in public. To clear these concerns, it has been stated that the cameras are not able to identify people (Kabell 2017).

Shared Autonomous Vehicles - Austin (USA) A conceptual study about autonomous driving and ride-sharing has been conducted in Austin (USA) (Fagnant and Kockelman 2018). Focusing only on car as transportation mode, specifically shared autonomous vehicles (SAV), the research aims at showing the potential impact that a combination of autonomous driving with ride sharing can have on reducing the amount of cars.

The used data was provided by a regional collaboration for metropolitan planning, containing a coded roadway network of the region and a trip table with origins and destinations (ibid., p. 145). Additionally, household travel diaries helped to know departure time distributions of citizens. These three datasets served as input to simulate traffic flow over 24 hours. For further analysis, only the most densely populated region had been selected with a random set of trips (see figure 5).

Based on this data-subset, an agent-based micro-simulation modelled a fleet of SAV that served all requested trips using a backward-modified Dijkstra's algorithm (ibid., p. 146). The model accounts for three different actors: travelers, the fleet management assigning travelers to cars, and the SAV taking the instructed routes (ibid., p. 146). By varying model parameters such as fleet size, conditions for trip and waiting times, demand and ride-sharing versus no ride-sharing, different scenarios were derived.

Major findings are that one SAV could replace about 10 conventional cars, assuming ride-sharing and a slight increase in travel times. Transport service provider could expect an estimated ROI of around 19% on fleet costs (ibid., p.157). However, total vehicle-miles traveled could rise due to empty drives. This can be optimized by increasing waiting and travel time.

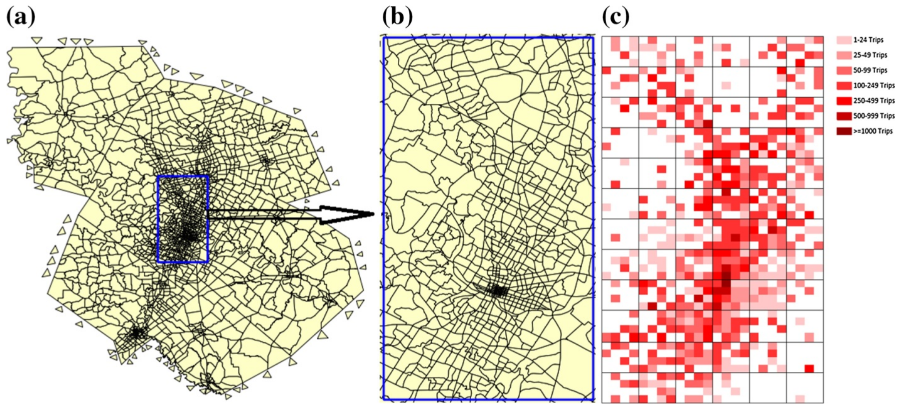

Figure 5: (a) Regional transportation network, (b) nework within the 12mi x 24mi geofence, (c) distribution of trip origins over 24h (Fagnant and Kockelman 2018)

3.5.4 Comparison and Summary

After the previous insights into usage of data in traffic, an overall summary and comparison of cases will convey a big picture about the state of the art of urban traffic flow optimization. This will be done along our framework.

The different approaches in this chapter 3.5 have in common an overarching objective: to improve the traffic flow on roads, cutting down commuting time and reducing transport emissions. Each case generated very specific information, but while two cases tackle the challenge of faster throughput of vehicles, the remaining focus on reducing the total amount of motorised vehicles on the road.

Similar stakeholders are affected by the different projects. Generally speaking, involved entities is the city with its citizens, companies and administration (especially the department of transportation). However, focusing on the people being directly addressed by the research to take action, cases one and three call on the city's transportation department in first place, as this is the entity to actually install traffic lights and use traffic predictions. On the contrary, parking measures and ride-sharing appeal to road users to adopt innovations, but also to private companies like Uber to roll out such projects, e.g. implementing an SAV fleet.

The cases applied very different methods to the data. The scientific articles specifically name the mathematical models applied (i.e. neural networks & Dijkstra's algorithm). Both rely on secondary spatiotemporal data. In contrast to this, the case studies presented by a private company and a city council respectively only briefly touch upon the details of the applied technology and focus on the projects' outcomes. Their data is mainly collected directly through sensors.

All of the analyzed cases show promising approaches towards smoother traffic, as the methods were proven to be successful, either conceptually or applied on small scale. However, most suggested approaches require some willingness of people and municipalities to change commuting patterns and working habits regarding the application of data. To conclude, above analysis shows that overall, smart traffic flow optimization is a topical issue and ongoing process. The usage of data can support this to a great extent.

4 Discussion[20]

There are several problems which cities of today are facing with regard to urban mobility. This is especially the case with rising populations, limited space and impending consequences of the climate crisis. How can we ensure an efficient and more sustainable transport system in this context? We are living however not only in a time of various challenges, but also in a time of numerous technological opportunities. These technological opportunities include the field of data science, which can provide solutions to the problems and challenges of our time. In our group we have identified five problem areas where data science methods could be applied to analyze and improve the mobility systems in our cities. Thereby the focus was on the nature as well as the handling of data in specific use cases.

Why did we think that the problems cities are facing can be tackled especially well with data science methods? This is because the issues at stake can be considered in many cases optimization problems or issues of pattern detection in complex and large amounts of data. The topic of traffic overload for example can be seen as a classical optimization problem where the traffic flow should be improved and road bottlenecks should be removed (see chapter 3.5). With the help of data science methods the infrastructure can be redesigned to reduce the number of vehicles on the road by promoting green alternatives. An example for the detection of overlooked patterns and structures would be the topic of segregation and social exclusion. The objective of the use cases of this area is to provide equal access to the transport system for all citizens and to detect inequalities of participation in the current structure. These inequalities were revealed for example in the use case of the city of Stockholm with respect to non-ethnic Swedes. Another classical issue is the handling of different types of data. In cities and the context of modern transport systems, there are countless possible sources of data to be collected and to be brought together. As the cases show, diverse sources as sensors to measure pollution levels, the GPS location of vehicles, monitoring the capacity utilisation of streets or railways or further information about time tables of buses and trams are all collected and analysed to yield information. That is why we considered the connectedness of data to be a crucial point for making cities safer and more sustainable. An example would be the development and implementation of the mobility data specification of the city of Los Angeles. In general we can say that the challenges of urban transportation are well suited to be tackled by data science methods.

To access data of different sources and to connect it, it is necessary to establish a system for the storage and the management of the data. This is especially the case if several organisations, institutions or different sectors are involved. We recognized that the topic of open or shared data was present in almost all our cases. The question arose how to make public databases accessible by different parties in order to realise the potential of data analysis. We have found that the challenges we selected might be of

[20] written by Lukas Groeninger

different angles, but the kind of data which is collected and analyzed is similar in many of our cases. The findings of the cases regarding the topic of road safety and accident prevention can also have implications on how cities can deal with congestion problems or air pollution. The same data that has been collected for one specific purpose can be used to address several other research questions. This is also true for other methods and models that have been generated. Therefore one could argue that there is a level of redundancy of the data, methods and models. In addition the collection of data through sensors and storing them is a difficult and cost-intensive task. So there is a lot of value in offering access to data for analysis and further data science methods and using synergy effects. What is also noteworthy when comparing our cases is that almost everywhere there has been some cooperation between public, societal and private actors. We found that it is mostly public data which is analyzed and used for training and the building of models by third sector parties or private companies. Afterwards the insight is brought back and used for example for informed decision making or regulation and planning purposes as the knowledge concerning data analysis is missing in administrations on the local governmental level. This approach is best demonstrated with the use cases of road safety and accident prevention in New York (see chapter 3.2). What also stood out when comparing our cases was the focus on a comprehensible and clear visualization of the results of the respective analysis.

The topics we selected had particularly one thing in common. It was the vision of a smart, sustainable, safe and connected city. This means as well that there is a focus on an urban transport system which does not rely solely on one mode of transport, but rather embraces the concept of multimodality. That includes “green” modes like cycling, walking and public transport like metro and buses but also new forms of transport like electric scooters and ride sharing. The car as the preferred means of transport around which the cities infrastructure is built, has become a thing of the past. However, it still has a place in the transport mix of our cities. The cities of our case studies aimed for a multimodal urban transport system while emphasizing that a change towards more sustainable modes of transport is necessary. How this system change towards sustainability can be managed in a controlled and effective way is a challenge that cities are facing. We are confident that data science methods can assist with exactly this challenge. Our case studies showed that the methods used, offer a tremendous potential to reach this vision of a smart, sustainable, safe and connected city.

5 Conclusion[21]

We analysed our research question through multiple problem areas, which were Segregation and Social Exclusion, Air and Noise Pollution, Road Safety and Accidents, Connectedness of Data and Traffic Overload. Big data has generated a significant effect on the urban mobility industry. The velocity, volume and variety of big data has pushed research and promoted data oriented decision making.

The first main trend that we found was the usage of open data. By making data sets

[21] written by Shaurya Dev Singh

publicly available, organisations can specialise in analysing the data and need not expend their resources in data collection. This would lead to benefits from specialisation and save time and costs such as the costs of installing sensors and conducting surveys. Open data also fosters research and collaboration between researchers, since they can operate with shared data and overcome the hurdles of data confidentiality. Hence, these factors contribute to an ecosystem of shared best practices, where parties can learn from other entities' studies utilising open data and even improve on existing analysis using the same data set. Furthermore, open data, if used in collaboration with open source software can enable anyone to contribute and present their results in a cost efficient manner. With the growing popularity of open source software for professional purposes and sophisticated analysis tools, their combination with open data sets can drive innovation. Additionally, a connected observation across our cases was the high reliance on visualisation tools like maps and APIs, leading to better information dissemination, enhancing the substance of the message being conveyed and ensuring that the data can be interpreted and utilised effectively.

The second key observation in our project was the cooperation between public and private sectors. In many of our cases, private companies conduct data analysis and provide their results to a public authority - such as the transportation department, planning section or pollution body. The public authorities can subsequently utilise this data for drafting policies and undertaking urban planning. Having policies grounded in data would lead to enhanced public support and improve transparency if the figures connected with the policy were made publicly available. Thus, public-private sector cooperation also has specialisation benefits associated with it as the government can divert their resources towards policy implementation, leaving the private entities to analyse data with their expertise and experience in the given field. However, these partnerships also have their shortcomings. In a report by the European Court of Auditors published in 2018, it was observed that Public Private Partnerships led to significant amounts of ineffective spending due to a lack of robust prerequisite analyses, poor strategic approaches, and delays due to institutional and legal frameworks (Ivanova et al. 2018).

The final insight that we came across in our sustainable urban mobility project was the issue of data privacy and anonymization of records. Despite all the positive implications of big data towards sustainable urban mobility, the problem of data privacy comes up in a large number of cases. Mobility data such as vehicle GPS data, taxi cab routes calculations and travel history are sensitive in nature and if such data falls into the wrong hands, it can pose security risks for the people involved. Moreover, with the advent of autonomous driving led by companies like Tesla, the reliance of companies on such data is increasing exponentially, raising the need for regulation. In some studies, there are claims of anonymizing records, so that the data cannot be traced back to people. But there is no concrete evidence provided as these are claims made by private companies. This is another downside to the involvement of private entities in the field of sustainable urban mobility, in addition to the demerits of Public Private Partnerships. The relevant pubic transportation authorities should work in tandem with commercial organisations to address and overcome concerns of data privacy. Priority should be given to raising awareness on regulatory requirements and how they affect different stakeholders, as there is a lack of clarity in the understanding of legal regulations. Comprehensively understanding the trade-off between data privacy and security, and smart mobility solutions is a crucial prerequisite to address and overcome related problems (Hill et al. 2017).

6 Reflection on Project Work[22]

The whole process of our project can be summed up and viewed by four stages:

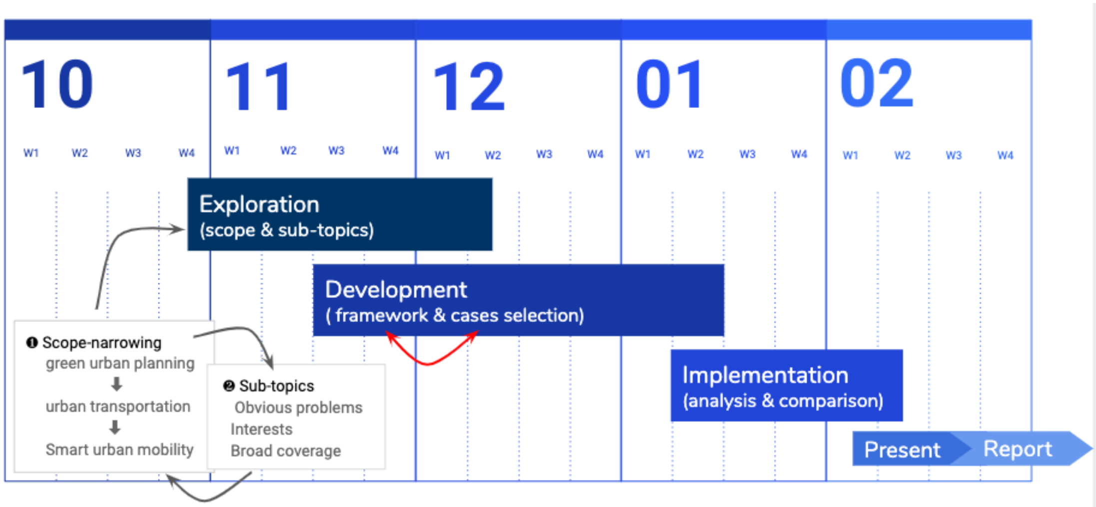

Figure 6: Overview of our project work

Exploration of Scope and Sub-topics:

In the exploration phase, we narrowed the focus of the paper to the research question. During the first month, we met on a weekly basis to share information. Starting from the general question of interests, to specifically formulated sub-topics, this stage we followed an exploratory approach. All ideas were documented in a google document for discussion and brainstorming. After two group meetings, distinct cases of urban transportation emerged from our research. After the third meeting, each of us had chosen one topic of interest. The only criterion set at this stage was that a case should have connection to a specific city. At this stage, we also formulated a preliminary research question.

Because we had not clearly delineated different sub-topics, we relied on a preliminary framework for further literature research. This framework looked at modes of transport, type of data used, methods used and solutions proposed.

Finally, we concluded that different cities share similar problems which they aim to solve with the help of data. As a result, we decided to classify sub-topics according to problems. The sub-topics of safety and pollution were chosen due to their clear positions

[22]written by Hsin-Yu, Ku

as problems to be tackled. The other sub-topics were developed after further research and discussion to ensure no over-laps and a broad coverage of problems.

Development of Framework and Case Selection

This stage proved to be the hardest. There had been confusion about the exact nature of the framework. After discussion, we decided it would be more reasonable to use a bottom-up framework to compare cases.

Having discussed more cases, we came up with a framework based on the preliminary classification of sub-topics. At this point, the distinction between the generation of sub-topics (according to problems) and the analysis of cases situated in the sub-topics (according to the framework) became clear. In the following weeks, the framework was further refined. At the same stage, the research question was finalised.

Finding relevant cases was easy, yet filtering abundant sources and pick cases was rather hard. Due to the diverse nature of the different sub-topics, we decided not to apply a single case-selection criterion for all sub-topics. Instead, selection of cases within sub-topics was done individually (as described in "Case selection procedure" of each sub-topics in section 3).

Implementation of analysis and comparison: After thorough analysis of cases in each sub-topic by our framework (see the appendix), we had an effective and inspiring meeting in which we shared all cases according to the framework. Discussion about similarity and essential aspects for data usage in sustainable urban mobility was provoked naturally. After going through the standstill of developing the framework the rest of the group work ran smoothly.

Overall Reflection

Reflecting on the whole work, from exploration to development, the approaches of dividing sub-topics under problems and developing bottom-up framework worked out effectively. Starting from the implementation phase, we set up plans to work for both the presentation and written report. Goals, to-dos, and deadlines for each of our meetings were set beforehand. The objective is for each individual to be on time. It turned out that most of us were on time, while some fell behind and needed extra support from the others. Though this resulted in some frustration, the support as a group worked out. We managed to fit our presentation in required time when some failed to keep in their time limits. We also sorted out the writing work to proofread others' work.[23]

Overall, although it would be better if everyone works equally on time, the whole group project reveals a learning process with mutual supports.

[23] Special appreciation to Lukas S. for technical support to fix the problems and styling in overleaf, the software used for the written report.

References

Bailey, Charlie (Apr. 22, 2019). *MDS, GBFS, and how cities can ask for data from micromobility providers.* remix. URL: https://blog.remix.com/mds-gbfs-and-how-cities-can-ask-for-data-from-micromobility-providers-7957ca639f16 (visited on 01/20/2020).

Botsman, Rachel and Roo Rogers (Sept. 14, 2010). *What's Mine Is Yours: The Rise of Collaborative Consumption.* Harper Collins. 310 pp.

Carpio-Pinedo, Jose (Oct. 14, 2019). "Multimodal transport and potential encounters with social difference: A novel approach based on network analysis". In: *Journal of Urban Affairs*, pp. 1–24.

Cassiers, Tim and Christian Kesteloot (July 2012). "Socio-spatial Inequalities and Social Cohesion in European Cities". In: *Urban Studies* 49.9, pp. 1909–1924.

Center for Democracy & Technology (Nov. 29, 2018). *Comments to LADOT on Privacy & Security Concerns for Data Sharing for Dockless Mobility.* Center for Democracy and Technology. URL: https://cdt.org/insights/comments-to-ladot-on-privacy-security-concerns-for-data-sharing-for-dockless-mobility/ (visited on 03/11/2020).

Chang, S. K. Jason, Hou-Yu Chen, and Hung-Chang Chen (Dec. 1, 2019). "Mobility as a service policy planning, deployments and trials in Taiwan". In: *IATSS Research* 43.4, pp. 210–218.

Chapman, Lee (Sept. 2007). "Transport and climate change: a review". In: *Journal of Transport Geography* 15.5, pp. 354–367.

Ciociola, Alessandro et al. (Aug. 2017). "UMAP: Urban mobility analysis platform to harvest car sharing data". In: *2017 IEEE SmartWorld.* 2017 IEEE SmartWorld. San Francisco, CA: IEEE, pp. 1–8.

Cloutier, Scott et al. (2017). "Measures of a Sustainable Commute as a Predictor of Happiness". In: *Sustainability* 9.7. Publisher: Multidisciplinary Digital Publishing Institute, p. 1214.

DataKind (2020). *Analyzing Traffic Crashes in New York City Using Machine Learning Based Approaches.* DataKind.

Daud, Muhammad et al. (2020). *Using real driving data for transportation policy: a proof of concept.* Robert F. Wagner Graduate School of Public Service.

Deutsche Bahn (2020). *REST Fahrplan-API Version Beta.*

Engels, Dirk et al. (Sept. 2019). *Topic Guide Urban Road Safety and Active Travel in Sustainable Urban Mobility Planning.*

Evan Mancini (July 3, 2019). *Telematics as a Vision Zero Tool.* Vision Zero Network. URL: https://visionzeronetwork.org/telematics-as-a-vision-zero-tool/ (visited on 01/04/2020).

Fagnant, Daniel and Kara Kockelman (Jan. 1, 2018). "Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in Austin, Texas". In: *Transportation* 45.1, pp. 143–158.

Filipponi, Luca, Silvia Santini, and Andrea Vitaletti (June 11, 2008). "Data Collection in Wireless Sensor Networks for Noise Pollution Monitoring". In: Distributed Computing in Sensor Systems, 4th IEEE International Conference, DCOSS 2008. Santorini Island, Greece.

Frost, Adam (Sept. 23, 2019). *Luxembourg upgrades travel app to full multimodal MaaS.* Traffic Technology Today. URL: `https://www.traffictechnologytoday.com/news/mobility-as-a-service/luxembourg-upgrades-travel-app-to-full-multimodal-maas.html` (visited on 02/23/2020).

Gilibert, Mireia and Imma Ribas (Nov. 18, 2019). "Synergies between app-based car-related Shared Mobility Services for the development of more profitable business models". In: *Journal of Industrial Engineering and Management* 12.3, p. 405.

Glasco, Jon (Aug. 21, 2018). *The long road to urban mobility safety.* Bee Smart City. URL: `https://hub.beesmart.city/solutions/en/smart-mobility/the-long-road-to-urban-mobility-safety` (visited on 01/09/2020).

Golub, Aaron et al. (Dec. 2019). "Assessing the barriers to equity in smart mobility systems: A case study of Portland, Oregon". In: *Case Studies on Transport Policy* 7.4, pp. 689–697.

Goodall, Warwick et al. (2017). "The rise of mobility as a service. Reshaping how urbanites get around". In: *Deloitte Review*, pp. 112–129.

GSMA (June 2018). *Case Study: Telefónica Brazil, Predicting air pollution levels 24 to 48 hours in advance in São Paulo, Brazil.*

Hasenfratz, David et al. (Jan. 2012). "Participatory Air Pollution Monitoring Using Smartphones". In:

Häußermann, Hartmut and Walter Siebel (2001). "Integration und Segregation - Überlegungen zu einer alten Debatte". In: *Deutsche Zeitschrift für Kommunalwissenschaften* 2001.1, pp. 68–79.

Hawkins, Andrew J. (Oct. 29, 2019). *Uber threatens to sue Los Angeles, as the fight over scooter data escalates.* The Verge. URL: `https://www.theverge.com/2019/10/29/20938212/uber-lawsuit-la-ladot-scooter-data-mds` (visited on 02/03/2020).

Heikkilä, Sonja (Apr. 28, 2014). *Mobility as a Service –A Proposal for Action for the Public Administration.*

HERE Mobility (2020). *Smart Urban Mobility: A Quick Start Guide.* HERE Mobility. Library Catalog: mobility.here.com. URL: `https://mobility.here.com/smart-urban-mobility-quick-start-guide` (visited on 02/25/2020).

Hill, Nikolas et al. (June 3, 2017). "Scoping Study into Deriving Transport Benefits from Big Data and the Internet of Things in Smart Cities". In: 4, p. 94.

Hirschhorn, Fabio et al. (Dec. 2019). "Public transport regimes and mobility as a service: Governance approaches in Amsterdam, Birmingham, and Helsinki". In: *Transportation Research Part A: Policy and Practice* 130, pp. 178–191.

Huawei Technologies (2020). *China Unicom Shanghai, Huawei, and TransInfo Cooperate in the Trial of NB-IoT-based Smart Parking.* Huawei. URL: `https://carrier.huawei.com/en/relevant-information/reshaping-b2b-together-through-cloud-services/unicom-huawei-smart-parking` (visited on 01/31/2020).

Huber, Christian et al. (2019). *DG4MaaS - Daten-Governance Regeln für Mobility-as-a-Service : Dokumentation.* Fachhochschule Potsdam, 1 Online–Ressource (116 Seiten, 4, 39 MB).

International Transport Forum (2015). *Big Data and Transport Understanding and assessing options.*

— (2020). *Urban Mobility System Upgrade. How shared self-driving cars could change city traffic.*

ITF (2019). *New Directions for Data-Driven Transport Safety.* International Transport Forum,

Ivanova, Iliana et al. (2018). *Public Private Partnerships in the EU: Widespread shortcomings and limited benefits.* Special Report 09. Luxembourg: European Union, Annex IV 1–4.

Jaffe, Eric (Aug. 27, 2019). *Whim is a mobility app with a modest proposal: Give up your car.* SidewalkTalk. URL: https://medium.com/sidewalk-talk/whim-is-a-mobility-app-with-a-modest-proposal-give-up-your-car-f2db0bf2ba65 (visited on 02/20/2020).

Janssen, Marijn, George Kuk, and René W. Wagenaar (Apr. 2008). "A survey of Web-based business models for e-government in the Netherlands". In: *Government Information Quarterly* 25.2, pp. 202–220.

Jezdović, Ivan et al. (Mar. 2018). "Smart city: A system for measuring noise pollution". In: *Smart Cities and Regional Development Journal (03-2018)*, pp. 79–85.

Kabell, Morten (Jan. 26, 2017). *Approaches to sustainable mobility. Copenhagen.*

Kaplan, Saul (Apr. 19, 2011). "Business Models Aren't Just For Business". In: *Harvard Business Review.*

Kaur, Gaganjot et al. (Oct. 2017). "Air Quality Prediction: Big data and Machine Learning Approaches". In: 017 5th International Conference on Sustainable Environment and Agriculture (ICSEA 2017). Vol. 1.

Krieg, Jean-Gabriel et al. (2018). "Unlocking the smartphone's sensors for smart city parking". In: *Pervasive and Mobile Computing* 43, pp. 78–95.

LADOT (June 13, 2018a). *LADOT announces a new digital playbook for mobility.*

— (Oct. 31, 2018b). *Mobility Data Specification: Information Briefing.*

Lucas, Karen (Mar. 2012). "Transport and social exclusion: Where are we now?" In: *Transport Policy* 20, pp. 105–113.

Meurs, Henk et al. (Jan. 1, 2020). "Organizing integrated services in mobility-as-a-service systems: Principles of alliance formation applied to a MaaS-pilot in the Netherlands". In: *Transportation Research Part A: Policy and Practice.* Developments in Mobility as a Service (MaaS) and Intelligent Mobility 131, pp. 178–195.

Michael, Craig (Jan. 8, 2018). *What is telematics?* geotab. URL: https://www.geotab.com/blog/what-is-telematics/ (visited on 01/09/2020).

Move Forward (Aug. 1, 2019a). *Along with Scooters, Portland Pilots Data Partnerships — Part 1.* Move Forward. URL: https://medium.com/move-forward-blog/along-with-scooters-portland-pilots-data-partnerships-part-1-e1706ead9c2f (visited on 01/21/2020).

— (Aug. 13, 2019b). *Along with Scooters, Portland Pilots Data Partnerships — Part 2.* Move Forward. URL: https://medium.com/move-forward-blog/along-with-scooters-portland-pilots-data-partnerships-part-2-5906be2552e2 (visited on 01/21/2020).

NACTO (2019). *NACTO Policy 2019: Managing Mobility Data.*
Neilson, Alex et al. (Sept. 2019). "Systematic Review of the Literature on Big Data in the Transportation Domain: Concepts and Applications". In: *Big Data Research* 17, pp. 35–44.
New York City Mayor's Office of Operations (Feb. 2017). *Vision Zero Year Three Report.* NYC.
— (Mar. 2018). *Vision Zero Year Four Report.* NYC.
Numetric (2020). *How Greater Data Accessibility is Driving New Jersey to Safer Roadways.*
Open Mobility Foundation on GitHub (2020). *Understanding MDS APIs.* GitHub. URL: `https://github.com/openmobilityfoundation/mobility-data-specification` (visited on 02/24/2020).
Osterwalder, Alexander and Yves Pigneur (2010). *Business model generation: a handbook for visionaries, game changers, and challengers.* Toronto: Flash Reproductions. 278 pp.
Osterwalder, Alexander, Yves Pigneur, and Christopher L. Tucci (2005). "Clarifying Business Models: Origins, Present, and Future of the Concept". In: *Communications of the Association for Information Systems* 16.
Peden, Margie et al. (2004). *World report on road traffic injury prevention.* Geneva: WHO, The World Bank.
Peters, B. Guy (2015). *Advanced introduction to public policy.* Cheltenhan; Northampton: Elgar.
Poel, Martijn, Andrea Renda, and Pieter Ballon (Aug. 21, 2007). "Business model analysis as a new tool for policy evaluation: policies for digital content platforms". In: *info* 9.5. Ed. by Pieter Ballon, pp. 86–100.
radforschung (Apr. 16, 2019a). *Mobility Data Specification für Kommunen, erklärt.* radforschung.de. URL: `https://radforschung.org/log/mds-fuer-kommunen-erklaert/` (visited on 01/20/2020).
— (Feb. 27, 2019b). *Rollersharing kommt nach Deutschland. Was deutsche Städte jetzt schon von anderen lernen können.* radforschung.de. URL: `https://radforschung.org/log/rollersharing-was-staedte-lernen-koennen/` (visited on 01/20/2020).
Rana, Rajib et al. (Jan. 2009). "Ear-Phone assessment of noise pollution with mobile phones". In: Proceedings of the 7th International Conference on Embedded Networked Sensor Systems, SenSys 2009, November 4-6, 2009. Berkeley, California, USA,
Ranerup, Agneta, Helle Zinner Henriksen, and Jonas Hedman (Jan. 1, 2016). "An analysis of business models in Public Service Platforms". In: *Government Information Quarterly* 33.1, pp. 6–14.
Ritchie, Hannah and Max Roser (June 13, 2018). "Urbanization". In: *Our World in Data.*
Rokem, Jonathan and Laura Vaughan (Nov. 2018). "Segregation, mobility and encounters in Jerusalem: The role of public transport infrastructure in connecting the 'divided city'". In: *Urban Studies* 55.15, pp. 3454–3473.

Rokem, Jonathan and Laura Vaughan (Sept. 2019). “Geographies of ethnic segregation in Stockholm: The role of mobility and co-presence in shaping the ‘diverse’ city”. In: *Urban Studies* 56.12, pp. 2426–2446.

Rolnick, David et al. (June 10, 2019). “Tackling Climate Change with Machine Learning”. In: *arXiv:1906.05433 [cs, stat]*. arXiv: 1906.05433.

Rybarczyk, Yves and Rasa Zalakeviciute (Dec. 20, 2017). “Regression Models to Predict Air Pollution from Affordable Data Collections”. In: *Machine Learning - Advanced Techniques and Emerging Applications*.

San Francisco Department of Public Health (2020). *San Francisco’s Transportation-related Injury Surveillance System*.

Schönfelder, Stefan and Kay W Axhausen (Oct. 2003). “Activity spaces: measures of social exclusion?” In: *Transport Policy* 10.4, pp. 273–286.

Sennett, Richard (1977). *The Fall of Public Man*. New York.

Shahum, Leah (Jan. 26, 2016). *10 Cities Lead National Effort to Eliminate Traffic Fatalities*. Vision Zero Network. URL: https://visionzeronetwork.org/vzn_focus_cities/ (visited on 02/03/2020).

Sharma, Nidhi et al. (2018). “Forecasting air pollution load in Delhi using data analysis tools”. In: *Procedia Computer Science* 132, pp. 1077–1085.

Singh, Mitashi (Jan. 1, 2020). “India’s shift from mass transit to MaaS transit: Insights from Kochi”. In: *Transportation Research Part A: Policy and Practice*. Developments in Mobility as a Service (MaaS) and Intelligent Mobility 131, pp. 219–227.

Statista (2020). *Number of smartphone users worldwide from 2016 to 2021*. Telecommunications. URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ (visited on 01/09/2020).

Storme, Tom et al. (Jan. 1, 2020). “Limitations to the car-substitution effect of MaaS. Findings from a Belgian pilot study”. In: *Transportation Research Part A: Policy and Practice*. Developments in Mobility as a Service (MaaS) and Intelligent Mobility 131, pp. 196–205.

Suatmadi, Anissa Yuniashaesa, Felix Creutzig, and Ilona M. Otto (June 2019). “On-demand motorcycle taxis improve mobility, not sustainability”. In: *Case Studies on Transport Policy* 7.2, pp. 218–229.

SUMC (June 2019). *When Uber Replaces the Bus: Learning from the Pinellas Suncoast Transit Authority’s "Direct Connect" Pilot. A First-Last Mile Case Study*.

Tao, Sui et al. (Sept. 2014). “Exploring Bus Rapid Transit passenger travel behaviour using big data”. In: *Applied Geography* 53, pp. 90–104.

TomTom International BV (2020). *Traffic Index 2018*. URL: https://tomtom.com/en_gb/traffic-index/ranking/ (visited on 01/24/2020).

United Nations (2015). *Transforming our World: The 2030 Agenda for Sustainable Development*.

Vestergaard Andersen, Anne (Mar. 1, 2016). *380 Intelligent Traffic Signals for Copenhagen*. State of Green. URL: https://stateofgreen.com/en/partners/state-of-green/news/380-intelligent-traffic-signals-for-copenhagen/ (visited on 02/18/2020).

Veurman, J. et al. (2002). "Emissions at different conditions of traffic flow". In: *WIT Transactions on The Built Environment.* Vol. 60.

Vision Zero Network (July 2019). *Telematics Final on Vimeo.*

— (2020a). *9 Components of a Strong Vision Zero Commitment.*

— (2020b). *What is Vision Zero?* visionzeronetwork. URL: https://visionzeronetwork.org/about/what-is-vision-zero/ (visited on 01/02/2020).

Wandall, Niclas Højgaard (Jan. 6, 2016). *Imflow in Valby. A case of implementing an adaptive traffic control system (Master Thesis).*

Wang, G, Frans van den Bosch, and Monika Kuffer (Jan. 2008). "Modelling Urban Traffic Air Pollution Dispersion". In:

WBCSD (Jan. 21, 2020). *Enabling data-sharing: Emerging principles for transforming urban mobility.* World Business Council for Sustainable Development (WBCSD). URL: https://www.wbcsd.org/tumremps (visited on 01/29/2020).

WHO (2004). *World report on road traffic injury prevention - Main messages.*

— (2018). *Ranking EU Profess on Road Safety: Global status report on road safety 2018.* WHO.

— (2020a). *Air pollution.* Health, environment, and sustainable development. URL: https://www.who.int/sustainable-development/transport/health-risks/air-pollution/en/ (visited on 01/16/2020).

— (2020b). *Noise.* Health and sustainable development. URL: https://www.who.int/sustainable-development/transport/health-risks/noise/en/ (visited on 01/14/2020).

WHO Regional Office for Europe (2020a). *Air pollution and climate change.* URL: http://www.euro.who.int/en/health-topics/environment-and-health/Transport-and-health/data-and-statistics/air-pollution-and-climate-change2 (visited on 01/12/2020).

— (2020b). *Noise.* URL: http://www.euro.who.int/en/health-topics/environment-and-health/noise/noise (visited on 01/17/2020).

World Bank (Dec. 19, 2016). *Press Release: The World Bank Launches New Open Transport Partnership to Improve Transportation through Open Data.*

Yao, Weiran and Zhen Sean Qian (2019). "Real-Time Traffic Monitoring and Prediction for Cranberry Township". In:

Zenkert, Johannes et al. (June 21, 2018). "Big data analytics in smart mobility: Modeling and analysis of the Aarhus smart city dataset". In: 2018 IEEE Industrial Cyber-Physical Systems (ICPS). IEEE, pp. 363–368.

Zhu, Li et al. (Jan. 2019). "Big Data Analytics in Intelligent Transportation Systems: A Survey". In: *IEEE Transactions on Intelligent Transportation Systems* 20.1, pp. 383–398.

Zipper, David (July 2, 2017). "Private Mobility Services Need To Share Their Data. Here's How." In: *citylab.com.*

— (Oct. 25, 2018). "Is Helsinki's MaaS App, Whim, the Future?" In: *citylab.com.*

— (Apr. 2, 2019). "Cities Can See Where You're Taking That Scooter". In: *Slate Magazine.*

Appendix: Overview of Covered Cases

	Segregation and Social Exclusion		
	Potential encounters with social difference	Segregation, mobility and encounters	Geographies of ethnic segregation
	Stockholm (Sweden)	Madrid (Spain)	Jerusalem (Israel)
Implementation	conceptual	conceptual	conceptual
Objective	Analyzing the impact that the transport system can have on segmenting and linking populations	Increase accessibility of the urban sphere for all demographics and increase social cohesion	Stress the importance of connectivity as an issue of spatial and social justice
Data	demographic data, grid data from public transport	demographic data, grid data from public transport, GPS Data for the automobile system	demographic data, grid data from public transport
Information	role of public transport infrastructure play separation between foreign-born and ethnic Swedes	How does transport enable individuals to engage in spatially diverse urban spaces?	Evaluation of newly established public transport lines
Method	spatial network analysis, model of mobility flows		
Transport Mode	public transport	public transport	public transport

Table 3: Overview of all cases covered in section 3.1, according to the framework

	Road Safety and Accidents		
	Telematics: As a Vision Zero Tool New York (USA)	Datakind: Predicting Crash Locations New York (USA)	Dash: App Data for Individual Driving Behaviour New York (USA)
Implementation	implemented	conceptual	conceptual
Objective	get information from city fleets towards vision zero safety goals for all road users	develop a fatality and injuries prediction framework for evaluation of city's Street Improvements Project towards vision zero	demonstrate potential uses and challenges of real-time data to address road safety towards vision zero.
Data	data from city fleets by telematics	open geospatial data	real-time data from Dash, open data.
Information	driving behaviors, collisions factors, road conditions, hazardous intersections	traffic exposure, high crash locations, road safety for different types of road users	locations of certain driver behaviors and built environment that affects behaviors
Method	machine learning, data analytics, GPS, visualization	machine learning algorithms, statistical models, visualization	GIS platform, data analysis in R, algorithms, visualization
Transport Mode	public transport	private transport	private transport

Table 4: Overview of all cases covered in section 3.2, according to the framework

	Air and Noise Pollution		
	Mobile Network Data for Air Pollution Sao Pãolo (Brazil)	Healthiest Route Visualisation Aarhus (Denmark)	Tackling Noise Pollution using IoT Belgrade (Serbia)
Implementation	implemented	prototype	prototype
Objective	efficient collection of air pollution data and providing policy support	to find the least polluted road route between two points	to create an IoT based data system for processing noise pollution data
Data	data generated by mobile networks, coupled with sensor data for pollution and weather	open data set containing sensor data for road traffic, pollution, weather and parking	sensor data for noise collected by fixed and mobile stations
Information	to model and predict air quality by using mobile network data	relative danger to health if a person commutes on a chosen route	noise measurements using IoT tools
Method	machine learning and predictive analytics	parallel processing of sensor data, statistical analysis and interactive visualisations	big data processing – data collection, cloud storage, analysis and visualisation
Transport Mode	All modes of road transport	All modes of road transport	All modes of road transport

Table 5: Overview of all cases covered in section 3.3, according to the framework

	Connectedness of Data		
	Mobility Data Specification	Direct Connect	Whim by MaaS Global
	Los Angeles (USA)	Pinellas County (USA)	Helsinki (Finland)
Implementation	implemented	pilot, implemented	implemented
Objective	regulating private mobility service providers	cost-efficiently operating public transport services	facilitating mobility as a service through private providers
Data	sensor and geospatial data from e-scooters	monthly aggregated reports	timetable and location data from public transport services
Information	scooter usage at the trip-level	numbers of users	real-time information about public transport
Method	near real time collection of data via API	monthly reports	opening up data about public transport
Transport Mode	e-scooters	ride-hailing services	public transport

Table 6: Overview of all cases covered in section 3.4, according to the framework

	Traffic Overload			
	Traffic Prediction Cranberry Township (USA)	Smart Curbside Parking Shanghai (China)	Smart Traffic Lights Copenhagen (Denmark)	Shared Autonomous Vehicles Austin (USA)
Implem-entation	unclear	pilot, implemented	implemented	conceptual
Objective	optimise traffic flows	optimise traffic flows	reduce number of vehicles	reduce number of vehicles
Data	crowdsourced data feeds, traffic sensors and weather reports	sensor data from magnetic vehicle detectors	sensor data from cameras	coded roadway network, trip tables, household diaries
Infor-mation	traffic delay prediction up to 30min in advance	curbside parking patterns	commuting mode of objects on street	impact of autonomous vehicles coupled with ride sharing
Method	artificial recurrent neural networks	IoT as technology	image recognition	agent-based modeling
Transport Mode	motorised vehicles	private cars	multi-modal	autonomous cars

Table 7: Overview of all cases covered in section 3.5, according to the framework

Declaration of Authorship

We hereby certify that the research work we are submitting is entirely our own original work, except where otherwise indicated. We are aware of the University of Leuphana's regulations concerning plagiarism, including those regulations concerning disciplinary actions that may result from plagiarism. Any use of the works of any other author, in any form, is properly acknowledged at their point of use.

All sections were written and, in case of the result sections, researched as indicated in the following table:

Section	Author
1	Eva Tesch
2	Lukas Schmid
3.1	Lukas Groeninger
3.2	Hsin-Yu Ku
3.3	Shaurya Dev Singh
3.4	Lukas Schmid
3.5	Eva Tesch
4	Lukas Groeninger
5	Shaurya Dev Singh
6	Hsin-Yu Ku

Lukas Groeninger ____________________

Hsin-Yu Ku ____________________

Lukas Schmid ____________________

Shaurya Dev Singh ____________________

Eva Tesch ____________________

(Lüneburg, October 16, 2020)