

# Public sentiment analysis and its explanatory power

Applications to new mobility planning  
and orchestration



# Contents

Research team	3
Executive summary	5
<b>1 Introduction</b>	<b>10</b>
1.1 Research basis	10
1.2 Research objectives, questions and approach	11
<b>2 An evolving mobility services landscape</b>	<b>15</b>
2.1 Mobility services and trip marketplace basics	15
2.2 Physical service impacts on the trip market	15
<b>3 Case study definition</b>	<b>19</b>
3.1 Overview	19
3.2 Greater Perth, Western Australia	20
3.3 King County, WA	24
3.4 Comparative diagnostics	28
<b>4 The quest to know more</b>	<b>30</b>
4.1 The data drive	30
4.2 Traditional travel and mobility assessment methods	31
4.3 New enablement/ data potential	33
4.4 Current regulatory environment – data access	35
4.5 Data protocols and standardisation	39
4.6 Potential limitations to data-driven insights	40
<b>5 The potential value of sentiment analyses in transport research</b>	<b>41</b>
<b>6 Findings and discussion</b>	<b>45</b>
6.1 Sample statistics	45
6.2 Results and discussion	47
<b>7 Key insights and conclusions</b>	<b>54</b>
List of acronyms	57
References	58
Appendix A	61
Summary of expert interviews	



# Research team

Arup is drawn to projects involving complex, atypical problems requiring non-standard solutions. We are uniquely positioned to serve a mix of public, private and industry clients with large stakes in the evolving mobility paradigm. Arup conducts internally funded research and produces thought-pieces regarding the future of transit, highways, rail and airports.

This research — along with targeted studies of impacts of driverless cars on urban streets and policy implications of driverless cars and the mobility-as-a-service ecosystem — demonstrates our commitment to exploring the outcomes of new and emerging mobility tools and services.

Oniracom is a data driven creative agency operating in Santa Barbara, CA. We sharpened our tusks at the crossroads of music and philanthropy, starting in 2001 as the agency of record for a then little-known artist named Jack Johnson. He aimed to use his talent to benefit environmental and humanitarian concerns. We immersed ourselves in supporting this dream. Connecting with fans, while taking full advantage of the social, streaming, apps, mobile and digital world exploding around us, we built a cause-driven global community. Jack's dream was realised: millions raised for compelling causes through sold-out stadiums, merchandise sales and non-profit donations. Today we back our creative marketing initiatives with Actionable Intelligence™ which uncovers insights for organisations and brands through a suite of social listening, data mining and analysis tools and human processes. We then apply recommended strategies and tactics to activate our findings.

## ARUP



### Chief Investigators

**Ryan Falconer** (primary author) holds a PhD in sustainability and technology policy. As former Cities Leader Western Australia and Integrated Planning Leader Canada for Arup, he has been involved with a broad range of sustainable cities and changing mobility projects and research. He has spoken to international audiences in New Zealand, Australia, Sweden, Hong Kong, China and North America on topics including mobility-as-a-service, automated vehicles and transit-oriented development.

**Jacob Tell** has expertise in positioning premiere entertainment, lifestyle and consumer brands through high-level marketing strategy. With an educational background in film and politics, Jacob seeks world-changing solutions for positive and creative people. He embraces technology-driven solutions and aims to make them widely accessible. With an appreciation for nature and hope for the future of human ecology, Jacob supports forward-thinking environmental and political leaders, and aims to further meld artistry and science.



Ryan Falconer



Jacob Tell

## Co-Investigators

**Dylan Melsom** is a rising Transport Engineer in Arup's Transport Planning team in Belfast, Northern Ireland. He has spent the early years of his career developing a technical specialism in emerging forms of mobility, receiving multiple awards for his research into the implications of autonomous vehicles for cities and urban form. Dylan's research has led to a number of major industry speaking and presentation engagements across Australia and has helped inform long-term government planning around future mobility. He is passionate about improving and future-proofing quality of life through the sustainable, intelligent enabling of mobility.

**Terry Lee-Williams** is Arup's Principal of Transport Strategy for Australasia with more than 25 years' experience in integrating technology, people-focused strategy and services design and delivery across Australia and Europe. He aims to dig beyond hype and ensure that real problems are solved by technology and data-derived insights. Terry has been project director of software and hardware systems for transport that continue to service more than 70 cities worldwide.

**Sean Campos** is Oniracom's Chief of Technology and has more than 15 years of experience in the music industry creating e-commerce, ticketing and marketing platforms for artists, labels and festivals including Disney Records, Warner Brothers Music and Universal Music Group. More recently, Sean has been working with UCSB's Media Arts and Technology program on machine learning applications and applying these emerging big data methods and technologies to create intelligence products for retail brand groups to better understand how their audiences perceive and respond to brands in both the physical and digital world.

**Mike Wald** is Oniracom's Chief Marketing Officer with over 20 years' experience developing campaigns and launching products for a wide range of verticals. His expertise and affection in ecological anthropology brings an innovative and thoughtful blend of science-meets-tech-meets-sociology to Oniracom's work. Mike's background in driving successful conversions across mixed media ensures effective solutions through a wide array of industries and platforms for nimble, first-mover companies, mass-focused businesses and international governments. Mike's work can be seen everywhere from Disney movie campaigns to government infrastructure dashboards in Australia.

### Declaration

This report represents the sole opinion of the authors and is not to be construed as representing directly the views held by the organisations listed. No interviewees or the organisations they work for exercised editorial control over the report.

### Privacy Statement

<https://www.arup.com/-/media/arup/files/publications/p/privacy-statement.pdf>

### Other credits

The authors would like to thank **Melissa Ruhl** and **Martha Koch**, Transport Planners in Arup's San Francisco office, and **David Beal**, Transport Planner in Arup's Seattle office, for their excellent technical support. Thanks to **Karen Scarborough** for her formatting and graphical support, and **Jennifer Shand** for her thoughtful editorial reviews. Lastly, we thank our interviewees, including representatives of the State Department of Transport and Transperth in Perth, Western Australia and the Department of Transportation in Seattle, Washington State, as well as mobility service and platform vendors in both geographies.



# Executive summary

## Approach

The White Paper is second in series of following the March 2018 publication of *Mobility-as-a-Service: The value proposition for the public and our urban systems*, co-authored by Arup and MaRS Discovery District in Toronto. It is also a partner document to research conducted by Arup associated with *demand-responsive transit* and *MaaS governance and orchestration*.

The research objectives addressed in the current White Paper include:

1. Enriching our understanding of the impacts of new mobility business models on passenger travel including why people choose to use services provided by Transport Service Providers (TSP) such as Uber, Lyft and Ola. This research focuses particularly on ride-hailing and its relationship with public transport
2. Appraising the opportunities and limitations associated with traditional and emerging data collection, and analytical methods and methodologies to yield these insights
3. Ascertaining whether new technologies - some of which are fuelling these new mobility business models - also provide the opportunity to address gaps in traditional and alternative data collection methods and methodologies allowing more robust interrogation of the travel market. We are interested especially in the potential of sentiment analysis in this regard.

Sentiment analysis is a data-driven technique becoming more common globally to assess with currency and over time, consumer experiences and appraisals of variables of interest. The technique gives insight into the ideas, opinions, wants, needs and concerns of members of the public relative to topics of interest. These personal expressions may be understood spatially and temporally by analysing the rich attributes of social data.

There is significant potential for sentiment analysis to provide insights to complex issues contemplated by transport planners and policy-makers but not well understood or explained.

These added insights are needed for three significant reasons. Firstly, disruption in the mobility sector including diversification of service offerings and vendors, and new business models are changing fundamentally supply and demand for travel.

Secondly, traditional data collection and analytical methodologies feature various biases, and tend to be limited by cost, inaccessibility and other variables. They therefore cannot provide sufficient explanatory power regarding the scale, nature and specific trajectory of disruption.

Thirdly, despite huge volumes of travel data being generated and retained by new mobility service providers, there are significant limitations to how accessible and applicable this information is. Specifically, pan-geographic data standards and handover requirements are lacking.

The value proposition of sentiment analysis as a data collection and analytical methodology includes source data being:

- Accessible
- Provided in real-time
- Dynamic, allowing changes and trends to be observed
- Expressive – relatable to travel experience; especially the how (mode) and why (purpose)
- Independent of any research framework – data volunteered by travellers, not elicited through any targeted questioning or survey
- Potentially voluminous and originating from a wide audience, adding to its explanatory power at population levels.

The empirical components of our research project involved use of sentiment analysis to compare and contrast trip marketplaces in Greater Perth, Western Australia and King County (including City of Seattle), Washington State, focusing on public transport and ride-hailing. Social media data generated between May 2017 and May 2018 was analysed to generate the specified insights. Key features of these marketplaces include:

- Until early 2018, when Ola entered the market, Uber was the sole ride-hailing vendor in Greater Perth. A third, much smaller vendor, Shofer operates in Greater Perth now as well. Otherwise, the metropolitan area has few other mobility services such as car-share or bike-share and weak strategic policy in relation to new mobility. Furthermore, despite increases in public transport service kilometres during the 2010s, ridership since 2012 has declined across the system
- King County is bucking the trend across the US with respect for public transport *ridership*, displaying annual increases following ambitious investment in infrastructure and services. This seems to be despite the diverse and growing local mobility services sector. In contrast to government in Western Australia, the City Government in Seattle has a clear *New Mobility Playbook* proposing principles and strategies to leverage positive change in the sector.



The steps employed to undertake the sentiment analysis included:

1. Finding relevant conversations by building queries to search through social media archives over the period of interest
2. Geo-filtering conversations to make sure content was relevant to the specified case study areas
3. Organising conversations by author, timestamp and location
4. Assigning polarity (positive, negative or neutral) using Bayesian sentiment analysis
5. Assigning emotion using a proprietary Bayesian approach
6. Creating topics according to the most frequently occurring keywords within the closest proximity using the Latent Dirichlet Allocation (LDA) algorithm. LDA is a mathematical method applicable to written content resembling how unsupervised clustering works on numerical data
7. Creating graphs to identify over time specific subsets of the data worth more detailed hand analysis
8. Analysing a subset of results manually, based on completion of the previous steps in the approach enabling formation of a more complete picture of both the sentiment and emotion expressed in social media mentions
9. Drawing conclusion about mode use and appraisal trends over time through rigorous dissection of what people have said, why they have said it and when they said it.

## Findings

Collectively, the data show a range of clear commonalities and differences in sentiment (and the variables it is attached to), between Greater Perth and King County. Commonalities include:

- A lot of mode choice is habitual. Variability is associated much more with disruptive events and discretionary/ atypical trips
- The modes appear to satisfy different segments of the trip marketplace but there is overlap. The overlap appears to be particular to individuals' balancing of time-money equations
- Customer service and satisfaction is a big influence on social media postings (including their polarity). Over time, this may lead consumers to alter their travel habits
- In broad terms, ride-hailing companies appear to be eliciting worse sentiment tied to some of their customer service practices.

For ride-hailing, the relatively strong negative sentiment pertaining to customer service and experience may herald a tapering of growth in the sector, particularly if public transport levels of service are improved further. This is a trend that should be watched closely.

Still, while the data is informative for transport policy, it provides limited insights into public transport usage trends in Greater Perth and King County. In Greater Perth, the State Government is performing well with respect for core services: where they are provided frequently and reliably, consumers are satisfied. Furthermore, customer service is rated very highly.



Our research project involved the use of sentiment analysis to compare and contrast trip marketplaces in Greater Perth, Western Australia and King County (including City of Seattle), WA.

In King County, there is a delta between negative and positive sentiment relating to public transport. Reasons for this are not particularly clear from our dataset but may relate to some perceptions of customer service as well as levels of service that require further improvement including time reliability via some routes.

## Key insights

### **Mode choice is a habit.**

The conversations held on social media in both Greater Perth and King County indicate that mode choice tends to reflect habits and has little dynamism. Significant disruption - which may include repeated negative experiences of one mode - is required generally to alter behaviour. These findings accord with the general transport literature.

### **Time and money trade-offs drive mode choice.**

People typically make a trade-off between time and money when making travel choices. Ride-sharing tends to be preferred for discretionary trips and if timing is a critical variable (e.g. someone is running late or faces an imperative to arrive at their destination at a specific time – especially if the trip is atypical). These findings are consistent with other research that shows some segmentation of the trip market in favour of public transport and some, ride-hailing.

### **Social experiences often dictate sentiment.**

Travel by either public transport or ride-hailing involves some level of social interaction. The particulars of these interactions correlate strongly with whether expressed sentiments are positive or negative. The largest volume of negative sentiment in both Greater Perth and King County was related to poor customer service. Conversely, the largest volume of positive sentiment in both locations related to good customer service and positive social experiences while travelling.

### **Public transport's affordability overcomes some other barriers to use.**

Public transport presents new and occasional riders with some barriers to use including the need to learn or interpret routes and schedules. This tends to be more acute when trip demands need to be satisfied by lower levels of scheduled service. As a more affordable option, in many cases, people address these barriers in lieu of paying more to utilise ride-hailing, unless the trip is discretionary or atypical. In such cases, ride-hailing becomes more appealing. These sorts of insights help demonstrate the appeal to some public transport agencies of select partnerships with TSP.

### **Sentiment reflects level of service.**

Public and private mobility service vendors need to appreciate that timeliness, convenience and reliability are all factors influencing perception of service and sense of satisfaction. Maintaining and enhancing service standards should be part of core public sector policy (and generally is, in both Greater Perth and King County).

### **Sentiment can help explain some of the current transport trends in Greater Perth and King County.**

The study highlighted the importance of time-budget trade-offs and emphases on levels of service. The various investments in public transport in King County, including in fixed and ancillary services are likely to have improved the balancing equation in favour of public transport.

We identified sets of strengths and potential applications, as well as limitations associated with sentiment analysis following our prototyping exercise. Strengths and potential applications within the transport policy and planning sphere include:

- Certain qualitative datasets can be compiled at-scale, which helps identify trends, risks and opportunities
- Change-over-time can be understood better; particularly in terms of how the travelling public reacts to changes in levels of service. The more disruptive these events, the more likely they are to be captured comprehensively by the technique
- Cross-sectional comparisons are enabled of differences in transport trends and how these are appraised by the travelling public
- In time, methodological refinement (including growing sophistication of the search and analytics algorithms as well as user expertise), will add to the applicability of the technique. Furthermore, a growing dataset will permit better inferences to be drawn as well as increasing predictive capability (especially in advance of known or likely events), to help shape responses and management practices
- Triangulating findings from other methodologies; particularly those unable to shed significant light on the decision-making of the travelling public, and their appraisal of services offered based on factors such as levels of service, time-of-day and location

- Avoiding introduction of surveyor/ framing and some recall bias into the data collection procedure and enabling an iterative research process, rather than analytics and insights being constrained by a static and/ or single-point survey. Furthermore, processing can be rerun using adjusted parameters, if on review, the researcher (or a peer-review) considers that the algorithm has introduced bias
- Facilitating capture of relevant feedback from an increasingly broad sample of the travelling public and deep-and-wide dataset, leading to greater internal consistency of sentiment (e.g. omission of outlying responses from commentators who make multiple posts, to appraise stronger trends).

## Limitations

- Commentators generally share content relating to more extreme or severe experiences (especially negative experiences), which skews the data. Nevertheless, transport planners and policy makers, again with a large dataset, can observe changes over time and the specific polarity of content to gain insights into factors such as satisfaction and potential behaviour change. These insights can be compared-and-contrasted with both other datasets and similar data from other contexts
- People are more inclined to comment when change occurs. Sentiment analysis is therefore better suited to measuring reactions to change than for day-to-day or habitual decisions. The potential power of the technique to gauge and in time, predict the impacts of changes, may well have merit
- Long-run data is needed to gauge change to habits: some negative sentiment may simply illustrate common frustration
- Sentiment cannot be used reliably to *quantify* mode use and the typical lag between experience and posting makes geo-location difficult. The technique therefore offers more as a means to understand the why, what and when of travel behaviour but much less so regarding where (with much geographic specificity) and how much. While there is potential to apply this technique to help answer questions about where services could be improved, this is more likely to be possible at the mesoscopic level
- Rhetorical devices and complex language (e.g. sarcasm/ irony/ memes) are very difficult to analyse correctly applying existing algorithms. Although this difficulty applies also to traditional techniques, oral surveys and focus groups do allow researchers to clarify meaning directly with the participant. This limitation on sentiment analysis is mitigated through manual analysis, which can become labour-intensive and requires the analyst to have sufficient domain expertise
- Some framing, researcher and recall biases can be avoided but analyst and processing biases remain. While these may be mitigated over time as the technique becomes more sophisticated, the technology and the analyst are required to contextualise sentiment. Such limitations are mitigated further through aggregation of more data, heuristic analyses and triangulation with other information. It is possible that reliability and validity limitations are ameliorated over time following sufficient cycles and machine learning
- Those without the internet or social media are not captured and personality types more inclined to post online are captured disproportionately compared to those that do not post as often
- ‘One mention’ does not constitute consistency in mode choices or experience.

Additionally, the methodology can enable triangulation of Personally Identifiable Information (PII), such as socio-demographic and contact details, even if all data is from public sources. This is something requiring appropriate management and control, and undoubtedly, civil and criminal law will continue to evolve to address this. The public and private sectors must avail themselves of statutory (and ethical) boundaries.









# 1 Introduction

## 1.1 Research basis

In 2017 and 2018, Arup conducted research with MaRS Discovery District in Toronto, Canada exploring the evolution and characteristics of the mobility services market. The resulting White Paper, *Mobility-as-a-Service: The value proposition for the public and our urban system*, provided insights into the impacts and opportunities associated with new mobility business models and ride-hailing.

The research was timely, joining an expanding body of knowledge relating to mobility services. Urgency and sensationalism pervade the mobility industry and policy-makers' thinking, given the fairly recent (post-2014) explosion of service vendors and service variants, in trips served, and of digital platforms providing journey planning, product access and payment methodology in a single application.

The full suite of primary and secondary (collateral) effects of this disruption are far from clear leading to a major quest for more data: data that sheds greater light on uptake of new mobility services (subject to consumer budgets, preferences, willingness-to-pay, trade-offs, pain-points, social circumstances etc.) and its effects (demand for services and service types, residual use of other modes, demand for infrastructure and impacts of these demands on variables like network performance, consequences for public health etc.) [1].

The quest for these sorts of data echoes the entrenched and mainstream pursuit of the transport planner and policy-maker to understand why, when, how and where passengers and freight travel. Such understandings are fundamental to policy and investment decisions relating to transport infrastructure and services, as well as land use policy, social services provisions and many other aspects of public governance. These understandings also provide the business case for all manner of private sector enterprise: enterprise relating directly to people and goods movement, and enterprises reliant on them.

Traditionally, transport and travel data has been either quantitative or qualitative. Put simply, the former provides statistical validity and reliability for wide datasets, and explanatory power for populations of interest while the latter, rich and deep insights into complex topics across much smaller samples.

Data collection methods have been limited by variables including:

- **Access** to the population of interest and pertinent feedback
- Availability of **labour** to execute the exercise
- **Budget** available to conduct the research
- **Currency** of data (e.g. how often is data collected and how long until it becomes unreliable?)
- **Timeliness** of feedback (e.g. reliance on a subject's recall ability and/ or accuracy of predictions regarding how they may travel in future)
- **Processing** capability once data is collected.

While adoption of mixed-methods approaches can ameliorate these limitations, industry strives continuously to learn more about transport and travel behaviour and seek better ways to enable this learning. Qualitative insights from large and powerful sample sizes have always been difficult to decipher and expensive to gain. Some contemporary, sophisticated products, such as *Mobility Mosaic* can address some of these challenges, combining quantitative and qualitative data at-scale.

For quantitative data, deployment of Internet-of-Things (IoT) devices is now pervasive. There is potential for access to datasets that are far greater than required for statistical significance. Still, these datasets have highly variable value depending on the circumstances of collection and often have gaps due to collection methodologies that make reliability a real issue.

By 2020, six billion smart phones will be in circulation worldwide according to IHS Markit [2]. As of 2018, about 3.2 billion people are users of social media [3]. The current global population is around 7.7 billion people, meaning one smart phone for four in every five of us and that one in every 2.5 tweet, post on Facebook or have used some other digital social platform [4].

This scale of digital consumption is impressive, given the great majority of growth has occurred since 2007 and 2008, after the launch of the original iPhone and Android operating platform [5]. In cities in developed economies this penetration is even greater, accounting for a significant majority of the travelling public.

GOVERNMENT  
Greater  
Perth

A challenge and an opportunity is the mass collection and analysis of valuable data. It is essential for successful operation in a public-private ecosystem.



A simple review of any common social media platform reveals commentary relating to any number of different topics; transport-related topics included. These cloud-based data may provide some of the insights the transport industry is looking for. They might provide the level of insight required with the density needed for reliability when used as evidence for decision-making.

Access to this data and sophisticated processing of it to provide insights into transport decision-making and its impacts, may address limitations manifest in the pursuit of insights through other digital means. For example, many governments understand the power of datasets held by various service vendors -telecommunications companies, Global Positioning System- (GPS) based service providers (e.g. TomTom), credit card companies as well as Transport Service Providers (TSP)<sup>1</sup>. But, often, many barriers to accessing and using this data apply including privacy limitations, and data collection by and retention within proprietary systems (price and/or data formatting barriers). These limitations and barriers are explored further in **Section 4**.

## 1.2 Research objectives, questions and approach

Our research focuses on two trip marketplaces - Greater Perth, Western Australia<sup>2</sup> and King County (including City of Seattle), Washington State (WA) - as case studies. The project team was intrigued by these geographies because:

- Until early 2018, when Ola entered the market, Uber was the sole ride-hailing vendor in Greater Perth. A third, much smaller vendor, Shofer operates in Greater Perth now as well. Otherwise, the metropolitan area has few other mobility services such as car-share or bike-share and weak strategic policy in relation to new mobility. Furthermore, despite increases in public transport service kilometres during the 2010s, ridership since 2012 has declined across the system. Other researchers have sought to explain the trend but with limited success and evidence [6]
- King County is bucking the trend across the US with respect for public transport *ridership*, showing strong, year-on-year increases off the back of ambitious investment in infrastructure and services<sup>3</sup>. This seems to be despite the diverse and growing local mobility services sector. In contrast to Government in Western Australia, the City Government in Seattle has a *New Mobility Playbook* proposing principles and strategies to leverage positive change in the sector.

---

<sup>1</sup> In this White Paper, we use TSP as a synonym for Transportation Network Companies (TNC), which is an alternative term used commonly in the mobility literature, particularly in North America.

<sup>2</sup> We refer to Western Australia in full throughout the paper rather than its common abbreviation, WA. This is to avoid confusion with Washington State (also, WA).

<sup>3</sup> Although there have been decreases in total passenger kilometres post-2015



More discussion is provided in **Section 3** regarding the selection of case study geographies as well as a comparative urban and transport system diagnostic. Data policies are discussed in **Section 4**, which are central to our research.

Our research objectives include:

1. Enriching our understanding of the impacts of new mobility business models on passenger travel including why people choose to use services provided by TSP. This research focuses particularly on ride-hailing and its relationship with public transport
2. Appraising the opportunities and limitations associated with traditional and emerging data collection, and analytical methods and methodologies to yield these insights
3. Ascertaining whether new technologies - some of which are fuelling these new mobility business models - also provide the opportunity to address gaps in traditional and alternative data collection methods and methodologies allowing more robust interrogation of the travel market. We are interested especially in the potential of *sentiment analysis* in this regard.

Fulfillment of objectives 1 and 2 provided a theoretical basis for the main, empirical thrust of our work (satisfaction of objective 3). For this, our research team designed and prototyped a study of public transport and private sector-provided mobility service use across the two case study trip marketplaces, using sentiment analysis.

A comprehensive description of sentiment analysis and its employment in this study is provided in **Section 5**. Briefly, sentiment analysis is a means of leveraging self-reported mobility information published through social platforms by applying sophisticated processing algorithms and technical knowledge to frame and answer questions of interest. The value proposition of this method includes source data being:

- Accessible
- Provided in real-time
- Dynamic, allowing changes and trends to be observed
- Expressive – relatable to travel experience; especially the how (mode) and why (purpose)
- Independent of any research framework – data volunteered by travellers, not elicited through any targeted questioning or survey

- Potentially voluminous and originating from a wide audience, adding to its explanatory power at population levels.

The remainder of this White Paper is structured as follows:

## **Section 2 – An Evolving Mobility Services Landscape**

What are mobility services? What does early research and analyses conclude may be some of the impacts of new and changing mobility services on the mobility ecosystem, people's travel predilections and the urban environment? What precipitates the current drive for transport planners and policy-makers to learn more about these effects?

## **Section 3 – Case Study Definition**

What are some of the main urban and transport characteristics of Greater Perth compared to King County with particular focus on public transport and private mobility services? What makes these two geographies of interest to study further?

## **Section 4 – The Quest to Know More**

What tend to be the gaps in knowledge owing to the benefits and limitations associated with traditional methods of transport data collection and analysis? What alternative methods are being enabled by technology and innovation? How does the current regulatory and business landscape facilitate or inhibit access to meaningful data? What may be the limits to what data and analytical tools can tell us?

## **Section 5 – The Potential Value of Sentiment Analyses in Transport Research**

What is the 'promise' associated with sentiment analysis? What gaps in knowledge may it fill? How is the technique employed?

## **Section 6 – Findings and Discussion**

What evidence does sentiment analysis yield with respect for travel choice, experiences of public transport and ride-hailing, and how dynamic mode choice may be? How powerful are these insights and what benefits may they provide transport planners and policy-makers?

## **Section 7 – Key Insights and Conclusions**

What insights have our research and the prototyping of sentiment analysis as a transport assessment method provided? What appear to be potential future applications of this method and further work to be completed?



Seattle congestion on 2nd Ave

© Oran Viriyincy







## Research approach

Conducting a research review focusing on associations between mobility services and public transport usage, differentiating between trip types wherever possible, and analysing any travel decision-making explanations provided by the travelling public (focusing especially on reasons relating to quality of the public transport system in their respective geographies)

Undertaking a cross-sectional review of key urban and mobility-related data for our two case study cities as part of developing a comparative diagnostic of 'health' of the mobility eco-system

Conducting in-depth stakeholder interviews (see below) soliciting views regarding the quality and coverage of public transport in the study cities, institutional knowledge of customer satisfaction and usage patterns, and attitudes towards mobility services (including existing or potential partnerships)

Completing a review of traditional and emerging data collection methodologies to understand better opportunities for, and limitations to, gaining insights into travel behaviour

Analysing public policy relating to TSP data accessibility and thereafter, reviewing TSP data acquired through requests according to current public data policy in each location (Greater Perth, Western Australia and King County WA)

Designing and implementing the sentiment analysis component of the research structured around our key research questions

Comparing our various datasets between Greater Perth and King County and analysing it for correlations between customer sentiment and public transport system quality variables

Developing insights based on our work according to the research objectives and questions set.

## Expert interviews

Seven senior industry experts from a range of backgrounds were interviewed in semi-structured interviews as part of this study including senior government officers in Greater Perth and King County responsible for future/ new mobility strategy and planning, and service development for public transport (Perth). Senior representatives from two TSP were interviewed as well. The purpose of the interviews was to add context to the research. Interviewees were asked the following questions:

What are today's major mobility/ transportation challenges and opportunities?

How does your agency/ company view the role of transportation in improving quality of life?

What are the relative value propositions of mobility services and transit (e.g. convenience, price, comfort)?

What data sources provide insights into ride-hailing and transit use, and what are their limitations?

What do you think are the most important things to improve with regard to user experience?

What are barriers to achieving better user experience right now?

How can industry build a better collective knowledge base regarding mode choice?

Responses provided by the interviewees informed the remainder of our research and we have included a series of responses throughout the body of the document. A more complete summary of the interviews is provided in **Appendix A**.

## 2 An evolving mobility services landscape

### 2.1 Mobility services and trip marketplace basics

Contemporary mobility services include both the physical transport service provided to a paying customer (such as a trip via public transport, taxi or ride-hailing, or a combination of modes), and the digital platform (typically in the form of an app) that enables journey planning, selection of service(s) and payment. Our research is concerned with both the physical and platform components, given the physical service impacts on the mobility system and the platform generates (or has the potential to generate) data relating to the travel undertaken.

While there are many and evolving physical service variants - ride-hailing, ride-sharing, car-sharing, scooter-sharing, bicycle-sharing and so on - we focus on ride-hailing, given it is the mobility offer receiving most attention in the literature with respect for its relationship with and impacts on public transport use. This is not to the exclusion of ride-sharing of the type available, as an option, through various vendor platforms (e.g. Lyft Line, UberPOOL or ViaVan).



#### Mobility Service Vendor

MaaS enables the consumer to make individual economic choices on a trip-by-trip basis rather than in long-term periods, e.g. buying a car.

We define the trip marketplace in any given geography to be the demand from consumers for physical travel. The size and make-up of the marketplace varies depending on factors such as:

- Time
- Day
- Location
- Changing socio-economic and socio-demographics
- Mobility options available in the marketplace
- Mobility price-points.

The entry and/ or exit of mobility service vendors into a marketplace can affect demand for travel and the physical travel demanded across different mobility options. There is potential also for demand to remain unfulfilled because of limitations to service offers including geographic or temporal coverage, and trip reliability, quality and cost.

### 2.2 Physical service impacts on the trip market

A growing body of literature discusses the impacts of and opportunities associated with availability of and growth in ride-hailing services in different locations. These findings are based, in many cases, on limited and short-run datasets:

- UberX and other app-based, on-demand, point-to-point mobility services have only been available for around half a decade
- New service vendors and the business models they employ have uncapped supply of mobility services compared to traditional licensing controls on the taxi industry. The same lack of licensing structure has limited the volumes of specific trip data available to government and industry to assess impacts.

Over this very short time, there has been massive growth in use of contemporary ride-hailing to fulfil travel demand and various efforts to assess what types of trips are being satisfied by new service vendors and service offerings. In August 2014, after about half a year of operations, UberX in Australia was estimated to satisfy about 6% of point-to-point trip demand [7]. The same research concluded that about 61% of these trips were new (induced), or reassigned from private cars, public transport or active transport [8]. These estimates relied on data supplied by a vendor (Uber).

About 64% were estimated to be trips originating more than 800 metres away from frequent public transport, leading the authors to conclude that ride-hailing services address segments of the trip market difficult to serve by traditional public transport, yielding a strong consumer value proposition and potential value to governments [9]. This reflects research conducted at the Chaddick Institute at De Paul University in Illinois, which proposed that mobility services (typically, ride-hailing), provided by TSP and public transport services tend to cater to different travel needs making them relatively complementary rather than competitive [10].



Mobility  
Service  
Vendor

Ride-sharing can provide an affordable, reliable and safe alternative travel option which solves the last-mile problem when people choose to drive if any part of their trip is outside walking distance.



Mobility  
Service  
Vendor

In transit-rich areas, TSP value proposition is providing service where transit isn't operating at optimal levels, e.g. BART in the San Francisco Bay Area, where customers buy a ticket, move downstairs to the platform, and have already spent their money without knowing when the next train will arrive since there's no sign outside telling you when the next one is.

The potential for operational synergies is driving a lot of government interest in public-private partnerships to satisfy a range of travel demands including late-night, special-needs, first/ last mile and suburban-suburban trips. Historically, these demands have been costly for the public sector to satisfy because of the low passenger compared to service kilometres. The 2018 report, [Partners in Transit](#) prepared by the Chaddick Institute evaluates 29 pilot private-public partnership schemes of varying types across continental US that represent attempts to address a range of these challenging scenarios [11].

Uber now hosts public transport service data on its app in Denver, Colorado, providing access to routes and timetables in addition to comparing public transport and ride-hailing fares. Over time, Uber intends to integrate ticketing as well [12]. In Australia, the ACT Government partnered with Uber in late 2018 to provide discounted rides connecting with Light Night Rapid bus services in some parts of the Territory.

Commonly, 'transit partnerships' require government to subsidise some of the cost of the ride-hailing trip on the proviso the trip addresses a defined service gap (e.g. a last-mile trip). This lowers the price paid by the consumer, making it more affordable and appealing. The intent is that the government subsidy offsets cost savings from reducing or removing conventional public transport services that would have been provided otherwise (to provide a basic level of service), while improving convenience - the value proposition - to the consumer. The different forms of partnerships and variance on business models/ mobility services provided are covered in other literature, including the Chaddick Institute's paper.

In Ontario, a partnership between the Town of Innisfil and Uber for subsidised Uber rides to be provided in lieu of a local bus service has generated ridership well in excess of forecasts. This has led to measures to limit ride-based subsidies per person. This is contrary to typical strategies to maximise 'public transport' usage, which may indicate the limits to efficacy of a public-private mobility partnership of the type found in Innisfil [13].

Irrespective of merits associated with partnerships, TSP do not limit their service provisions to trip market segments that complement public transport. TSP exploit a trip marketplace using a business model offering convenience and portability to consumers, engaging a willing workforce and (in many cases and at least for now), exempt from traditional price barriers to vehicle-based mobility choices: expensive licensing requirements (e.g. medallions or plates), and parking costs [14, 15]. The convenience, comfort and cost offer to consumers therefore appeals for many different trip types, subject to when and where the mobility services are made available and the buying power of consumer groups.

Research conducted at the University of Kentucky concluded that for every year after ride-hailing companies enter an urban market, rail ridership may fall by about 1.3 percent, and bus ridership about 1.7 percent, placing greater strain on many systems that require already significant public investment [16].

In Boston, research by the Metropolitan Area Planning Council (MAPC) involving mobility service user surveys found that the average rail hailing trip reduces Massachusetts Bay Transportation Authority (MBTA) revenue by about 35 cents, which exceeds the legislated 20 cent surcharge per ride [17].



Across the US, Bruce Schaller predicted that by the end of 2018, the point-to-point share of the trip market would be about 4.74 billion trips, surpassing local bus services and representing a 241% increase over six years [18: p1]. TSP share of these trips increased from marginal in 2012 to about 4.2 billion in 2018. Schaller's research found about 60% of TSP users in large, dense cities would have taken public transport, active transport or not made the trip if not for TSP availability. About 40% would have used a personal vehicle or taxi [19, 20].

Access to contemporary ride-hailing services tends to generate more passenger travel Vehicle Kilometres (or Miles) Travelled (VKT/ VMT), associated with lower-occupancy trips and dead-heading, but less urban parking demand [21, 22, 23]. In dense and busy cities, the added VKT can have undesirable impacts on network operations [24], while many cities are yet to appraise how to take advantage of reduced demand for parking. In some cases, this reduced parking demand may impact on municipal bottom lines, which rely on parking revenues.







The disruption being caused within trip marketplaces is not uniform. It also manifests as a mix of opportunities and challenges for governments and the urban transport systems, and social programmes they are responsible for.

Transport and Infrastructure in South Australia used some of a Au\$10M Future Mobility Lab sandbox fund to investigate MaaS pilot opportunities. Both are proposed as programmes, similar to US mobility innovation sandbox funds, to allow mobility vendors to innovate while fulfilling agreed key performance indicators developed with the public interest in mind.

While these programmes are both important and commendable, they are not necessarily scalable, nor will they necessarily capture all the key data they could or should for comparison against key performance indicators. Moreover, as the literature is beginning to show, the externalities associated with changes within the trip marketplace may be unintended and/ or unpredictable.

Just as service vendors see the value of rich travel data and sophisticated analytics for improving their business modes, so the context is there for governments to continually retool policy in pursuit of effective mobility management. Such a truism underpins the current flurry of revision to (or creation of) data access legislation and our own investigations into young forms of data capture and analysis.

The next section of the paper overviews the case study geographies selected for more detailed analysis and applied research. The urban and transport systems diagnostics focus on core elements of each location's public transport and private sector-provided mobility services ecosystems.



Congestion and pollution are significant challenges to infrastructure capacity and liveability. TNC strives to tackle this by getting more people into fewer cars.

Our 2018 White Paper argued that much more work is needed to define how the interests of and value to service vendors, consumers and governments can be balanced through more effective management of the mobility marketplace [25]. This work requires collation and analysis of suitable data with appropriate geographic focus.

Some government policy is evolving in reflection of deepening understanding of both impacts and value creation, and applying a suitable geographic lens. This is essential, so regulation is calibrated to balance innovation and competitiveness, and public value [26].

In 2018 in Australia, Transport for New South Wales (TfNSW) held a Mobility-as-a-Service (MaaS) Innovation Challenge, selecting five propositions for piloting locally and the Department for Planning,



## 3 Case study definition

### 3.1 Overview

Perth, Western Australia and Seattle, Washington State (WA) were selected as our case study cities for several reasons:

- International relevance of the research topics and the opportunity to compare and contrast US and Australian urban experiences
- Both are second-tier population centres in their respective countries and have been subject of limited<sup>3</sup> study previously with respect for the issues described
- Perth has exhibited recent decline in public transport patronage whereas Seattle has exhibited growth
- Seattle has a relatively rich mobility services ecosystem whereas it is comparatively basic in Perth with many fewer vendors.

With respect for ride-hailing, both cities have conventional taxi services with well-established, associated regulatory frameworks. Seattle was the third city to host Uber, and Sidecar and Lyft followed. Uber launched operations in Australia in 2012. Contemporary ride-hailing was introduced to Australia (UberX operating in Melbourne and Sydney) in April 2014, sometime after Seattle, and these services become available in Perth in October of the same year [27; 28].

For the purposes of our analysis, we have analysed each city within its broader geographic context. In the case of Perth, we have analysed Greater Perth and Seattle, King County. The broader geographic resolution allows assessment of greater spatial variability in terms of mobility service and travel behaviour. Also, cities have close relationships with their immediate hinterlands and there tend to be complex movements of freight and people between the two.

▼  
Perth, Western Australia



<sup>3</sup> This is truer for Perth; however, cities including San Francisco (and the Bay Area) and New York tend to feature more prominently than Seattle in the related North American literature that we have reviewed.



Furthermore, in each case, public transport systems extend across broader areas than the cities themselves as administrative areas (for example, City of Perth is only about 20 square kilometres while Greater Perth is about 6,418 square kilometres [29]). Comparatively, Seattle is about 217 and King County, 5,479 square kilometres in land area [30]. Current populations are relatively similar (see **Sections 3.2** and **3.3**).

Clearly, alternative spatial divisions could have been selected. For example, the Perth Metropolitan Region is a defined administrative area and major statistical division according to the Australian Government. In contrast, Greater Perth, which is larger, is defined as a Capital City Statistical Area and larger than the Perth Metropolitan Region, encompassing the City of Mandurah and part of the Shire of Murray. In this case, Greater Perth was chosen because of its population and relative spatial consistency with service coverage provided by Transperth, the applicable public transport agency.

### 3.2 Greater Perth, Western Australia

Greater Perth (shown in **Figure 1**) is a relatively young urban area and immediate hinterland. With a population in 2018 of 2.04 million [31], much of its early long-term planning occurred during the middle-late 20th century when suburbanisation and car-centric transport policies were the norm. Today's built environment reflects these influences, including a low-density land use footprint and heavy dependence on private vehicles for passenger travel [32; 33; 34].

The local public transport system, operated by Transperth features five radial metropolitan rail lines, a comprehensive network of bus services and limited ferry services operating on the Swan River, the river bisecting the city. The public transport system is planned and administered by the State Government while bus services are contracted to operators. The State Government provides the rolling stock and all infrastructure, as well as conducting route planning and timetabling, and maintaining ticketing and digital media. This allows services to be coordinated and all bus drivers to wear the same uniform, despite being employed by contractors. The fare system is zonal in nature.



**Figure 1**  
Greater Perth area

*Imagery source:*  
WA Government







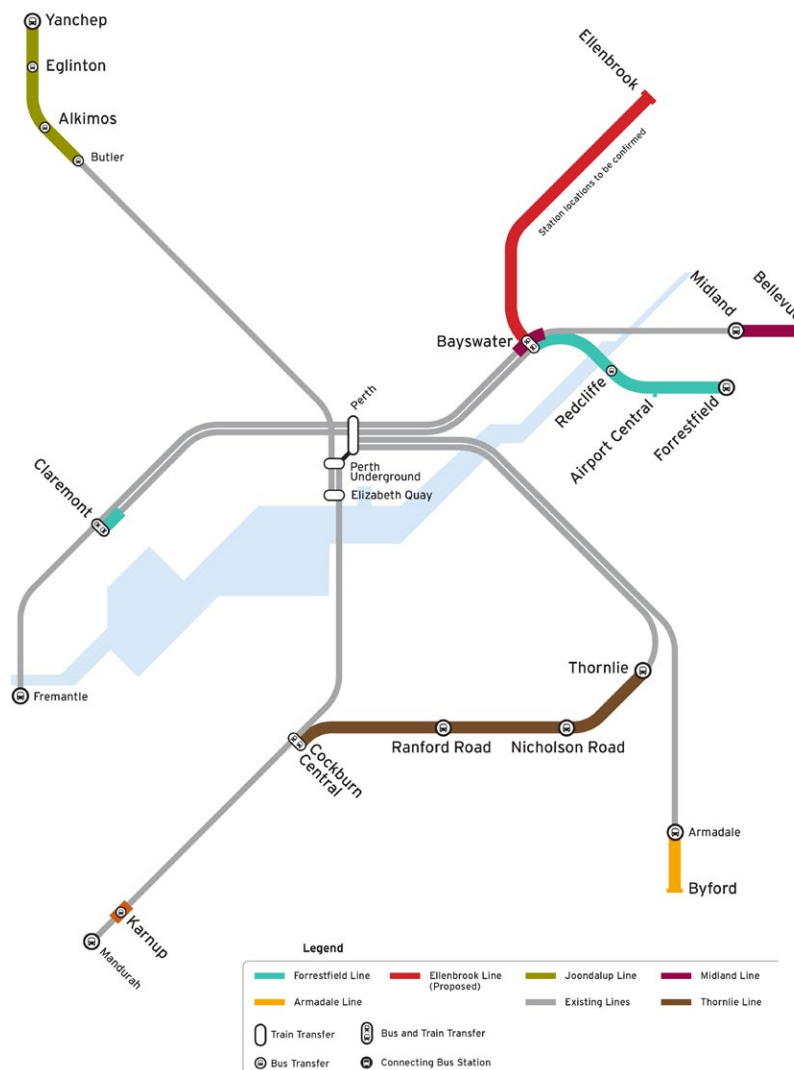


Public transport funding for services across Greater Perth has risen over recent years (including for bus service kilometres) and the current State Government has commenced an ambitious programme of extension to the rail network called METRONET (Figure 2). Between 2012-2013 and 2016-2017, service kilometres increased by 29% for train and 12% for bus. Between 2016-2017 and 2017-2018, bus service kilometres increased another 1.3% associated largely with new train station access kilometres and bus services connecting to Optus Stadium (Perth's new major events stadium), for events<sup>6</sup>.

In contrast, a 2018 study found that public transport patronage in Greater Perth has decreased since 2012-2013 despite the increased service levels [35]. The authors consider that the downturn may have been due to a range of factors including the growth of ride-hailing services in the trip marketplace as well as changes to population growth, central business district office vacancy and State economic growth rates, and parking and pricing variables. The authors were unable to demonstrate a causal relationship between any of the variables and the public transport ridership decline.



The recent drop-offs in transit have related to full-fare journeys. CBD trips are right down, which can skew results. This actually relates to a positive passenger experience, as it creates additional capacity and therefore ability to sit down.



**Figure 2**  
Proposed METRONET system including existing and proposed new rail alignments.

Source: METRONET

<sup>6</sup> Source: Transperth Annual Reports.

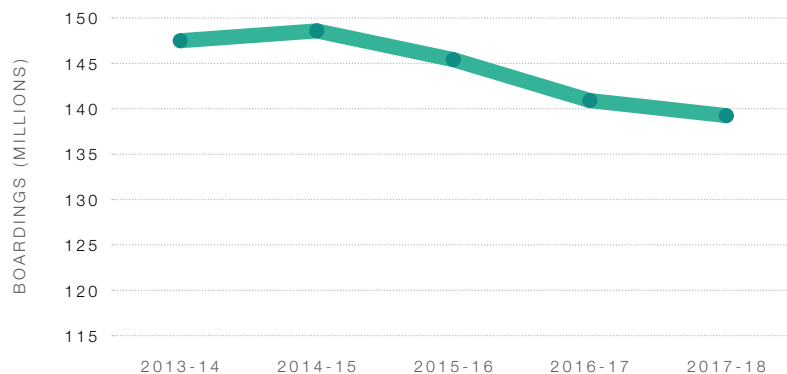
**Figure 3** shows annualised passenger boardings and **Figure 4**, annualised passenger kilometres, based on data published by the State Government. These data demonstrate the downturn, with the exception of the passenger kilometres data for 2017-18, which exhibits a significant increase (around 44%).

This uptick may be associated with the opening of Optus Stadium and the associated new train station in December 2017. A significant number of bus services operate on event days to help patrons travel to and from the stadium. Increases in boardings may have been anticipated as well but these have not occurred.

The latest passenger satisfaction data collected by the State Government shows reduced bus fare value ratings compared to the previous year attributable to a fare increase and decrease in discount offered for ticket purchases using the 'Smartrider' smartcard; however, this follows a significant increase in satisfaction from the year before. Other user satisfaction data published in the State Government's *Passenger Satisfaction Monitor 2018*, includes:

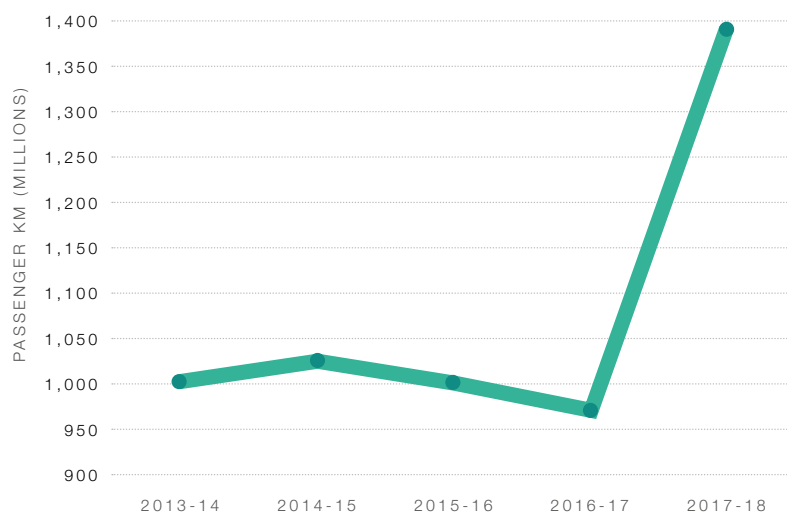
- Bus and train passenger satisfaction is at an all-time high; dissatisfaction is at an all-time low
- The two main reasons given by patrons for their likelihood to recommend Transperth services are:
  - Convenient/ easier/ less hassle/ better than driving (41% bus, 42% train)
  - Reliable – on-time/ usually on-time (34%)
- The largest factors preventing patrons from recommending Transperth services:
  - Need a more frequent service (general) (42% bus, 20% train)
  - Need more frequent service during off-peak times (33% bus)
  - Too expensive/ reduce the fares (22%)
- Peak and off-peak train patrons are satisfied equally
- Personal safety ratings are at an all-time high.

While the availability of ride-hailing may be a factor contributing to the public transport ridership downturn in Perth, the mobility marketplace has relatively little diversity in terms of private vendors and service variants. Ola and Shofer, competitors to Uber, only entered the marketplace in 2018 (Ola) and are small-scale (Shofer).



**Figure 3** Annualised public transport system passenger boardings, 2012-2018 (Greater Perth public transport system)

Source: Public Transport Authority Annual Reports



**Figure 4** Annualised public transport system passenger kilometres, 2012-2018 (Greater Perth public transport system)

Source: Public Transport Authority Annual Reports



Transperth has been monitoring passenger service for around 27 years, and it involves about 14,000 passengers. We've never been rated better.

There is no publicly-available data to understand trips made by ride-hailing<sup>7</sup> nor market share between vendors. Based on the authors' own experiences, many drivers operate for both Ola and Uber<sup>8</sup>, and jump between the two vendors depending on demand for rides.

Otherwise, Perth features no bicycle-share outside of a small pilot in Joondalup (a northern urban centre) and localised, property-related enterprises. It features no scooter-share and no car-share excepting a campus-based model at the University of Western Australia (Student Car Share).

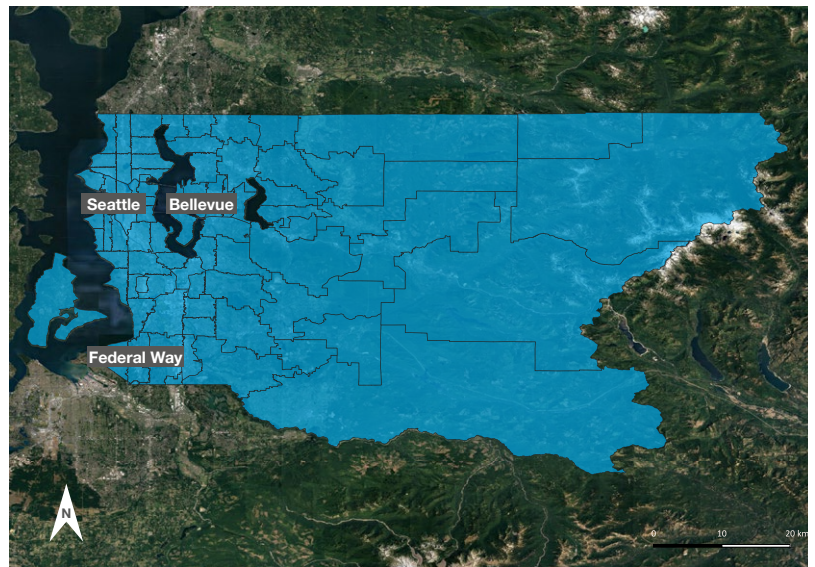
In an early 2019 study, TTF Australia and LEK appraised the 'new mobility readiness' of six of Australia's eight states and territories (Tasmania and Northern Territory were not assessed). The appraisal covered ten variables and assigned a score from one to four per variable. Western Australia's total of 16 points was lowest by some margin (Victoria [VIC] was second-lowest, with 21 points) [36]. The appraisal was also generous with respect for the sophistication of current WA State Government transport planning strategy, which includes very limited mention of new mobility trends and does not set any key performance indicators, objectives or trajectories.

TTF Australia and LEK concluded that Western Australia is behind trend across seven mobility variables including car-share, bike-share, **Mobility-as-a-Service**, on-demand mobility, driverless rail, digital driver licenses and payment innovation.

### 3.3 King County, WA

King County in WA, US (shown in **Figure 5**) is one of the fastest-growing areas of the country [37]. As of 2017, the population of King County was about 2.19 million [38]. Seattle, the major city in the County, has received significant media attention in recent years because of an upsurge in public transport ridership bucking a strong trend across the US of decline [39].

The County is served by two major public transport agencies: King County Metro, operating local and commuter buses within King County, and Sound Transit, operating commuter rail, light rail and regional express buses in the region of Puget Sound (as well as Bus Rapid Transit in the future [40]). King County Metro operates 237 bus lines carrying over



400,000 passengers per day, the sixth-busiest public bus network in the US [41]. Metro operates the City of Seattle-owned streetcar light rail system.

Seattle is home also to the iconic Seattle Center Monorail connecting downtown Seattle and the Seattle Centre. Additionally, Washington State Ferries, another government agency runs the largest ferry system in the United States [42] and the third largest in the world. The system includes 10 routes and 20 terminals located around Puget Sound and the San Juan Islands.

Between 2015 and 2016, the public transport commuter mode-share in King County increased from 12.5% to 13.1% comparing to the 2016 national average of 5.1%. In Seattle, public transport mode share was at 21% as of 2016 [43].

▲ **Figure 5**  
King County, WA

Imagery source: Google

▼ Bus services in Seattle operating in a mixed traffic environment.

© Arup



<sup>7</sup> Except the data available through Uber's *Movement* platform.

<sup>8</sup> Shofer uses specially-branded vehicles.

Figure 6 and Figure 7 show annualised public transport system boardings and passenger kilometres, respectively. Public transport boarding growth has recently begun to slow for King County Metro alongside a decline in passenger service kilometres in 2016 and 2017<sup>9</sup>.

Sound Transit and Washington State Ferries ridership has also grown over recent years. Together, in 2017 the services provided by these agencies had about 71 million boardings (46.8 million on Sound Transit and 24.2 on Washington State Ferries services) [44]. The operating areas for both agencies are much larger than King County (e.g. the Sound Transit operating area includes neighbouring Pierce and Snohomish Counties).

According to Streetsblog, “Even with low gas prices, even with population growth, even with Uber and Lyft circling 24/7, Minneapolis and Seattle have reduced the amount of driving in their cities” [45].

Growing transit ridership in Seattle is the result of significant investment in getting frequent transit to as many people as possible and in transit priority at intersections and along corridors.



Transit still wins in Seattle when it is done well – highly useable service – and invested in.

Streetsblog concludes that the uptrend in public transport use in King County (and neighbouring counties) can be explained by new vehicle levies and sales tax increment to contribute to a holistic public transport investment package including both rail and bus [46]. In 2016, the US\$54B *Sound Transit 3* initiative was approved, providing funding to expand the public transport system that when delivered, is anticipated to fuel further ridership rises [47].

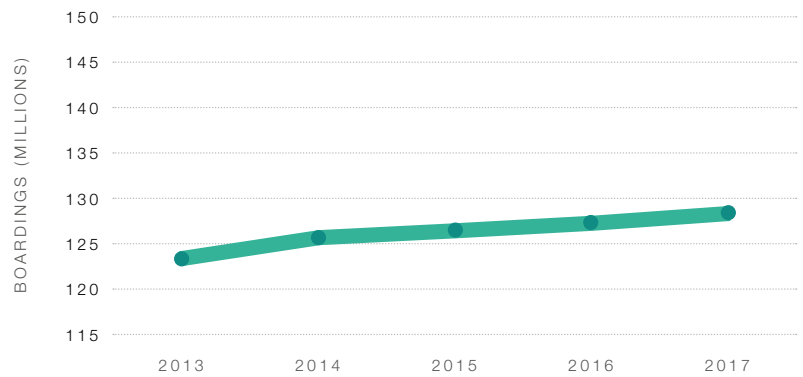


Figure 6

Annualised public transport system passenger boardings, 2012-2017 (King County Metro)

Source: *Federal Transit Administration – National Transit Database*

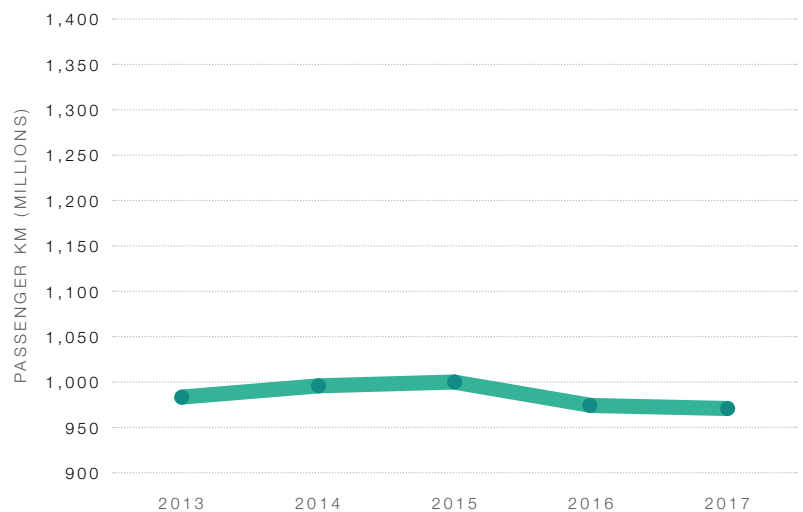


Figure 7

Annualised public transport system passenger kilometres, 2012-2017 (King County Metro)

Source: *Federal Transit Administration – National Transit Database*

Bruce Schaller observes, based on American Community Survey data, that “the best evidence that some Americans are actually ditching their cars can be seen in Seattle, where the ranks of car-free/ car-light households increased by 23 percent. That’s consistent with the city’s standing as one of very few with growing transit ridership. Yet even in Seattle, total vehicles edged up faster than population (16 percent versus 14 percent) since 2012” [48].

<sup>9</sup> King County Metro implemented “stop-based scheduling” in late 2017, which more accurately captures travel between layover and the first and last stop on a route. As a result, hours and mileage that were previously counted to revenue are now tracked as deadhead time – the full, annualised effects of which are first reflected in the 2018 reporting year.



In most cases, public transport patrons appear happy with services provided helping explain the growth in ridership. Passenger satisfaction data reported in the [King County Metro Transit 2016 Rider Survey Report](#) shows a sustained uptrend between 2013 and 2016 in the percentage of those describing themselves as 'somewhat' or 'very satisfied' (85% to 93% in the aggregate). Other headline findings of the survey (p11) included:

- Riders remain highly favourable of most variables relating to fare payment and bus operator satisfaction
- Satisfaction with information-related element is lower than in previous years
- Level of service satisfaction (including on-time performance, travel time, service frequency and availability), while also lower than 2015, has returned to 2014 levels
- All service elements have net favourability ratings, meaning far more riders were satisfied with those elements than dissatisfied.

Within the new mobility branch of SDOT and the broader mobility options and transit team, staff are searching for a way to provide more affordable and sustainable options. This includes improving the transit experience in particular, at the expense of single-occupancy drivers, so people not only have cheap options but also sustainable options.



The mobility ecosystem in King County is much more diverse than in Greater Perth. King County hosts car-sharing companies such as car2go and Zipcar, dockless bicycle-share (e.g. Ofo, Lime and Spin), ride-hailing pool options (e.g. Uber Express POOL and Lyft Line), P2P car hire (e.g. Turo) and P2P parking spot hire (e.g. Justpark). King County also hosted dynamic shuttle services provided by Chariot before the vendor's shutdown in early 2019 while in April 2019, King County Metro, Sound Transit and City of Seattle announced a mobility partnership with Via to provide improved access to regional transit stations [49]. Notably, the City Council has banned shared scooters from Seattle's streets.

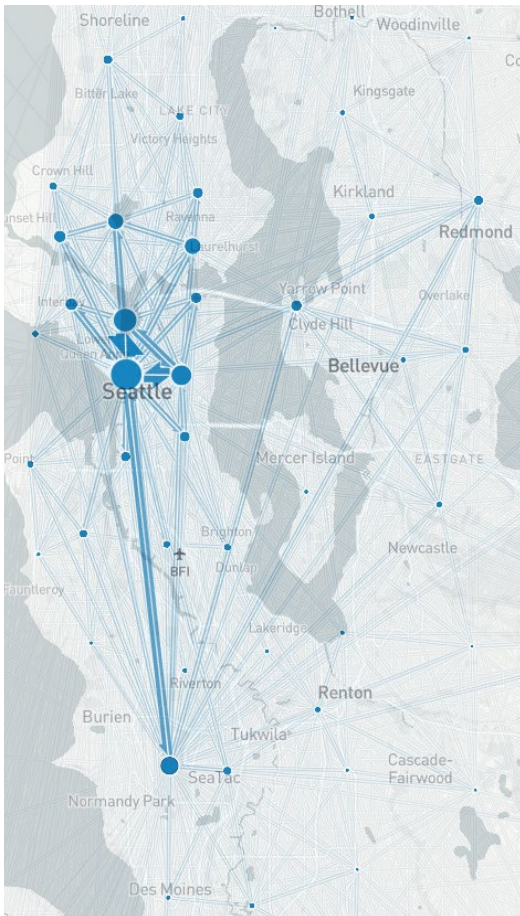
Schaller notes that in 2017, about 20 million trips were provided by TSP in the City of Seattle at a rate of about 33 per person [50]. Schaller concludes that about 94,000,000 additional miles were driven because of the presence of TSP in the marketplace (as alternatives to conventional taxis) [51].

In the second quarter of 2018, TSP provided an average of 91,000 rides per day (a large majority as single-passenger rides). While a fraction of total daily travel, trips were concentrated in downtown areas and usage has grown more than threefold since the second half of 2015. Furthermore, TSP trips total about six times the number of taxi trips when they were at their peak [52]. The high concentration of trips in the downtown contrasts with the limited and now older data available regarding ride-hailing patterns of use in Australia reported by Deloitte Access Economics.

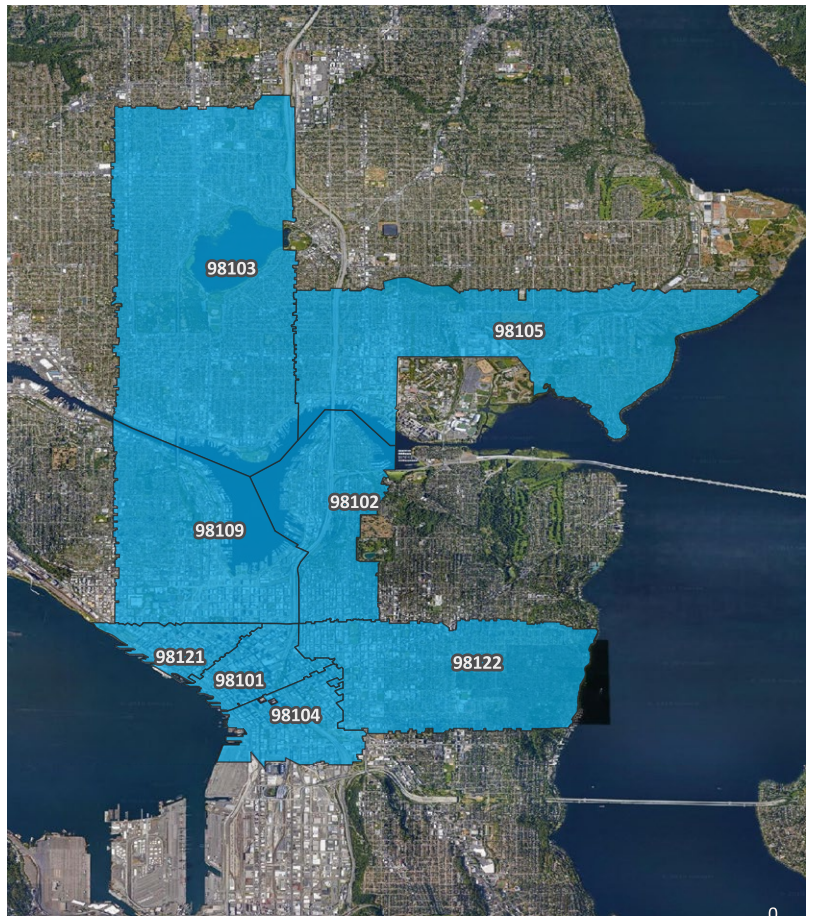
Aggregated trip data collected by Lyft for the second quarter of every year from 2015 to 2018 and obtained through a public request was analysed for trends in flows between zip codes. The number of trips increased by a factor of 15 (1,424% growth) between 2015-2016, a factor of over two (117% growth) between 2016-2017, and a factor of 1.6 between 2017-2018 (62% growth). Although the rate of growth is slowing following initial uptake, the overall growth in the number of trips continues to increase.

The data show that while tremendous growth in the total number of trips has occurred, the distribution of these trips has not changed significantly from the major flows shown in **Figure 8**. Trips are weighted heavily to origin and destination pairs including eight of King County's 85 zip codes (shown in **Figure 9**), all of which are in inner Seattle.

Government in King County is more progressive than in Greater Perth with respect to forward-looking mobility policy and strategy. In September 2017, the Seattle Department of Transportation (SDOT) published a [New Mobility Playbook](#). The Playbook sets noble goals, including that "*the transportation system in an affordable city improves the lives of all travellers – those with the latest model smart phones in their pockets and those without*" [53: p14].



▲ **Figure 8**  
 Major origin and destination pairs for Lyft trips have remained consistent between 2015 and 2018.  
 Generated using [flowmap.blue](https://flowmap.blue)



▲ **Figure 9**  
 Zip codes (origins and destinations) associated with the majority of Lyft rides, second quarter data 2015-2018  
 Source imagery: Google

The Playbook is principles- and actions- based, albeit it does not specify particular key performance indicators. Principles articulated include:

- Put people and safety first
- Design for customer dignity and happiness
- Advance race and social justice
- Forge a clean mobility future
- Keep an even playing field.

Five associated actions (or plays) include:

- Ensure new mobility delivers a fair and just transportation system for all
- Enable safer, more active and people-first uses of the public right-of-way
- Reorganise and retool SDOT to manage innovation and data
- Build new information and data infrastructure so new services can ‘plug-and-play’
- Anticipate, adapt to and leverage innovative and disruptive transportation technologies

### 3.4 Comparative diagnostics

**Table 1** compares key urban and transport system data for Greater Perth and King County. The data show that the case study areas are of reasonably similar size and have reasonably similar population numbers, although gross urban population density is noticeably higher in King County (owing to being somewhat smaller but with more people).

Both have similar private car driver commuter mode shares while other commuter mode shares differ more significantly, with King County exhibiting higher levels of public transport, active transport and car passenger. Greater Perth statistics show a much higher share for 'other'.

Based on 2017 data, public transport passenger kilometres<sup>10</sup> were similar for Transperth and King County Metro, although the latter attracted fewer boardings. When Sound Transit and Washington State Ferries boardings data are added, overall boardings associated with King County travel are higher. The King County Metro trend data shows growth in public transport boardings in King County despite a decline in passenger kilometres in recent years. In contrast, boardings on the Transperth system in Greater Perth have declined.

▼  
**Table 1**  
Comparative diagnostic:  
Greater Perth and King  
County

	Greater Perth, Western Australia	King County, WA
<b>Current population</b>	2.04 million [54]	2.19 million [55]
<b>Population change rate (5-year annual average)</b>	1.1% (5-year) [56]	1.7% [57]
<b>Median household income (per annum)</b>	Au\$85,436 US\$63,018 [58]	Au\$116,721 US\$86,095 [59]
<b>Land area</b>	641,786 ha [60] 2,478 square miles	547,930 ha 2,116 square miles [61]
<b>Gross population density</b>	3.2 people/ ha 823 people/ square mile	4 people/ ha 1,035 people/ square mile
<b>Commuter mode shares</b>	Car 64% Transit 10% Car pass. 5% Active trans. 3% Other 18% [62]	Car 63% Transit 13% Car pass. 10% Active trans. 7% Other 8% [63]
<b>Fixed public transport infrastructure km/ miles (guideway/ bus + rail)</b>	173 kilometres 107 miles	1,435 kilometres 892 miles <sup>11</sup> [64]
<b>Public transport fare system</b>	Zonal	Flat (King County Metro services); Distance-based (Sound Transit light rail)

<sup>10</sup> With the exception of 2017-2018 data for Perth, when passenger kilometres increased and may be associated with a newly-opened event-based rail station and events-based bus services.

<sup>11</sup> Includes King County Department of Transportation and Central Puget Sound Regional Transit Authority data – King County fixed distance alone has not been isolated for this study.

Investment in public transport infrastructure in King County has been outstripping that in Greater Perth, albeit the current Western Australian Government's METRONET plan is intended to yield significantly more kilometres of train track. Still, King County hosts a much more diverse and competitive mobility services market and a population with higher incomes on average, inferring more disposable income *ceteris paribus*, for spending on preferred transport choices.

Together, the data set an intriguing context for further study. In particular, our difficulty acquiring detailed mobility-related data for each study area added impetus to the review of mobility data access

and applicability (**Section 4**) and the potential for sentiment analysis to provide much more detailed insights into the varying travel and mobility trends (**Section 5**).

The next section of the paper explores in more detail gaps in knowledge owing to the benefits and limitations associated with traditional methods of transport data collection and analysis. It continues with discussion of alternative methods being enabled by technology and innovation as well as how the current regulatory and business landscapes facilitate or inhibit access to meaningful data. The section concludes with consideration of potential limits to what data and data tools can tell us.

	Greater Perth, Western Australia	King County, WA
<b>Public transport operating model</b>	Centrally-regulated; route contracts	Centrally-regulated; publicly-run service
<b>Public transport system boardings</b>	140 million [65]	128 million <sup>12</sup> [66]
<b>Public transport system boardings (5-year trend rate)</b>	-1% [67]	+5% <sup>13</sup> [68]
<b>Public transport system passenger km/ miles</b>	1,392 million kilometres 865 million miles [69]	967 million kilometres 601 million miles <sup>14</sup> [70]
<b>Public transport system passenger km/ miles (5-year trend)</b>	-1% <sup>15</sup> [71]	+4% <sup>16</sup> [72]
<b>Private mobility services available (excluding taxis)</b>	Ride-hailing <sup>17</sup>	Ride-hailing Ride-pooling Car-sharing Bicycle-sharing P2P car-hire P2P parking-hire
<b>Private mobility vendors operating</b>	Three – Ola, Uber and Shofer	Many vendors
<b>Does government have a public New Mobility Strategy?</b>	No	<u>Yes</u>

<sup>12</sup> Ridership data is for King County Metro, only as per previous figures. In 2017, Sound Transit ridership was 46.8M and Washington State Ferries, 24.2M. Our study did not identify what percentage of these riderships had an origin and/ or destination in King County.

<sup>13</sup> Applicable to King County Metro ridership only

<sup>14</sup> King County Metro only.

<sup>15</sup> Excludes 2017-18 (a significant increase occurred that may be associated with the opening of Optus Stadium).

<sup>16</sup> King County Metro only.

<sup>17</sup> Bicycle- car-share are not included as these are limited, location-specific offerings in Perth.



## 4 The quest to know more

### 4.1 The data drive

The public sector seeks data relating to people's travel behaviour and impacts of this behaviour on transport systems and infrastructure for many and obvious reasons:

- Asset management purposes
- Understand the impacts of, and to inform, policy
- Assess investment priorities
- Gauge efficiency, equity and efficacy of those investments
- Appraise externalities of transport and mobility (e.g. impacts on the environment, development patterns and public health)
- Facilitate estimation of effects of future changes to the system.

Likewise, the private sector seeks similar data to refine service offerings, allowing growth, service refinement and profit maximisation.

Historically, three significant challenges have limited access to and the explanatory power of mobility data:

1. Collating applicable data at-scale
2. Being able to unpick the complex reasoning behind travel choices to inform future policy and strategy
3. Refreshing datasets to gauge impacts of new policies, investments and/ or other interventions.

GOVERNMENT  
Greater  
Perth

Where trends are evident, there has not been enough deep digging to fully understand what they – the trends – mean: e.g. young people driving less – what are young people doing if they are driving less, but also taking public transport less?

Some digital tools permit access to mobility services, coordination of those services and collation of vast pools of data hereto unimaginable. Yet, data alone is insufficient. It requires processing and contextualisation by sophisticated digital programmes. Critically, those designing and applying the programmes need to understand the mobility queries requiring answers and what data is important in those respects.

Late British architect, academic and critic Cedric Price once argued “*technology is the answer, but what was the question?*” In the context of the current mobility landscape, perhaps the word data could be substituted for technology.

For both the public and private sector, data is growing as a commercial commodity. Yet, data has no intrinsic value. Its value is derived from potential application including increasingly, how it could be applied to predict future travel behaviour [73].

Useful data sources depend on what problem you are trying to solve, and what your metrics are. They vary on place, type of issue, what the city's values are and so on.

GOVERNMENT  
King  
County

GOVERNMENT  
Greater  
Perth

Government needs to be sceptical with big data. What can we make it tell us? What is the question that needs answering? Where is the data coming from? Who is it helping? Also, who benefits from open data? The risk is nefarious use - so what are the motives?

**Section 4.2** reviews briefly traditional travel and mobility assessment methods, highlighting their benefits and limitations. **Section 4.3** provides more discussion regarding the potential for new methods in our current digital era. **Section 4.4** explores the current regulatory and business landscape and how this abets or hinders access to meaningful transport data. **Section 4.5** discusses data protocols and standards, which are essential for facilitating meaningful processing and application of transport data. Finally, **Section 4.6** considers some of the potential limitations to the explanatory power of new datasets.

## 4.2 Traditional travel and mobility assessment methods

Insights into mobility preferences, experiences and factors have been often gleaned from surveys, interviews/ focus groups and counter-data (e.g. tube counters, door-step counters and other location-specific devices). Typically, these methods are employed for specific purposes and limited by cost, and access to labour and the population of interest.

They are limited also by those asking the questions of the data such as transport planners and systems owners, because they are not looking at the core reasons mobility is required. Data is used often to appraise what people are doing, which tends to be limited (or at least influenced) by mobility choices available as well as other characteristics of the built environment. So, in many cases, traditional methods have created confirmation bias rather than insight into potential needs.

In some instances, a mixed-methods approach is employed to address challenges such as representativeness, response rates, data accuracy and costs [74]. Utilising primary as well as secondary data has been shown also to increase the overall efficiency and quality of the research [75].

**Table 2** summarises a range of traditional data collection methods including their relative benefits and limitations. In this summary, we are concerned with these features in broad terms: many other papers evaluate in detail the merits of different techniques. For example, in many cases, the larger the sample size, the more representative the data

GOVERNMENT  
Greater  
Perth

In the past, transport habits were relatively consistent. Today, they are much more dynamic and different journeys are more complex. Simplistic models of bus services and roads doing the same thing every day do not work as well anymore compared to the dynamic nature of private mobility services.

that is collected may be. Furthermore, the efficacy of any investigation depends on the nature of the research questions, the calibration of the research tools and application of data collected into an analytical framework (e.g. model) and so on.

These traditional methods are limited in terms of their power to provide insights into rapid and dynamic changes within trip marketplaces. Mixed methods approaches can address some of these limitations, but these become expensive and have lags between data collation, analysis and reporting.

New, more powerful means of collecting, assessing and exploiting trip and travel data are required. These are both necessary and possible because of:

- The disruptive effects of new mobility services including how these are redefining supply and demand for travel (such as fulfilling latent demands and providing choices consumers did not have previously); and
- Up- or down-trends in public transport use.





**Table 2**

Comparison of traditional data collection methods

Method	Examples	Benefits	Limitations
<b>Single-response surveys</b>	Census Household travel surveys: e.g. the US 2016/ 2017 National Household Travel Survey (NHTS) and Perth And Regions Travel Survey (PARTS)	Statistically representative and powerful; key means to chart transport trends at a population-level; some gauge of motives and strength of motives when applying Likert scale or similar approach	Time and labour-intensive to administer; will not capture short-run changes; usually include limited data fields; cannot typically address localised and novel mobility issues; do not dig into travel behaviour complexities
<b>Stated and revealed preference surveys</b>	Specific survey or interview to gauge personal choices associated with specific issues	Can be statistically representative and powerful; can drill into some of the complexities of travel decision-making; can gauge some possible consequences of change within the mobility system	Some stated preferences do not mirror decisions made in real markets [76]; revealed preferences rely on recall; time and labour-intensive to conduct at-scale; usually provide insights limited to a few specific issues
<b>Multiple-response surveys</b>	Panel/ longitudinal surveys to gauge change-over-time	Can reduce influence of self-selection; statistically representative and powerful if sample size is maintained; means to chart transport trends at a population-level; some gauge of motives and strength of motives when applying Likert scale or similar approach	Particularly time and labour-intensive to administer; problematic to address localised and novel mobility issues; limited opportunity to dig into travel behaviour complexities require measures to overcome problems of attrition [77]
<b>Interviews</b>	Focus groups; Interviews; Delphi surveys (population of interest or domain experts)	Allows complex travel behaviour to be explored and rich explanatory data to be collected; researcher can clarify meaning in real-time	Time and labour-intensive to administer; limited sample size - findings do not scale to population level; length, frequency and complexity of interviews can lead to participant refusal/ attrition [78]
<b>Counter data</b>	Gate counters; Tube counters; Doorway counters	Gate counters; Tube counters; Doorway counters	Requires researchers to make significant inferences about travel (e.g. purpose); localised in nature
<b>Visual audits/ observations</b>	Video surveys; Traffic cameras; Field observations	Can provide significant, statistically powerful datasets applicable to a specific location; relatively limited time and labour requirements depending on study parameters	Requires researchers to make significant inferences about travel (e.g. purpose); localised in nature; significant cost and labour-intensive to conduct at-scale



### 4.3 New enablement/ data potential

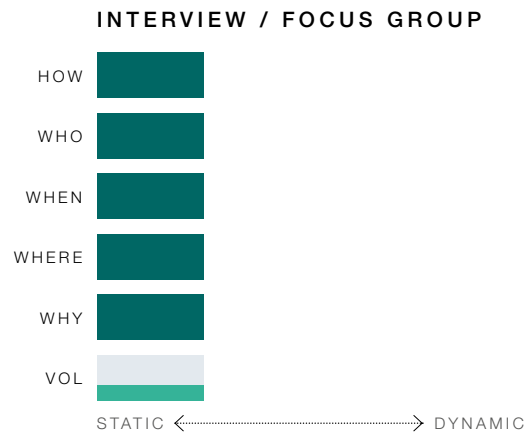
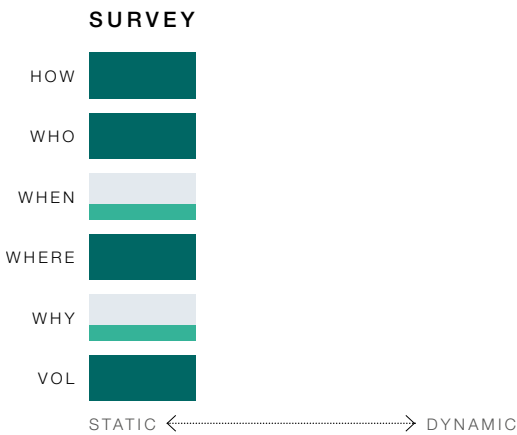
Innovation in the private sector and development of new technology has precipitated enormous growth in all manner of data collection, retention and applications. For example, an increasingly popular method of combating plunging participation rates and high costs is the use of Skype- or VOIP-mediated interviews. These can complement interview research methods enabling wider reach at lower costs although studies show they cannot completely replace face-to-face interaction [79]. Also, as well as capturing vast amounts of trip data, public transport smartcards are assigned to users and linked to various socio-demographic and/ or socio-economic data (subject to privacy limitations) [80].



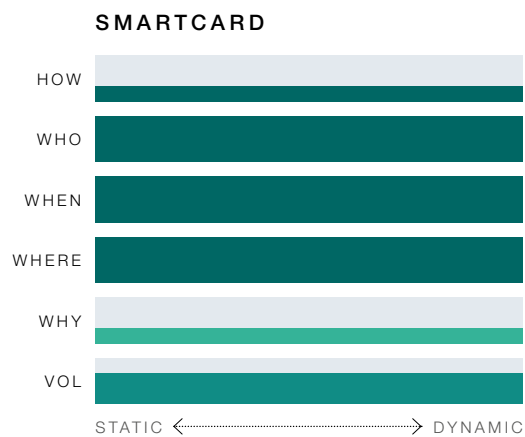
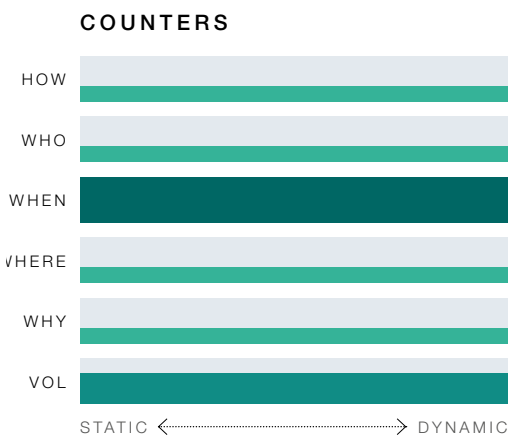
SmartRider is a very rich dataset but requires specific queries. About 10 years of data is now available.

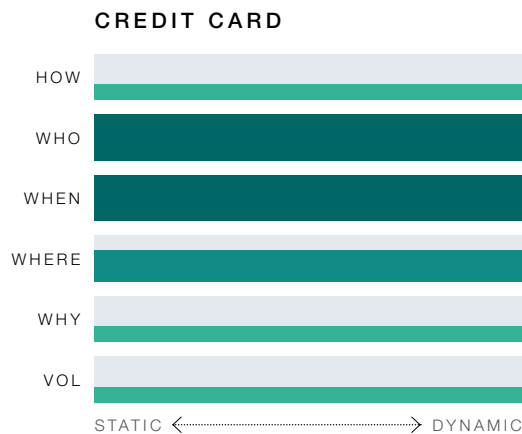
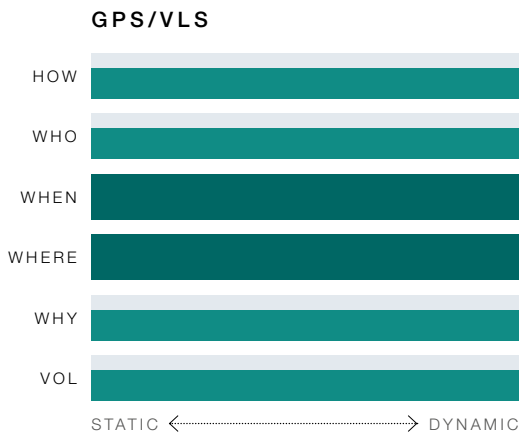
Computer processing power, digital databasing and ubiquitous Web coverage and access allows aggregation and display of transport data from many different jurisdictions. For example, TransitCenter.org hosts a tool enabling users to compare various time-series transport, land use and other statistics for different locations across the US.

Many other, newer methods rely on GPS, chip-card, Bluetooth and WiFi-based products and services. These facilities can provide deep insights, when data is processed appropriately, into all manner of dimensions of travel behaviour: how, who, when, where and why. Moreover, the datasets refresh constantly and in real-time, addressing two of the major constraints on traditional collection techniques: lag (currency) and stasis (lack of dynamism to gauge change). **Figures 10 to 17** are theoretical diagrams comparing broadly the relative power of various traditional and newer data collection methods to provide insights across a typical array of travel dimensions.

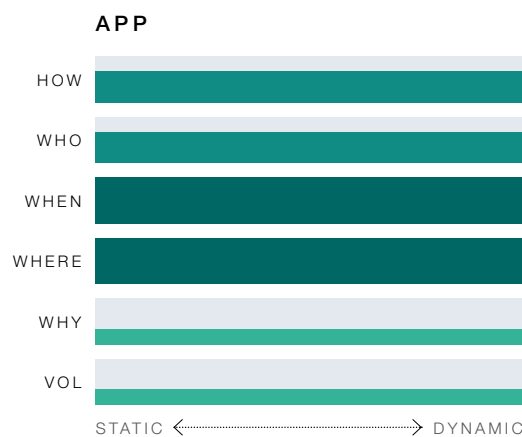
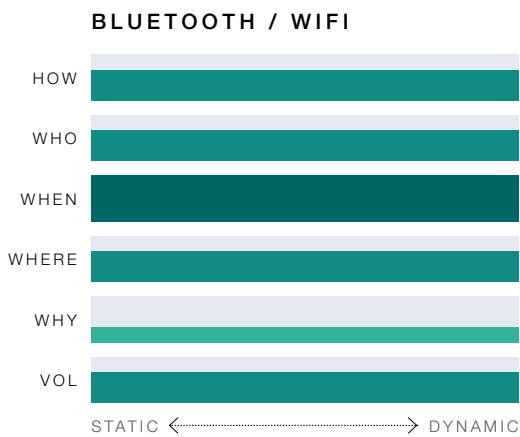


**Figures 10-13**  
Theoretical diagrams comparing the efficacy of different methods of data collection. VOL refers to volumes of data that tend to be generated by each method. The X axis indicates whether data is measured once or continuously.





**Figures 14-17**  
Theoretical diagrams comparing different methods of data collection. VOL refers to volumes of data that tend to be generated by each method. The X axis indicates whether data is measured once or continuously.



TSP can analyse issues in real-time - what was the trip ETA? Was there a functionality issue? What was the cost of the trip? This enables a better understanding of what causes people to take/ not take the ride.

Much of the data that can be used to generate insights is collected incidentally to the primary service or good provided to the consumer (e.g. in the case of credit cards) and/ or yielded voluntarily by consumers as a condition of receiving a good or service (e.g. to TSP to enable them to supply point-to-point mobility). In some cases, companies are able to convince consumers to supply data voluntarily to supplement or create their core business. For example, the crowd-sourced and aggregation-based mobility platform Moovit claims to have more than 400 million users, globally and leverages many of these users to contribute data about the mobility network within their own cities [81]. Without these volunteers, the company would have a much weaker offer.



As people rely more on app-enabled mobility, TSP use of extremely accurate real-time data can sway people towards private mobility services and away from transit. Investing in true real-time data and making it available to mapmakers is critical.

Governments are being courted by service vendors that offer access to data collected through a variety of different sources including smart phone pings and analysis of this data through AI. For example, Portland in the US has contracted Sidewalk Labs to analyse urban transport using their *Replica* software [82]. Despite its sophistication, the software still makes inferences regarding many dimensions of travel including mode and purpose. Input data is also generated without real sentiment attached, which can shed light on why people travel the way they do and how they react to changes to the mobility environment.

There is particular public-sector interest in datasets retained by TSP compared to smart phone data, because the former includes much more specific attributes regarding people's trip origins and destinations (where), time of travel (when) and how the travel was conducted (how), alongside personal information about each consumer (who). Travel time information can be used as a proxy for network performance and subject to sophisticated analytics, the collective dataset could provide insights into all manner of other complex mobility questions like mode substitution (e.g. substituting ride-hailing for public transport for particular trips).

#### 4.4 Current regulatory environment – data access

The UK Government has identified that access to and sharing of transport data is essential to enable the public sector to fulfil its responsibilities effectively given trends and trajectories in mobility [83], but the public and private sectors have different stakes in the data game.

Government's primary objective should be servicing the public good. With such *modus operandi*, more data is generally better enabling more informed decision-making. The public good is not the primary objective of private-sector enterprise including TSP [84]. For TSP, their trip data can provide a competitive advantage, informing service assignments by time and place, service quality requirements and price points.

GOVERNMENT  
Greater  
Perth

TSP can be agile and dynamic in responding to customers through data analysis. They can collect, analyse and respond very quickly. A good example is when TSP vehicles were banned from Optus Stadium when events were on post-opening. They responded quickly with discounted rates to drop people nearby.

While TSP may be signatories to virtuous mobility frameworks, like the *Shared Mobility Principles for Livable Cities*, the prerogatives of the public and private sectors can be in stark contrast in a variety of ways:

- TSP seek a share of the massive commuter trip market segment, but governments would prefer commuters to use mass transit along main corridors to mitigate congestion
- TSP would prefer to minimise costs and maximise revenues, which means *inter alia*, providing more drivers in areas of higher demand. This may contravene public spatial equity goals
- TSP apply pricing as a means to try to maximise revenue even if this means short-term discounts to attract new riders. Government aims to protect the integrity of the public transport system and would prefer for TSP to not pull riders from marginal bus services and exploit operational failures<sup>18</sup>
- Governments are responsible for the access and mobility needs of whole populations while TSP are interested only in profitable markets.

Seattle has been sued and identified as a target of the companies.

GOVERNMENT  
King  
County

Furthermore, without legislation or contract requiring provision of data, there is a lack of incentive for mobility service vendors to hand data over. For them, this equates to cost and inconvenience as well as potential for reduced competitive advantage [85]. So, do TSP datasets represent the 'golden egg' for transport planners and policy-makers? If so, what is the contractual and legislative landscape like with respect for access?

In Australia, regulations differ between States and Territories regarding TSP operations and data reporting requirements. In all cases, the regulatory environment is still evolving and there are no common data formatting or handover protocols.

<sup>18</sup> See <https://slate.com/business/2016/03/washington-d-c-subway-shutdown-is-good-for-uber-and-lyft-but-bad-for-commuters.html> and <https://www.blogto.com/city/2018/02/lyft-offering-cheap-rides-during-ttc-subway-closures/>



Government is playing catch-up and it is now unlikely that any jurisdiction will be able to enact thorough 'pay (supply data) to play' legislation, even if this may seem to be fair recompense for TSP services using public infrastructure like city streets [86]. According to Transportation for America, "*Cities did not act quickly enough to set data-sharing standards when transportation network companies (TSP) arrived almost a decade ago...*" [87].

Furthermore, it is too early to know how any data provisions will work in practice, with enforcement and validation protocols to be determined. Given the currency of the challenges and dynamism of the mobility services sectors, there is also no real government track-record regarding processing and meaningful application of any data that may be handed over. The efficacy of these activities remains to be seen.

In Victoria, the State Government has instituted a levy of Au\$1 on every commercial passenger trip as a source of revenue. Regarding data, as per *Commercial Passenger Vehicle Regulations 2018*, certain data must be stored by service vendors for a period not less than three years and handed over on request. These data include:

- Date and time the hirer requested the commercial passenger vehicle service for
- Driver accreditation number of the driver of the commercial passenger vehicle
- Registration number of the commercial passenger vehicle booked
- Commencement and end dates and times of the booked trip
- Commencement and end addresses or GPS coordinates of the booked trip
- Full amount charged for the trip
- Whether or not the fare for the trip was paid for in whole or in part under a subsidy scheme administered by the regulator in respect of which the driver of the commercial passenger vehicle or the booking service provider processed that payment
- Whether or not the hiring was provided using a wheelchair accessible commercial passenger vehicle.

In New South Wales (NSW), the Point to Point Transport (Taxis and Hire Vehicles) Act 2016 requires TSP to be authorised by Government as a booking service provider, which makes that vendor subject to data retention and supply requirements under the Act. Additionally, a dollar per trip is levied on TSP, which is used to fund a compensation scheme for taxi operators affected by the disruption in the sector.

In Western Australia, a case study geography for our study, legislative reform commenced in late 2015 with release for public comment of a Green Paper pertaining to mobility sector innovation. The Green Paper, *On-demand Transport A discussion paper for future innovation* prepared by Department of Transport Western Australia (DoT WA) focused on licensing reform rather than data provision. This reflected regulatory priorities at the time.

Thereafter, State Parliament passed the *Transport (Road Passenger Services) Act 2018* with staged implementation commencing late February 2019. On-demand booking services (including TSP) were required to apply for authorisation before 1 April 2019. Following full implementation by 2020, the Act is intended to:

- Deliver an industry that is accountable for the provision of safe and on-demand transport services
- Allow individuals and companies working in the on-demand transport industry to determine their own business operating models
- Have consistent requirements for players within the industry based on the nature of the service being provided
- Make it easier to enter and exit the on-demand transport industry and to operate fairly within it
- Allow government to have a good understanding of the size, nature and performance of on-demand transport services for planning purposes and ensuring that regional and vulnerable groups have reasonable access to services.

Other provisions of the Act will institute new fees and charges relating to the authorisation of on-demand booking services including an annual authorisation levy, which varies according to the scale of the enterprise (e.g. vehicles in service fleet).









Shofer has services aligned with coming transport on-demand Bill and indicated willingness to share data.

Under the Act and associated Regulations (*Transport [Road Passenger Services] Regulations 2019*), service vendors must keep records for no less than two years in a form and manner approved by the Regulator including:

- Day and time at which the booking was taken or facilitated
- Day of the associated journey and the times it began and ended
- Locations where the associated journey began and ended
- Name and driver's licence number of the driver of the vehicle
- Vehicle licence number or interstate vehicle licence number of the vehicle
- Any contact details provided by the person who made the booking or to whose account the booking was charged
- Number of passengers carried who were seated in a wheelchair (if any)
- Whether the vehicle was an electric vehicle
- Amount payable for the on-demand passenger transport service and the components of that amount.

Much like in NSW, Western Australian legislation *may* facilitate handover of specific, voluminous, ordered and consistently-formatted TSP data; however, application of the provisions of the Act is yet to be tested and data will not shed light on historic nor current travel. TSP may choose to test validity of data requests in court as well. Furthermore, as we expressed earlier in this section, data handover in and of itself does not provide answers to pressing mobility queries.

In the US, access to TSP and other transport data varies by state and city. Boston, Massachusetts (MA) is one of few cities in the country to have legislated for TSP data to be made available publicly. Still, the datasets available are coarse, as TSP are only required to report on town/ city of trip origin/ destination, as well as aggregated data regarding routes and time. Furthermore, public agencies in MA have significant issues addressing non-disclosure limitations on richer datasets collected or derived otherwise by the private sector, including TSP [88].

From early 2017, New York City's Taxi and Limousine Commission enacted requirements for handover of data from TSP relating to trip origins and destinations. Bruce Schaller has analysed and reported on such data extensively in his prominent series of research papers.





In Chicago, Illinois (IL), the City has commenced publishing publicly via its Open Data Portal detailed TSP data including information relating to trip origins and destinations, fares and times. This aligns TSP data sharing requirements with traditional taxicabs [89].

In Seattle, an agreement was reached in 2014 between the City and TSP for limited data handover. Quarterly, TSP are required to provide information regarding:

- Total number of rides provided by each taxi, for-hire vehicle license holder or TSP
- Type of dispatch for each ride (e.g. hail, phone, online app)
- Percentage or number of rides picked up in each ZIP code
- Pick-up and drop-off ZIP codes for each ride
- Percentage by ZIP code of rides that are requested by telephone or applications but do not happen
- Number of collisions, including the name and number of the affiliated driver, collision fault, injuries and estimated damage
- Number of rides when an accessible vehicle was requested
- Reports of crimes against drivers
- Records of passenger complaints
- Any other data identified by the Director of the Department of Finance and Administrative Services to ensure compliance.

The reports provided to the City include trip date and time fields as optional for mobility service vendors to populate: in our experience, they tend to opt not to provide this data<sup>19</sup>.

In 2015, TSP applied legal pressure to block the public release of a City report examining the state-of-play with respect for ridership and an assessment of ridership trajectory. The mobility service vendors relented only after an unfavourable Supreme Court ruling pertaining to withholding of 'trade secrets' articulated in government documents [90].

For researchers such as us, a public disclosure request to the City to obtain TSP data can trigger notification of the parties from whom the data has been supplied. Those parties may then seek a temporary restraining order preventing the disclosure if they so elect<sup>20</sup>.

Washington State has the most conservative public disclosure laws in the country, so mobility companies have fought the City to keep private the data submitted to the City as a condition of operation.

The legislation in Seattle requesting this data was enacted in 2014, so what it asks for (Zip-code-level data) is not very useful from a planning and operations perspective, but it is from a legislative perspective.

#### 4.5 Data protocols and standardisation

Strategy 4.2 in the City of Seattle's New Mobility Playbook is to "work with regional and national partners to establish a neutral trusted data platform that houses data from new mobility service providers, sensors, and other data sources, automates data analytics, and enables predictive analytics" [91: p44]. This is noble, bold goal but one difficult to satisfy given the challenges and limitations explored already in this paper.

For one, there is no established, best-practice with respect for TSP dataset formatting or disclosure and this is evident in the variability of legislation [92]. It is also because government (and the research community) is still working out the questions to ask of the data and whether the data can provide the answers sought.

In time, we anticipate that the International Standards Organization (ISO) or SAE International may establish such a standard. Currently, SAE International is convening cross-sector input to develop best practice standards for micro-mobility – efforts pertaining to ride-hailing etc. may follow in time [93].

In the meantime, industry looks to data vendors (which can include entities such as credit card and telecommunications companies), government, research institutes (e.g. University of Washington's [Transportation Data Collaborative](#)) and quasi-government and independent open data platforms such as [OpenTraffic](#), [SharedStreets](#) and [oneTRANSPORT](#). All have limitations relating to standardisation, opt-in reporting by industry, pay-to-access (in some instances) and so on.

---

<sup>19</sup> Evident in the data provided to our research team following a public disclosure request.

<sup>20</sup> In our case, the parties elected not to.

The National Association of City Transportation Officials (NACTO) in the US has published a series of city data sharing principles. These aim to “*facilitate better working relationships between cities and private transportation providers by elevating and **standardising** [our emphasis] the process of data sharing. These will enable more proactive, data-driven transportation planning solutions and allow private mobility providers to exchange data with cities in the most secure, cost effective, and efficient manner possible*” [94: p1]. Among other things, NACTO’s principles include mobility service vendors providing city officials with data regarding corridor-level trip routing (speed, travel time and volume), pick-up and drop-off locations, vehicle occupancy, non-revenue VKT/ VMT and vehicle dwell times [95].

Standardisation includes both uniform sharing *practices* and uniformity to the *format* of the data provided. Positively, with respect for the latter, there is increasing industry buy-in to formatting standards such as the General Transit Feed Specification (GTFS). Transportation for America’s proposed national standards for micro-mobility (e-scooters, for example), could be applied to other mobility services and include:

- Fleet API – public software intermediary as common requirement
- Authenticated and standardised data format such as Los Angeles’ Mobility Data Specification
- Access to historic ride activity data
- Consistent privacy practices focusing on data from vehicles, not individuals and protection of Personally Identifiable Information (PII)
- GPS tracking based on vehicles, not individuals.

In 2018, the City of Los Angeles launched a tool to assist with governance of the mobility market: the Mobility Data Specification (MDS). MDS is a data and API standard permitting the gathering, analyses and comparison of real-time data from mobility service vendors. The City also intends for the specification to help enforce regulations when needed. MDS was released publicly on [Github](#) to stimulate adoption by other governments, encouraging some degree of cross-geographic standardisation [96].

In Australia, TfNSW has developed a data specification for publication and sharing of mobility service vendor data associated with pilots included in the State’s [MaaS Innovation Challenge](#) [97]. The specifications are fairly broad-ranging, covering both current and historic datasets for analytical and planning purposes including “*gain[ing] insights into the viability of different schemes which different MaaS operators and providers are running to determine the future roadmap for Transport*” [98: p1].

TfNSW is seeking to publish data including real-time vehicle positioning and occupancy in addition to trip details (origins and destinations, travel times and costs) via [Open Data Hub](#). Yet, there is potential for access to various data to be limited to participants in TfNSW’s Challenge. Despite its limitation, TfNSW’s initiative may function as a litmus test for data reporting standardisation at a local jurisdictional level.

More broadly, there will be reluctance and cost-constraints to adoption by different governments of ‘uniform’ standards, despite the merits of doing so. For TSP that rely on efficient and standard, pan-geographic business models, a lack of pan-geographic standards for data supply to government will perpetuate reluctance to accede to requests for data because of the effort required for these to be satisfied.

#### 4.6 Potential limitations to data-driven insights

This section has demonstrated an extremely uneven playing field with respect for data access, formatting and completeness depending on location. In Australia, States and Territories are in the process of legislative reform to address these issues but reforms are variable and remain subject to road-tests, both in practical terms and in the courts.

Legislative requirements are more mature in some parts of the US, including Seattle/ King County; however, the data specifications are not as rigorous as transport planners and policy-makers may prefer. Typically, data available today either publicly or via Act or Statute, is unlikely to shed sufficient light on complex questions applicable to the new mobility landscape:

- What is really influencing mode substitution and under what circumstances?
- How dynamic is substitution and what factors create propensity for change (including both push and pull factors)?

In the next section, we define the potential value proposition of sentiment analysis, given that through its leveraging of rich social commentary, it may be means to drill deeply into reasoning behind certain travel decisions and appraisal of travel experiences. In this way, it may offer access to meaningful datasets and enable meaningful analyses of them.

## 5 The potential value of sentiment analyses in transport research

Sentiment analysis is a data-driven technique becoming more common globally to assess with currency and over time, consumer experiences and appraisals of variables of interest. It has particular traction in the consumer goods and services industries as means to provide vendors with rich customer feedback. TSP business models are data-driven and therefore, data-reliant. It is therefore expected that TSP use sentiment analysis to refine their offers and this is confirmed by our study's interviewees.

Sentiment analysis gives insight into the ideas, opinions, wants, needs and concerns of members of the public relative to topics of interest. The technique yields additional insights by including the geospatial and timestamp attributes of social data to contextualise further the thoughts and feelings people have expressed about a particular topic area.

Mobility Service Vendor

User experience (UX) is key to survival of TSP as a private company. Therefore, any and all information is needed regarding why customers do or do not use the service.

TSP employs sentiment analysis to determine what aspects of service people comment on and whether it is positive or negative feedback, which feeds into product development through the Net Promoter Score (NPS) survey. Sentiment is a metric that a large section of the company is focused on enhancing and is gauged regularly.

The ability to assign both emotions and polarity to mentions discovered in social data can yield a rich assessment of the emotional push and pull guiding people's actions. When enough of such comments are gathered from online posts over a period of time, a quantitative value of sentiment and emotion is built up next to various keywords, topics and phrases.

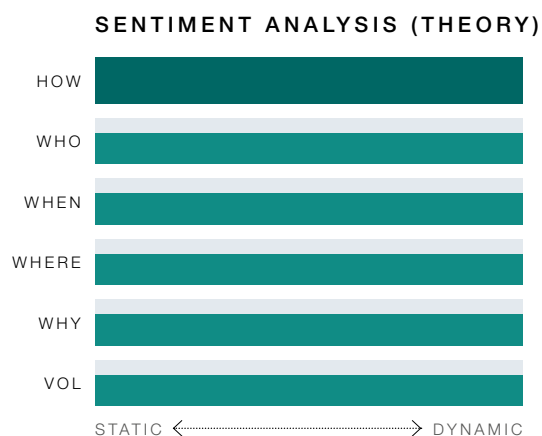
There is significant potential for broader applications of sentiment analysis - particularly to provide insights into complex issues contemplated with alacrity by transport planners and policy-makers but not well understood or explained. Earlier sections of the paper discussed traditional methods of data collection and analysis, and presented some system user satisfaction statistics from Greater Perth and King County.

Typically, government agencies rely on single-point, structured surveys to get feedback from consumers regarding the operational and experiential aspects of the public transport system. Limitations of these techniques include lack of comparison of mode alternatives (e.g. public transport and ride-hailing) and how (and why) people's experiences vary when using them.

The value proposition of sentiment analysis as a data collection and analytical methodology includes source data being:

- Accessible
- Provided in real-time
- Dynamic, allowing changes and trends to be observed
- Expressive – relatable to travel experience; especially the how (mode) and why (purpose)
- Independent of any research framework – data volunteered by travellers, not elicited through any targeted questioning or survey
- Potentially voluminous and originating from a wide audience, adding to its explanatory power at population levels.

In the aggregate and when subjected to relevant analyses, source data may provide strong insights into prevalence of mode choices and appraisal of those choices depending on factors such as trip purpose, travel time, location and characteristics of the user, as well as changes (including disruptions) to mode-based levels of service (see **Figure 18**).



**Figures 18**  
Potential sentiment analysis strengths





User experience is extremely important, particularly as customer expectations grow and change.



Understanding the customer and how the customer is changing requires ongoing assessment of customer trends, customer attitudes, and global technology trends and how these affect people’s decision-making.

Furthermore, the ability to geo-locate and timestamp sentiment data may provide a great deal of insight into the variability of service quality and appraisal depending on spatial, temporal and population factors. People express themselves frankly and candidly online, and often to make a point regarding specific brands (or service offerings). A brand today is perceived increasingly through the lens of the online community, so graphing sentiment/ emotion over time may be particularly useful for measuring brand equity of various modes of transport

Our research approach included definition and employment of a prototype study of public transport and ride-hailing user behaviour and commentary in Greater Perth and King County. Our intent was to define an approach that would be replicable and scalable, should the initial prototype satisfy our third research objective. A series of research questions helped frame our study further with regard for study objective 3 (refer **Section 1.2**):

- How dynamic is mode choice? E.g. do people discuss making mode choice decisions on the spur of the moment and if so, for what reasons?

- What affects mode choice in terms of consumers electing for public transport versus private mobility services, or vice versa?
- Do choice variables differ depending on the characteristics of trips, such as time-of-day, location and trip purpose?
- How do users report their experiences of public transport versus ride-hailing during or post-journey?
- What trip factors - comfort, privacy and so on - do sentiments correlate to?
- How has sentiment towards these modes changed over time and why?
- What are locational differences in sentiment and mode choices (Greater Perth compared to King County), and do these, alongside other insights, help to explain mobility trends evident in published data?

Objective 3 of the study was ascertaining whether new technologies - some of which are fuelling these new mobility business models - provide also the opportunity to address gaps in traditional and alternative data collection methods and methodologies allowing more robust interrogation of the travel market. We are interested especially in the potential of sentiment analysis in this regard.

The prototype study approach involved the following steps:

1. **Find relevant conversations.** Queries were built to search through social media archives in Greater Perth and King County from May 2017 to May 2018. The queries narrowed in on conversations held on specific topics of relevance to the study and its objectives. Many rounds of these searches were conducted, adjusting the query parameters slightly each time allowing gathering of the most conversations possible relevant to the given topics while also weeding out irrelevant commentary
2. **Geofilter conversations.** Metadata in the content as well as the author's social media accounts and conversational context were checked to make sure content was relevant to the specified geographic areas
3. **Organise conversations.** Content was divided into individual rows for each author, timestamp and location
4. **Assign polarity.** Each row was assigned a polarity level of positive, negative, or neutral using Bayesian sentiment analysis



5. **Assign emotion.** Each row was assigned one emotion: Anger, Anticipation, Disgust, Fear, Joy, Sadness, Trust using a proprietary Bayesian approach called ‘Convey’ as part of the *Conversus* platform. For example, someone may say, “I’m tired of how late the buses are running these days...it’s easier to share a ride with my co-worker who leaves an hour earlier than I am used to but at least I get to work on time!! Come on [service provider], get with it!” In such a case, a negative sentiment tag would be assigned plus the emotion of disgust to the keywords “[service provider]”. A positive emotion tag would have been assigned during step 4 plus the emotion of joy/ anticipation to the phrase of “share a ride with my co-worker”

6. **Create topics.** Each row was labelled by topic according to the most frequently occurring keywords within the closest proximity using the Latent Dirichlet Allocation (LDA) algorithm. LDA treats each row as a mixture of topics and each topic as a mixture of words. This allows rows to ‘overlap’ each other in terms of content rather than being separated into discrete groups in a way that mirrors typical use of language. LDA is a mathematical method applicable to written content resembling how unsupervised clustering works on numerical data. LDA’s biggest strength is that it provides a method for building a distribution of topics across a large number of documents – in this case, social media posts and comments. Instead of creating clusters where each document must have single membership, LDA’s ‘fuzzy memberships’ provide a more nuanced way of finding similar items or discovering user profiles/ personas. LDA’s main weakness when it comes to analysing topics in social media is that the documents (such as comment fields on a social media post) are often quite short. This can create a sparsity problem when the algorithm is trying to measure how close a document is to previous observations. This weakness is counteracted by gathering vast quantities of data

7. **Create graphs.** Sentiment and emotion was graphed over time to identify specific subsets of the data worth more detailed hand analysis. This may have included, for example, instances of mode substitutions following on from previously-expressed negativity regarding trip types

8. **Analyse results.** A subset of conversations was selected to analyse by hand based on completion of the previous steps in the approach enabling formation of a more complete picture of both the sentiment and emotion expressed in social media mentions. This process helps to address the limitations of algorithms, as these do not actually ‘understand’ the words they are analysing and nuances of language like sarcasm can be missed<sup>21</sup>. Concurrently, project team members refined our understanding of additional factors relevant to the study such as the reasons behind a person’s mode choice and their motivation for posting

9. **Draw conclusions.** Specific conclusions about mode use and appraisal trends over time were identified through rigorous dissection of what people have said, why they have said it and when they said it.

An added advantage of this methodology is the potential to build sentiment profiles associated with socio-demographic variables of interest. This is done through analysis of use of public-facing social media user profile data and socio-demographic stratification of sentiment.

The next section of the paper presents findings and discussion of our prototyping exercise.

---

<sup>21</sup> Intriguingly, Google researchers have unveiled recently AI that is showing great potential to appreciate nuances of language including emotion, sarcasm and so on: <https://medium.com/the-new-york-times/finally-a-machine-that-can-finish-your-sentence-b322a16d4199>



# 6 Findings and discussion

## 6.1 Sample statistics

The 12-month study period (May 2017 to May 2018) yielded around 17,000 *polarised* social mentions of public transport and 21,500 of ride-hailing in Greater Perth. Comparatively, there were about 15,000 regarding public transport and 34,000 ride-hailing in King County. **Figure 19** to **Figure 22** show trends in mentions in both study areas over the reporting period based on polarity.

Flat periods can be explained by the ways in which technology and social media policy evolves over time, requiring subtle recalibration of search and analytical algorithms to gauge sentiment. Furthermore, mentions can be associated with differing conditions within the mobility ecosystem, leading to relative peaks and troughs in the expression of sentiment. Over our study period, lulls were more observable for ride-hailing mentions in Greater Perth than for other mode/ location pairs.



◀ **Figures 19**  
Greater Perth public transport mentions.



◀ **Figures 20**  
Greater Perth ride-hailing mentions.



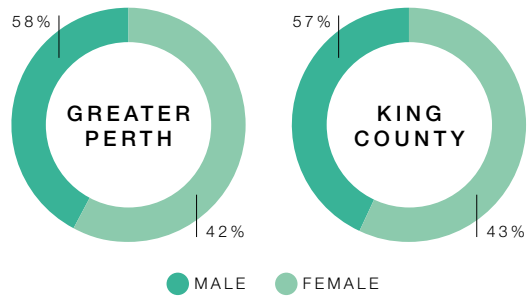
◀ **Figures 21**  
King County public transport mentions.



◀ **Figures 22**  
King County ride-hailing mentions.

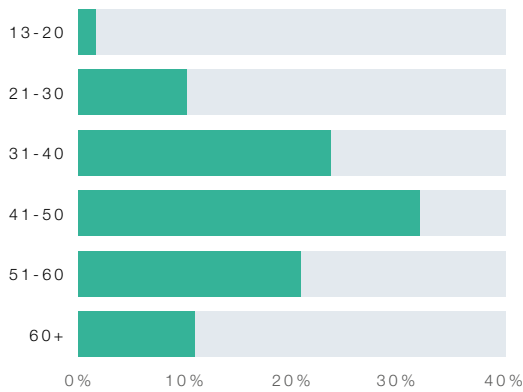
Based on publicly-available demographic information attached to social media posts, the following gender splits applied to the samples (**Figure 23**).

Age profiles are shown in **Figure 24** and **Figure 25**. These are not included inferring representativeness of the population: they are for information, only. Other sample demographics are much more onerous to identify and summarise and tend to be less reliable because of limitations to publicly-available attached to individuals.

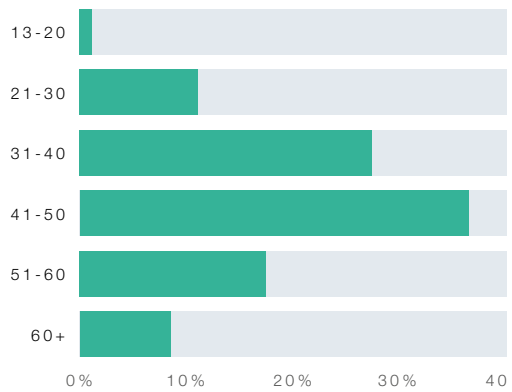


◀ **Figures 23**  
Social media posts gender splits.

**GREATER PERTH AGE BREAKDOWN**



**KING COUNTY AGE BREAKDOWN**



◀ **Figures 24 and 25**  
Age breakdown associated with relevant posts

## 6.2 Results and discussion

This sub-section reports and discusses results based on the research questions applicable to our specific prototyping of sentiment analysis (see Section 5).

### How dynamic is mode choice? E.g. do people discuss making mode choice decisions on the spur of the moment and if so, for what reasons?

A broad number of conversations were identified and analysed relating to the reasons people chose a specific mode. A large majority of commentators were found to choose modes based on habit and the data showed minimal dynamism in mode selection.

The clear majority of mode choice decisions were made ahead of time and fit a consistent pattern. There was virtually no spur-of-the-moment decision-making discussed. Moreover, the conversations inferred that some sort of significant, atypical event would have to occur to cause a reconsideration of mode choice - for example, a major service disruption like a rail line closure.

Figure 26 to Figure 29 show reasons that passengers reported making their mode choice demonstrating the strength of habit. The number of mentions analysed per passenger was limited to one (strongest, most direct statement or inference) so that each passenger was weighted equally.

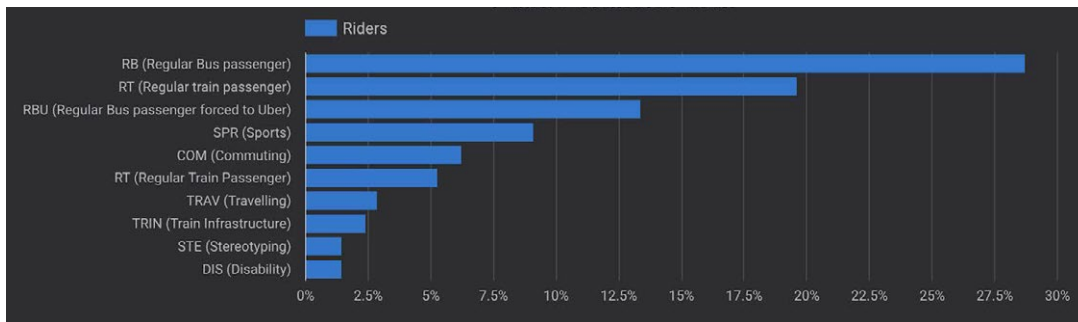


Figure 26  
Reported reasoning for mode choice – Greater Perth public transport

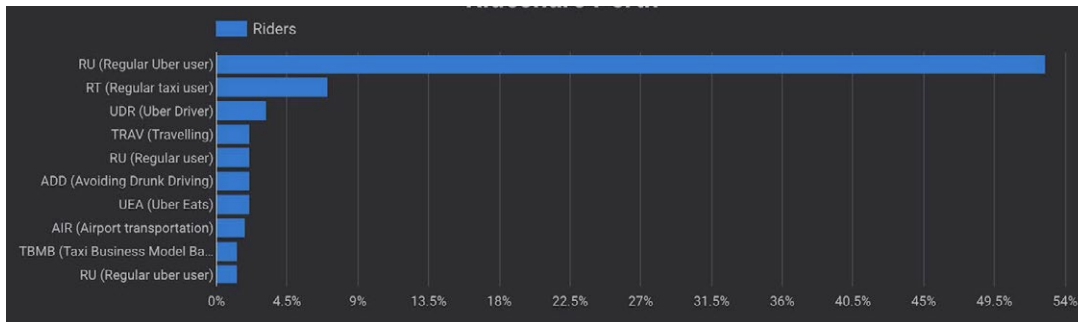


Figure 27  
Reported reasoning for mode choice – Greater Perth ride-hailing.

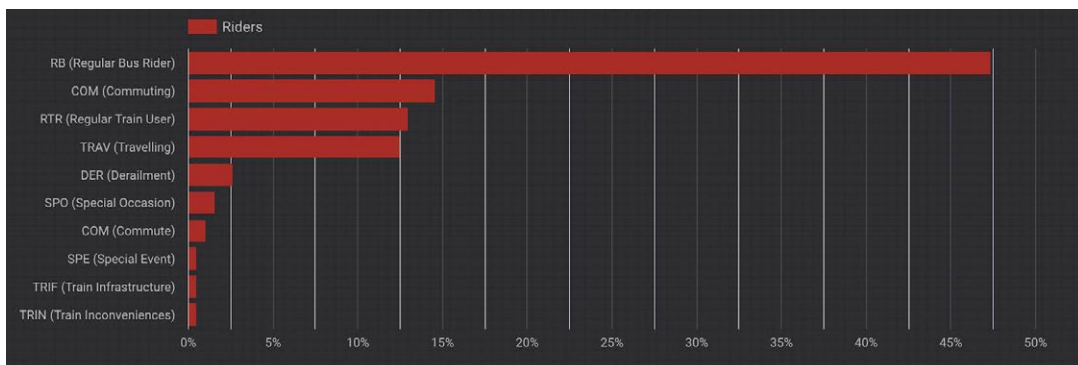
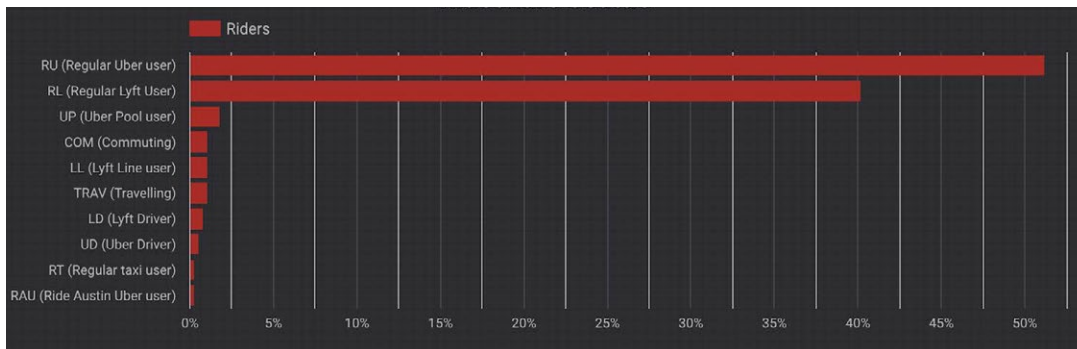


Figure 28  
Reported reasoning for mode choice – King County public transport.





◀ **Figures 29**  
Reported reasoning for mode choice – King County ride-hailing.

.....  
**What affects mode choice in terms of consumers electing for public transport versus private mobility services, or vice versa?**  
 .....

The primary factor identified affecting mode choice was trade-off between time travelling and the cost of the trip. Ride-hailing options were almost always chosen over public transport when the traveller did not consider the price to be prohibitive, but this was in select circumstances (see c), below).

Conversely, public transport users almost always expressed that they had more time than money available to budget. These riders either adjusted their schedules to accommodate the longer travel time or already had the time to spare.

The following are examples of conversations and comments that relay this tension between time and money:



**Weighing up time and cost budgets in specific circumstances:**

*“Trying to get to a friend’s for a party tonight and i can either catch a bus for 70 mins or pay \$35 for an uber”.*

*“...I’ve had business meetings with Uber, there are a lot of people that catch them to and from work daily”.*

*“I’ve been asked to go into work earlier... ha ha I’m going to have to go by uber because if I catch a bus I’ll be late... yay”.*

*“@Transperth I’ve been sitting at stop 14496 for 22 mins already. No buses. I’ve missed one connection already. Looks like I’ll miss the second. Are you going to pay for my uber to get home?”*

*“...Aww that sucks :( I think uber may be the better option knowing transperth bus timetables”.*

*“My morning has been terrible. My bus never showed. I was late to work I had a guy try to get me in his car at 5am AND my lyft driver was cray”.*

*“Let’s not talk about how often I Uber home from work because I’ve worked too late and it’s too dark/rainy to catch the bus or train”.*

*“Two hours on the bus. I hate getting up at sparrowfarts AM to get to work”.*

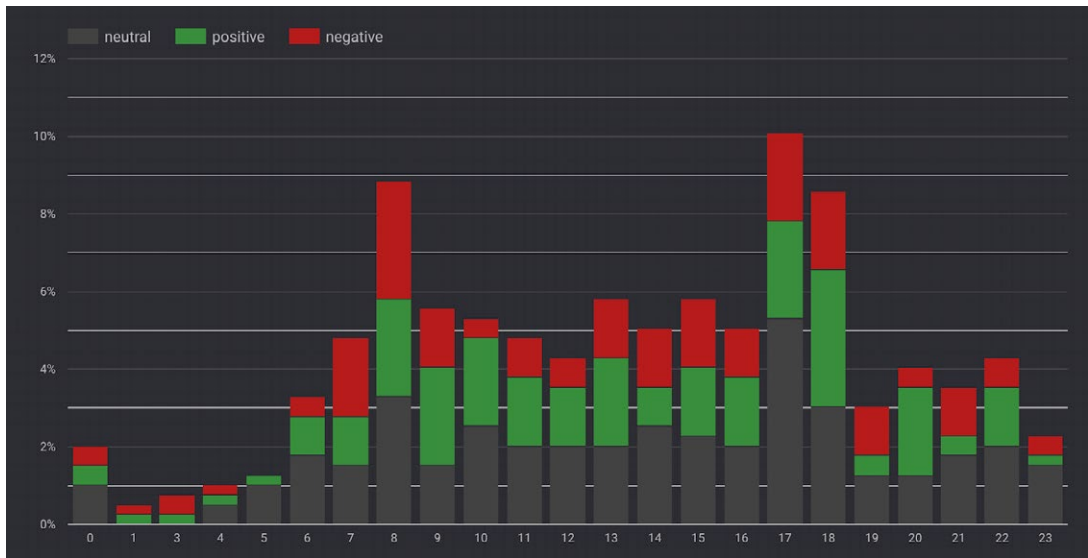
**Do choice variables differ depending on the characteristics of trips, such as time-of-day, location and trip purpose?**

The trade-off between time and money is a dominant factor in people’s travel choices. Our data showed that the characteristics of trips – including time-of-day, location, and trip purpose – only impact choice by the way they factor into the time-money trade-off.

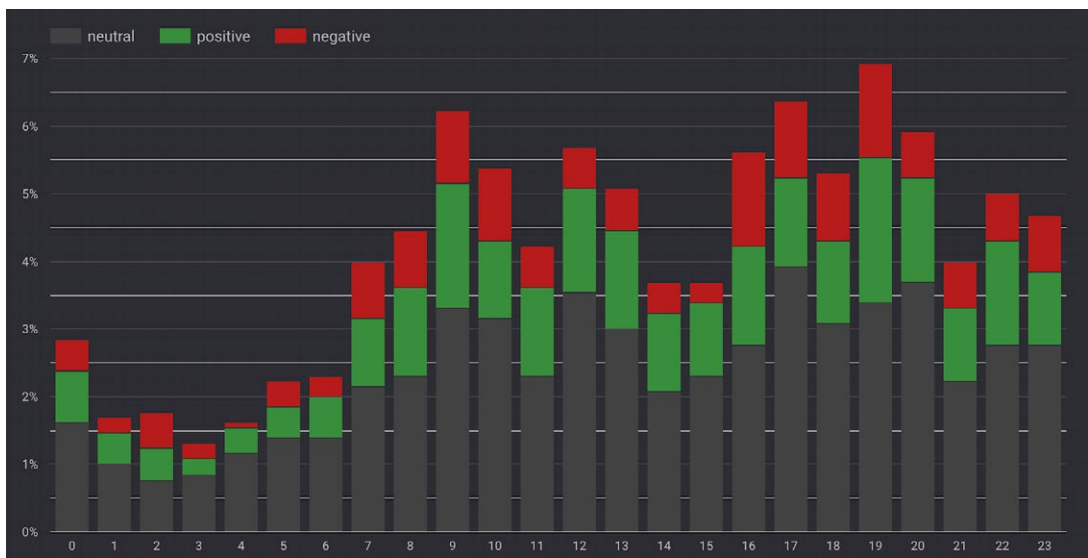
For example, frequent trips like commuting tend to be conducted by public transport because the cumulative cost is high. Some occasional trips, like those happening late at night, tend more often to be by ride-hailing. This reflects service availability during the off-peak and for certain origin/ destination pairs (convenience/ wait time/ travel time) and the less routine/ more discretionary nature of the travel (meaning passengers are more willing to spend more for a one-off ride than a daily commute).

The data points are much less useful for predicting mode splits than reasoning for mode choice because there is no correlation between mentions and frequency of use. This may be expected given use of a mode in and of itself does not prompt social media users to comment.

Profiling commentary by time-of-day (Figure 30 to Figure 33) does show that the frequency of social media posts and to some degree, polarity of comments, accords with those typical pain-points for the travelling public: peak hours. There was a spike in negative sentiment associated with public transport in King County during the morning peak. The profiles show also the higher frequencies of posts relating to ride-hailing during evening hours (i.e. during the public transport off-peak).



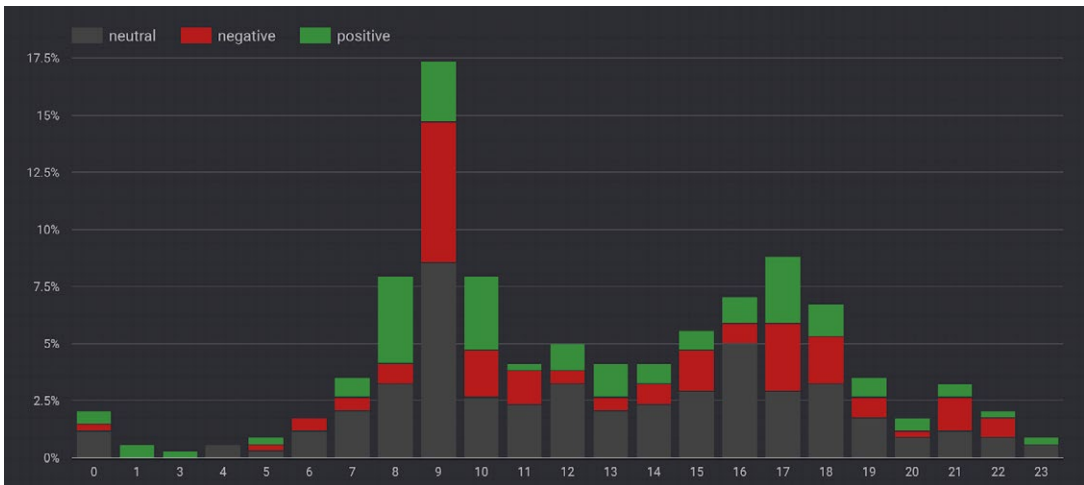
◀ **Figure 30**  
Greater Perth public transport commentary profile by time-of-day.



◀ **Figure 31**  
Greater Perth ride-hailing commentary profile by time-of-day.



◀ **Figures 32**  
King County public transport commentary profile by time-of-day.



◀ **Figures 33**  
King County ride-hailing commentary profile by time-of-day.

How do users report their experiences of public transport versus ride-hailing during or post-journey?

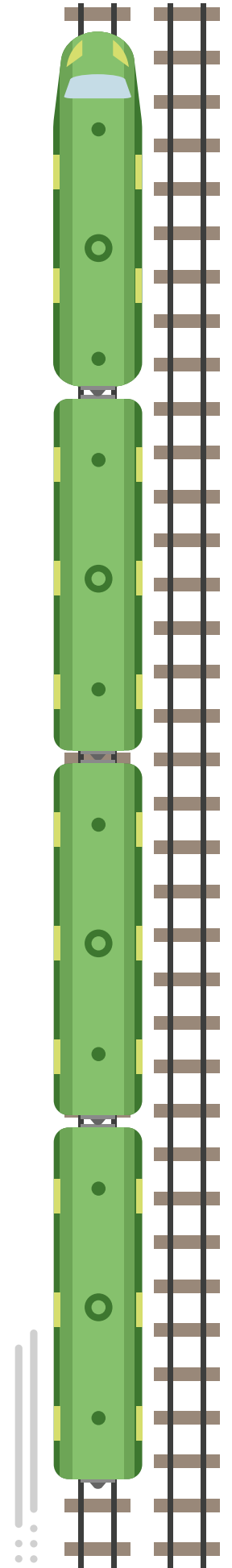
A significant majority of posts that discussed travel experiences were made post-journey - those associated with both public transport (92%) and ride-hailing (87%). This reflects people’s tendency to post when reflecting on travel not during the travel experience. This creates difficulties locating where specific events occurred and relating sentiment to specific origin/ destination pairs (especially for ride-hailing, given vehicles do not have unique identifiers relatable to routes as public transport services do), since the location metadata from the post is often not relevant to the trip taken.

What trip factors - comfort, privacy etc - do sentiment correlate to?

In Greater Perth, the most positive sentiment was attached to convenience of the public transport system, both in terms of the proximity of public transport options to users and the timeliness of services. Another factor that created positive sentiment was positive social interactivity with others during travel – applicable to both public transport and ride-hailing.

The most significant negative sentiment was attached to the cost of and customer service provided by ride-hailing companies. There was also strong negative sentiment associated with bad social experiences, with ride-hailing drivers referred to most frequently as cause for unhappiness.





For public transport, negative sentiment was posted most frequently in relation to schedules including variables such as buses running late and insufficient service headways. Also, some people expressed frustration with the impacts of public transport on the urban environment such as tracks, large barriers and noise.

In King County, the strongest factors influencing sentiment were social. People shared often about the positive social interactions they had while travelling: more often when using ride-hailing and less frequently, public transport. Similar to Greater Perth, the strongest negative sentiment correlations were with ride-hailing customer support issues.

Negative interactions with drivers were reported most commonly indicating the importance of social interaction in determining people's level of satisfaction with a trip. Traffic and delays were two other (interdependent) factors that correlated to negative sentiment and tended to be reported during peaks.

---

### How has sentiment towards these modes changed over time and why?

---

Refer to **Figure 19** to **Figure 22**. In Greater Perth, the data showed consistent reporting over time of both negative and positive sentiment relating to public transport, albeit, towards the end of the study period, positive sentiment trended closer to negative. A factor that may have contributed significantly to this trend is Transperth and its contractors' focus on customer service.

In contrast, negative sentiment relating to Greater Perth's ride-hailing marketplace outpaced positive by some margin. Multiple factors appeared to have contributed to this sentiment such as the social impacts of poor customer service including perceptions that drivers can 'game' the system to earn more by cancelling rides.

From a cost perspective, the higher and often unpredictable rates of surge pricing create strong negative sentiment. Lastly, but importantly, is some news coverage about drivers committing sexual assault, which can create a sense of fear and distrust toward the ride-hailing system.



In King County, in contrast to Greater Perth, negative sentiment associated with public transport outpaced positive. The major factors appeared to be lack of reliability with a growing perception that delays are increasing both in frequency and length. Simultaneously, concerns regarding irregularities in schedules and unreliability of these schedules appeared to have impacts.

King County's ride-hailing system generated a steep increase in negative sentiment over the period. The major contributing factors appeared to be customer service issues (with some vendors, especially), with some people believing that some company support structures were not scaling sufficiently to meet demands. Moreover, some people view select vendor business models as being unfair to drivers.

An overall increase in traffic over the period appeared to contribute to longer delays for some ride-hailing trips, making the time-money equation less favourable for users. Finally, there is another 'customer service' issue peculiar to ride-hailing: passengers appeared to become increasingly concerned that they will receive bad reviews from their drivers for not being social enough.



**What are locational differences in sentiment and mode choices (Greater Perth compared to King County), and do these alongside other insights help to explain mobility trends evident in published data?**

Collectively, the data show a range of clear commonalities and differences in sentiment (and the variables it is attached to), between Greater Perth and King County. These are summarised below.

**Commonalities:**

- A lot of mode choice is habitual. Variability is associated much more with disruptive events and discretionary/ atypical trips
- The modes appear to satisfy different segments of the trip marketplace but there is overlap. The overlap appears to be particular to individuals' balancing of time-money equations
- Customer service and satisfaction are big influences on social media postings (including their polarity). Over time, this may lead consumers to alter their travel habits
- In broad terms, ride-hailing companies appear to be eliciting worse sentiment tied to some of their customer service practices.

For ride-hailing, the relatively strong negative sentiment pertaining to customer service and experience may herald a tapering of growth in the sector - particularly if public transport levels of service are improved further. This is a trend that should be watched closely.

**Differences:**

The data does not shed light on public transport usage trends in Greater Perth and King County; however, it is informative for transport policy. In Greater Perth, the State Government is performing well with respect for core services: where they are provided frequently and reliably, consumers are satisfied. Furthermore, customer service is rated very highly.

In King County, there is a delta between negative and positive sentiment relating to public transport. Reasons for this are not particularly clear from our dataset but may relate to some perceptions of customer service as well as levels of service that require further improvement including travel time reliability via some routes.

The final section of our paper summarises our findings, comparing results from our prototyping exercise with those from the formative stages of our research. This assessment provides the basis for a series of proposed future applications for sentiment analysis based on its apparent strengths, a description of limitations and propositions for further work.



## 7 Key insights and conclusions

This White Paper discusses research conducted by Arup and Oniracom exploring the growing, disruptive effects of passenger services - particularly ride-hailing - provided by TSP on the mobility systems of cities. The exploration included case studies of Greater Perth, Western Australia and King County, WA, which were selected because of the medium-term downtrend in public transport ridership in one (Greater Perth), and uptrend in the other (King County). In King County's case, this is despite there being a much more diverse mobility services ecosystem than in Greater Perth.

Additionally, the study explored sources of knowledge and means to appraise these disruptive effects. Specifically, many governments understand the power of datasets held by TSP; however, many barriers are apparently typical to accessing and applying this data in ways that would shed light on the directionality, scale and significance of the disruption. Consequently, our research team prototyped sentiment analysis, a data collection and analytical methodology employed increasingly as part of good and services marketing practices, to ascertain whether it may have applicability for transport planners and policy-makers.

The prototype, focused on a year-long study of social media sentiment geo-tagged to Greater Perth and King County, yielded the following insights.

### **Mode choice is a habit.**

The conversations held on social media in both Greater Perth and King County indicate that mode choice tends to reflect habits and has little dynamism. Significant disruption - which may include repeated negative experiences of one mode - is generally required to alter behaviour. These findings accord with the general transport literature.

### **Time and money trade-offs drive mode choice.**

People typically trade off time and money when making travel choices. Ride-sharing tends to be preferred for discretionary trips and if timing is a critical variable (e.g. someone is running late or faces an imperative to arrive at their destination at a specific time - especially if the trip is atypical). These findings are consistent with other research that shows some segmentation of the trip market in favour of public transport and some, ride-hailing.

### **Social experiences often dictate sentiment.**

Travel by either public transport or ride-hailing involves some level of social interaction. The particulars of these interactions correlate strongly with whether expressed sentiments are positive or negative.

The largest volume of negative sentiment in both Greater Perth and King County related to poor customer service. Conversely, the largest volume of positive sentiment in both context cities related to good customer service and positive social experiences while travelling.

Public transport's affordability overcomes some other barriers to use. Public transport presents new and occasional riders with some barriers to use including the need to learn or interpret routes and schedules. This tends to be more acute when trip demands need to be satisfied by lower levels of scheduled service. As a more affordable option, in many cases, people address these barriers in lieu of paying more to utilise ride-hailing, unless the trip is discretionary or atypical. In such cases, ride-hailing becomes more appealing. These sorts of insights help demonstrate the appeal to some public transport agencies of select partnerships with TSP.

### **Sentiment reflects level of service.**

Public and private mobility service vendors need to appreciate that timeliness, convenience and reliability are all factors influencing perception of service and sense of satisfaction. Maintaining and enhancing service standards should be part of core public sector policy (and generally is, in both Greater Perth and King County).

### **Sentiment can help explain some of the current transport trends in Greater Perth and King County.**

The study highlighted the importance of time-budget trade-offs and emphases on levels of service. The various investments in public transport in King County, including in fixed and ancillary services, are likely to have improved the balancing equation in favour of public transport.

Many cities make the mistake of focusing on 'customer experience noise', such as public WiFi and better branding and wayfinding. These are important components to the overall transit experience but the primary service investments should be in service and in capital. The number one determinant of people's willingness to take transit more often is service.

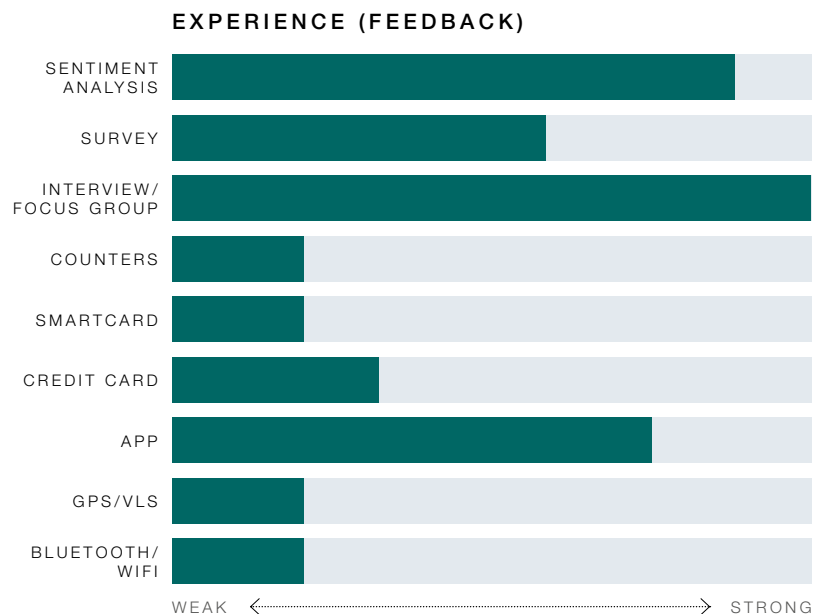


Based on our prototyping exercise, **Figure 34** illustrates the relative strength of feedback we believe that sentiment analysis leverages from the travelling public relative to other methods. More generally, sentiment analysis appears to be a useful supplement to other data collection methods; ideally, those that quantify public transport and TSP ridership statistics. It therefore neither replaces the need for industry to establish more uniform standards relating to data sharing nor sophisticated queries of data when it is made available.

Our appraisal of the key strengths and limitations of sentiment analysis, applied as we have, is as follows. These, alongside our propositions for further work, represent key conclusions from our applied research.

### Strengths and potential applications

- Certain qualitative datasets can be compiled at-scale, which helps identify trends, risks and opportunities
- Change-over-time can be understood better, particularly in terms of how the travelling public reacts to changes in levels of service. The prototype conducted by our research team analysed data captured over a year-long period, which is insufficient to effectively gauge changes to levels of service unless they are significant (order-of-magnitude) in busy areas of cities or across a city/ region-as-a-whole. These events could include major service disruptions (e.g. a line closure), new service types or competitors entering the mobility services market and/ or significant ticket price changes
- Cross-sectional comparisons are enabled of differences in transport trends and how these are appraised by the travelling public
- In time, methodological refinement (including growing sophistication of the search and analytics algorithms as well as user expertise), will add to the applicability of the technique. Furthermore, a growing dataset will permit better inferences to be drawn as well as increasing predictive capability (especially in advance of known or likely events), to help shape responses and management practices
- Triangulating findings from other methodologies, particularly those unable to shed significant light on the decision-making of the travelling public, and their appraisal of services offered based on factors such as levels of service, time-of-day and location



- Avoiding introduction of surveyor/ framing and some recall bias into the data collection procedure and enabling iterative research process, rather than analytics and insights being constrained by a static and/ or single-point survey. Furthermore, processing can be rerun using adjusted parameters, if on review, the researcher (or a peer-review) considers that the algorithm has introduced bias
- Facilitating capture of relevant feedback from an increasingly broad sample of the travelling public and deep-and-wide dataset, leading to greater internal consistency of sentiment (e.g. omission of outlying responses from commentators who make multiple posts, to appraise stronger trends).

### Limitations

- Commentators generally share content relating to more extreme or severe experiences (especially negative experiences), which skews the data. Nevertheless, transport planners and policy makers, again with a large dataset, can observe changes over time and the specific polarity of content to gain insights into factors such as satisfaction and potential behaviour change. These insights can be compared-and-contrasted with both other datasets and similar data from other contexts

**▲**  
**Figures 34**  
Comparison of different data collection methods in terms of theoretical strength of user experience feedback



A reporting bias exists where people are more likely to reach out when they have had a negative experience and less likely when the experience was positive.

- People are more inclined to comment when change occurs. Sentiment analysis is therefore better suited to measuring reactions to change than for day-to-day or habitual decisions. The potential power of the technique to gauge and in time, predict the impacts of changes, may well have merit
- Long-run data is needed to gauge change to habits: some negative sentiment may simply illustrate common frustration
- Sentiment cannot be used reliably to quantify mode use and the typical lag between experience and posting makes geo-location difficult. The technique therefore offers more a means to understand the why, what and when of travel behaviour but much less so regarding where (with much geographic specificity) and how much. While there is potential to apply this technique to help answer questions about where services could be improved, this is more likely to be possible at the mesoscopic level
- Rhetorical devices and complex language (e.g. sarcasm/ irony/ memes) are very difficult to analyse correctly applying existing algorithms. Although this difficulty applies also to traditional techniques, oral surveys and focus groups do allow researchers to clarify meaning directly with the participant. This limitation on sentiment analysis is mitigated through manual analysis, which can become labour-intensive and requires the analyst to have sufficient domain expertise
- Some framing, researcher and recall biases can be avoided but analyst and processing biases remain. While these may be mitigated over time as the technique becomes more sophisticated, the technology and the analyst are required to contextualise sentiment. Such limitations are mitigated further through aggregation of more data, heuristic analyses and triangulation with other information. It is possible that reliability and validity limitations are ameliorated over time following sufficient cycles and machine learning

- Those without the internet or social media are not captured and personality types more inclined to post online are captured disproportionately compared to those that do not post as often
- ‘One mention’ does not constitute consistency in mode choices or experience.

Additionally, the methodology can enable triangulation of PII, such as socio-demographic and contact details, even if all data is from public sources. This is something requiring appropriate management and control, and undoubtedly, civil and criminal law will continue to evolve to address this. The public and private sectors must avail themselves of statutory (and ethical) boundaries.

#### Further work

In many respects, the opportunities above represent the bases for further work. Each of the specified applications needs to be prototyped for veracity. The less onerous, resource-intensive prototypes include further cross-sectional analyses as well as gauging impacts on sentiment associated with disruptive events, such as a major line closure or service change (including triangulating sentiment data with other explanatory information such as ticketing/ smartcard data).

Ultimately, as Cedric Price may have observed, more data neither translates directly to more insight nor value-for-money. Further prototyping should be subject to specific framing of questions that this technique - sentiment analysis - can help to answer in efficient and effective ways. Such creative application is all-the-more important as the mobility landscape changes and industry continues to struggle to resolve much broader data ownership, formatting, access and value questions.



## List of Acronyms

API	Application Programming Interface
DoT WA	Department of Transport, Western Australia
GPS	Global Positioning System
GTFS	General Transit Feed Specification
IoT	Internet-of-Things
ISO	International Standards Organisation
LDA	Latent Dirichlet Allocation (algorithm)
MA	Massachusetts (State of, US)
MaaS	Mobility-as-a-Service
MAPC	Metropolitan Area Planning Council (Boston, MA)
MBTA	Massachusetts Bay Transportation Authority
MDS	Mobility Data Specification
NACTO	National Association of City Transportation Officials
NHTS	National Household Travel Survey (US)
NPS	Net Promoter Score
NSW	New South Wales (State of, Australia)
P2P	Person-to-Person/ Peer-to-Peer
PARTS	Perth And Regions Travel Survey
PII	Personally Identifiable Information
SDOT	Seattle Department of Transportation
TfNSW	Transport for New South Wales
TNC	Transportation Network Company
TSP	Transport Service Provider
UX	User Experience
VKT/ VMT	Vehicle Kilometres Travelled/ Vehicle Miles Travelled
VIC	Victoria (State of, Australia)
WA	Washington (State of, US)

# References

- [1] Ho C., Hensher D., Mulley C. and Wong Y. (2018) *Potential uptake and willingness-to-pay for Mobility as a Service (MaaS): A stated choice study*. Transportation Research Part A, 117, pp302-318
- [2] IHS Markit (2017) *More than Six Billion Smartphones by 2020, IHS Markit Says*. January 18, 2017: <https://news.ihsmarkit.com/press-release/technology/more-six-billion-smartphones-2020-ih-s-markit-says>
- [3] We Are Social (2018) *Digital in 2018: World's Internet Users Pass the 4 Billion Mark*. January 30, 2018: <https://wearesocial.com/blog/2018/01/global-digital-report-2018>
- [4] Worldometers (2019) *Current World Population*. <https://www.worldometers.info/world-population/>
- [5] Reed, Brad (2010) *A Brief History of Smartphones*. June 18, 2010: <https://www.pcworld.com/article/199243/a-brief-history-of-smartphones.html>
- [6] Piotrowski S. and Wallis I. (2018) *Recent Public Transport Patronage Trends in Perth*, paper presented at Australian Institute of Traffic Planning and Management national conference, July 24-27, Perth, Australia
- [7] Deloitte Access Economics (2016) *Economic Effects of Ridesharing in Australia*, report prepared for Uber <https://www2.deloitte.com/content/dam/Deloitte/au/Documents/Economics/deloitte-au-economics-economic-effects-of-ridesharing-australia-150216.pdf>
- [8] Deloitte Access Economics (2016)
- [9] Deloitte Access Economics (2016)
- [10] Schwieterman J.P., Livingston M. and van der Slot S. (2018) *Partners in Transit: A review of partnerships between Transportation Network Companies and Public Agencies in the United States*, August, Chaddick Institute for Metropolitan Development at DePaul University
- [11] Schwieterman et al. (2018)
- [12] Bliss L. (2019) *Uber Wants to Be Your One-Stop Transit Stop*, January 31 <https://www.citylab.com/transportation/2019/01/uber-app-map-directions-public-transportation-denver/581725/>
- [13] Bliss L. (2019) *'Uber Was Supposed To Be Our Public Transit'*, April 29 <https://www.citylab.com/transportation/2019/04/innisfil-transit-ride-hailing-bus-public-transportation-uber/588154/>
- [14] Cortright J. (2019) *Angie's List: The problem isn't ride hailing, it's the lack of road pricing*, February 7 <http://cityobservatory.org/angies-list-the-problem-isnt-ride-hailing-its-the-lack-of-road-pricing/>
- [15] Schaller B. (2018) *The New Automobility: Lyft, Uber and the future of American cities*, July, New York: Schaller Consulting
- [16] Bliss L. (2019) *Another Study Blames Uber and Lyft for Public Transit's Decline*, January 24 <https://www.citylab.com/transportation/2019/01/uber-lyft-ride-hailing-impact-public-transit-ridership/581062/>
- [17] Gehrke S., Felix A. and Reardon T. (2018) *Fare Choices: A survey of ride-hailing passengers in Metro Boston*, Report #1, February, MAPC
- [18] Schaller (2018)
- [19] Schaller B. (2017) *Unsustainable? The growth of app-based ride services and traffic, travel and the future of New York City*, February, New York: Schaller Consulting
- [20] Schaller (2018)
- [21] Henao A. and Marshall W. (2018) *'The impact of ride-hailing on vehicle miles traveled' in Transportation* <https://link.springer.com/article/10.1007/s11116-018-9923-2>
- [22] Henao A. and Marshall W. (2019) *'The impact of ride hailing on parking (and vice versa)'* in *The Journal of Transport and Land Use*, 12(1), pp127-147
- [23] Clewlow R. and Mishra G. (2017) *Disruptive Transportation: The adoption, utilization, and impacts of ride-hailing in the United States*, Research Report UCD-ITS-RR-17-07, October, Davis: Institute of Transportation Studies, University of California Davis
- [24] Schaller (2017)
- [25] Falconer R., Felder M. and Zhou T. (2018) *Mobility-as-a-Service: The value proposition for the public and our urban systems* Arup and MaRS Discovery District.
- [26] West Yorkshire Combined Authority and Arup (2017) *The Millennial City Region for All: Leeds City Region's journey towards future mobility*, UK
- [27] Deloitte Access Economics (2016)
- [28] Crispin S. (2014) *UberX Launches Illegally*, October 9 <https://www.businessnews.com.au/article/uberX-launches-illegally>
- [29] Australian Bureau of Statistics (2017) *Region Summary: Greater Perth, Western Australia*. Accessed 21-08-2018. [http://stat.abs.gov.au/itt/r.jsp?RegionSummary&region=5GPER&dataset=ABS\\_REGIONAL\\_ASGS&geoconcept=REGION&measure=MEASURE&datasetASGS=ABS\\_REGIONAL\\_ASGS&datasetLGA=ABS\\_NRP9\\_LGA&regionLGA=REGION&regionASGS=REGION](http://stat.abs.gov.au/itt/r.jsp?RegionSummary&region=5GPER&dataset=ABS_REGIONAL_ASGS&geoconcept=REGION&measure=MEASURE&datasetASGS=ABS_REGIONAL_ASGS&datasetLGA=ABS_NRP9_LGA&regionLGA=REGION&regionASGS=REGION)
- [30] United States Census Bureau (2018) *QuickFacts King County, Washington*, <https://www.census.gov/quickfacts/kingcountywashington>
- [31] Australian Bureau of Statistics (2017)
- [32] Newman, P. (2011) *The Perth Rail Transformation: Some Political Lessons Learned*, Fremantle: CUSP.
- [33] Stephenson, G., and Hepburn, J. ed. (1955) *Plan for the Metropolitan Region Perth and Fremantle*, Perth: Govt. Printing Office.

- [34] Gregory, J. (2012) *Stephenson and Metropolitan Planning in Perth*, *Town Planning Review* 83 (3): 297-318. doi: 10.3828.
- [35] Piotrowski and Wallis (2018)
- [36] TTF Australia and LEK (2019) *Are We There Yet? New mobility readiness scorecard*, March
- [37] United States Census Bureau (2018) *Census Bureau Reveals Fastest-Growing Large Cities*, May 24 <https://www.census.gov/newsroom/press-releases/2018/estimates-cities.html>
- [38] United States Census Bureau (2018)
- [39] CityLab (2017) *How Seattle Bucked a National Trend and Got More People to Ride the Bus*, October 16 <https://www.citylab.com/transportation/2017/10/how-seattle-bucked-a-national-trend-and-got-more-people-to-ride-the-bus/542958/>
- [40] Sound Transit (2019) *Who we are*, <https://www.soundtransit.org/get-to-know-us/who-we-are>
- [41] The Seattle Times (2019) *As the snow melts, King County Metro running more bus routes Wednesday*, February 12 <https://www.seattletimes.com/seattle-news/transportation/as-the-snow-melts-king-county-metro-will-run-90-percent-of-its-bus-routes-wednesday/>
- [42] Washington State Department of Transportation (2017) *Ferry ridership surges to highest level in a decade*, January 18 <https://www.wsdot.wa.gov/news/2017/01/18/ferry-ridership-surges-highest-level-decade>
- [43] Data USA (2016) *Data USA: Seattle, WA & King County, WA*, <https://datausa.io/profile/geo/seattle-wa/?compare=king-county-wa#top>
- [44] TransitCenter (2019) *Transit Insights*, accessed March 2019 <http://insights.transitcenter.org/>
- [45] Schmitt A. (2019) *How Two Cities Actually Reduced Driving*, February 8 <https://usa.streetsblog.org/2019/02/08/minneapolis-and-seattle-have-achieved-the-holy-grail-for-sustainable-transportation/>
- [46] Schmitt (2019)
- [47] Seattle Design Commission (2019) *Sound Transit 3*, <https://www.seattle.gov/designcommission/project-reviews/current-project-reviews/sound-transit-3>
- [48] Schaller B. (2019) *In a Reversal, 'Car-Rich' Households are Growing*, January 7 [https://www.citylab.com/perspective/2019/01/uber-lyft-make-traffic-worse-more-people-own-cars-transit/579481/?utm\\_content=edit-promo&utm\\_medium=social&utm\\_source=twitter&utm\\_campaign=citylab&utm\\_term=2019-01-07T12%3A39%3A07](https://www.citylab.com/perspective/2019/01/uber-lyft-make-traffic-worse-more-people-own-cars-transit/579481/?utm_content=edit-promo&utm_medium=social&utm_source=twitter&utm_campaign=citylab&utm_term=2019-01-07T12%3A39%3A07)
- [49] Wanek-Libman M (2019) *Via brings microtransit to southeast Seattle and Tukwila*, April 18, <https://www.masstransitmag.com/alt-mobility/shared-mobility/car-sharing/article/21076968/via-brings-microtransit-to-southeast-seattle-and-tukwila>
- [50] Schaller (2018)
- [51] Schaller (2018)
- [52] Gutman D. (2018) *How Popular are Uber and Lyft in Seattle? Ridership numbers kept secret until recently give us a clue*, November 5 <https://www.seattletimes.com/seattle-news/transportation/how-popular-are-uber-and-lyft-in-seattle-ridership-numbers-kept-secret-until-recently-give-us-a-clue/>
- [53] Seattle Department of Transportation (2017) *New Mobility Playbook v1.0*, September, SDOT
- [54] Australian Bureau of Statistics (2017)
- [55] United States Census Bureau (2018)
- [56] Australian Bureau of Statistics. 3218.0 – Regional Population Growth, Australia, 2017-2018 (27 March 2019). Accessed 22-04-2019. <https://www.abs.gov.au/ausstats/abs@.nsf/latestProducts/3218.0Media%20Release12017-18>
- [57] United States Census Bureau (2018)
- [58] Australian Bureau of Statistics (2016) *2016 QuickStats: Greater Perth, Western Australia*. Accessed January 2019. (extrapolated from weekly MHI) [http://quickstats.censusdata.abs.gov.au/census\\_services/getproduct/census/2016/quickstat/5GPER?opendocument](http://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/quickstat/5GPER?opendocument)
- [59] US Census Bureau. *American Community Survey 1-year estimates*, 2016. MOE +/- \$1584. Summary table S1903 for King County, WA.
- [60] Australian Bureau of Statistics (2017)
- [61] United States Census Bureau (2018)
- [62] Australian Bureau of Statistics (2016)
- [63] United States Census Bureau (2017) *2013-2017 American Community Survey 5-year Estimates*. Table B08301.
- [64] United States Federal Transit Administration National Transit Database [NTD] (2017) *Annual Database Service*. <https://www.transit.dot.gov/ntd/data-product/2017-annual-database-transit-way-mileage>
- [65] Western Australia Public Transport Authority (2018) *Transperth Performance*, July 2017-2018. <http://www.pta.wa.gov.au/about-us/priorities-and-performance/transport-performance>
- [66] United States Federal Transit Administration (2017)
- [67] Western Australia Public Transport Authority (2017). *Transport performance: Transperth performance 2012-2017* <http://www.pta.wa.gov.au/about-us/priorities-and-performance/transport-performance>



- [68] United States Federal Transit Administration (2017)
- [69] Transperth (2018) *Annual Report, 2017-2018*. <http://www.pta.wa.gov.au/annualreport2018>
- [70] United States Federal Transit Administration (2017)
- [71] Transperth (2019) *Annual Reports*. <https://www.pta.wa.gov.au/about-us/priorities-and-performance/annual-reports>
- [72] US Federal Transit Administration (2017).
- [73] Government Office for Science (2019) *A Time of Unprecedented Change in the Transport System*, January, London: Government Office for Science [GOS]
- [74] Inbakaran, C. and Annette Kroen. (2011) *Travel Surveys – Review of international survey methods*. Australasian Transport Research Forum 2011 Proceedings, 28-30 September 2011, Adelaide, Australia
- [75] Mir Hussain Talpur et al. (2012) *Transportation Planning Survey Methodologies for the Proposed Study of Physical and Socio-economic Development of Deprived Rural Regions: A Review*. Canadian Center of Science and Education: Modern Applied Science; Vol 6, No. 7. <http://dx.doi.org/10.5539/mas.v6n7p1>
- [76] Collins, A, Rose, JM and Hess, S (2012) *Interactive stated choice surveys: a study of air travel behaviour*. Transportation, 39 (1). 55 - 79 . ISSN 0049-4488
- [77] Comendador, J. and Lopez-Lambas, M.E. (2016) *Pain and Joy of a Panel Survey on Transport Studies*. Transportation Research Procedia 18 (2016) 248-255.
- [78] Hubrich, S., and Wittwer, Rico. (2014). *Household or individual – Advantages and disadvantages of different interview selection strategies*. Technische Universitat Dresden.
- [79] Lo Iacono, Valeria, Symonds, Paul and Brown, David H.K. (2016). *Skype as a Tool for Qualitative Research Interviews*. Sociological Research Online 21(2)12 <http://www.socresonline.org.uk/21/2/12.html>
- [80] Ma X., Liu C., Wen H., Wang Y. and Wu Y. (2017) *Understanding commuting patterns using transit smart card data*. *Journal of Transport Geography*, 58, pp135-145
- [81] Sawers P. (2019) *Microsoft, Moovit, and TomTom Team up for Multi-modal Transport Platform*, February 12 <https://venturebeat.com/2019/02/12/microsoft-moovit-and-tomtom-team-up-for-multi-modal-transport-platform/>
- [82] Kaye K. (2019) *Portland quietly launches mobile location data project with Alphabet's controversial Sidewalk Labs*, May 28 <https://www.geekwire.com/2019/Portland-quietly-launches-mobile-location-data-project-alphabets-controversial-sidewalk-labs/>
- [83] Government Office for Science (2019)
- [84] Transport Systems Catapult (2016) *Mobility as a Service: Exploring the opportunity for mobility as a service in the UK* <https://ts.catapult.org.uk/wp-content/uploads/2016/07/Mobility-as-a-Service-Exploring-the-Opportunity-for-MaaS-in-the-UK-Web.pdf> [CAT]
- [85] House of Commons Transport Committee (2018) *Mobility as a Service*, Eight Report of Session 2017-2019, HC 590, House of Commons, UK
- [86] House of Commons Transport Committee (2018)
- [87] Transportation for America, *Shared Micromobility Playbook* <https://playbook.t4america.org/>
- [88] Gehrke et al. (2018)
- [89] Greenfield J (2019) *Emanuel: Chicago Is Now the First U.S. City to Publish Detailed Ride-Hailing Data*, Streetsblog, April 12 <https://chi.streetsblog.org/2019/04/12/emanuel-chicago-is-now-the-first-u-s-city-to-publish-detailed-ride-hailing-data/>
- [90] Gutman D. (2018) *Uber and Lyft May Have to Disclose Seattle Data They Claim Secret, Supreme Court Rules*, May 31 <https://www.seattletimes.com/seattle-news/transportation/uber-and-lyft-may-have-to-disclose-data-they-claim-secret-supreme-court-rules/>
- [91] Seattle Department of Transportation (2017)
- [92] Government Office for Science (2019)
- [93] SAE International (2019) *SAE International Brings Together Public and Private Partners to Address Mobility Data-Sharing Principles*, May 21 <https://www.sae.org/news/press-room/2019/05/sae-international-brings-together-public-and-private-partners-to-address-mobility-data-sharing-principles>
- [94] National Association of City Transportation Officials (2017) *City Data Sharing Principles: Integrating new technologies into city streets* <https://nacto.org/wp-content/uploads/2017/01/NACTO-Policy-Data-Sharing-Principles.pdf>
- [95] National Association of City Transportation Officials (2017)
- [96] Pyzyk K. (2019) *Los Angeles' Mobility Data Specification Unlocks Potential for Data Management*, January 24 <https://www.smartcitiesdive.com/news/los-angeles-mobility-data-specification-dockless-rideshare/546692/>
- [97] Transport for NSW (2019) *TfNSW MaaS Data Sharing Specification (TMDSS)*, Draft v0.6, February [https://opendata.transport.nsw.gov.au/sites/default/files/TfNSW%20MaaS%20Data%20Sharing%20Specification\\_v0.6.pdf](https://opendata.transport.nsw.gov.au/sites/default/files/TfNSW%20MaaS%20Data%20Sharing%20Specification_v0.6.pdf)
- [98] Transport for NSW (2019)

Appendix A  
Summary of expert interviews



# Approach

Interviews – all of those in Perth, face-to-face and those in Seattle, by teleconference – followed a semi-formal structure design to elicit interviewee feedback on the following questions:

1. What are today's major mobility/ transportation challenges and opportunities?
2. How does your agency/ company view the role of transportation in improving quality of life?
3. What are the relative value propositions of mobility services and transit (e.g. convenience, price, comfort)?
4. What data sources provide insights into ride-hailing and transit use, and what are their limitations?
5. What do you think are the most important things to improve with regard for user experience?
6. What are barriers to achieving better user experience right now?
7. How can industry build a better collective knowledge base regarding mode choice?

All interviews were conducted in 2018.



# 1. What are today's major mobility/ transport challenges and opportunities?



## Challengers

- In the context of Perth, the decline in public transport patronage – in terms of both personal mobility options available and government revenue, as well as understanding the causes
- Defining and addressing gaps in the market for travel
- Understanding the role of mobility services and reasons for increased uptake in Perth compared to other states
- Congestion and the fundamental causes
- Uncertainty around the timing of coming changes and awareness of what is coming. Western Australia is a unique environment. It is hard to see how changes will affect the State until service offers broaden out: E.g. bike share and car share
- Current prototypes in other states (e.g. New South Wales) are tending to be relatively expensive, discouraging further trials on a cost-per-trip basis
- Developing appropriate responses to changes based on big-picture thinking. Understanding the role of stakeholders in mass transit going forward. Current thinking is that the Public Transport Authority in Western Australia can work with providers of mobility services but not necessarily as a partner or facilitator. Other State agencies may view things differently
- Future accessibility
- Environmental impacts (including user perceptions of these and political responses)
- Public transport effectiveness and efficiency
- Increased transport demand due to growing cities
- Understanding and applying systems approaches.

## Opportunities:

- MaaS and the range of value propositions it provides: Additional choice, private sector opportunities, filling gaps in transport, cost savings, meeting travel needs, transport pricing opportunities, responding more directly to public need to help people shift modes
- Leveraging new technology to provide improved service, through new applications and use cases
- New space for funding mechanisms due to reduction in fuel excise funding
- Increased mobility for those currently mobility-impaired
- New business models
- Changing consumer preferences and the opportunity to respond to those
- Big data/ predictive analytics. A challenge and an opportunity is the mass collection and analysis of valuable data. It is essential for successful operation in a public-private ecosystem
- Creating lasting successful utilities, through good and fit-for-purpose facilities, maximising operations, appropriate policy frameworks and managing demand
- Creation of additional mode choice/ transport ecosystem involvement for the public.



GOVERNMENT  
King  
County

Answers to this question can only be answered context-specifically, with a lens more globally about what is going on in the tech space.

**Challenges:**

- Equity, and how services are deployed and defined – and whether these services respond to the actual needs of customers (of all backgrounds)
- Digital equity - getting tools in the hands of people who do not currently have them (e.g. devices, connectivity, digital literacy)
  - particularly with the explosion in variety and ridership of services and the non-uniform manner in which it is occurring. It is happening more for white males with an income over \$100,000

- Ensuring government, particularly at a local level, is able to ingest data from providers as well as trend data on how customers view and use the transport system. Cities need to also be able to maintain privacy and preserve the integrity of individual information, and present their insights in a digestible manner to clearly convey their metrics and evaluation of mobility services.



Mobility  
Service  
Vendor

**Challenges:**

- A growing population such as that in Perth, combined with a heavy reliance on private vehicle travel, creates challenges in terms of congestion
- Since 2000, private vehicle ownership has increased by 43% in Australia
- Congestion and pollution are significant challenges to infrastructure capacity and liveability. TSP strive to tackle this by getting more people into fewer cars
- The current model of car ownership is entrenched into street design, making it difficult to get people out of private vehicles
- City policies also present a challenge; particularly those that impact land use and development: E.g. parking ratios; free and under-priced parking
- The use of streets for personal vehicle storage instead of activity.

**Opportunities:**

- Shifting from ownership to ridership
- Enabling mobility across multiple modes.

## 2. How does your agency/ company view the role of transportation in improving quality of life?



- We view trying to improve quality of life through transportation as the reason for our existence. We have a fundamental community service obligation – provision of a social good, especially for those without a private car. Provision of a reasonable universal level of service
- We should be acting in terms of systems and interactions rather than in isolation. No one agency is able to provide transport outcomes by itself, but we keep trying
- Transportation's role lies in accessing opportunities, and being able to move around and build a connection to and within society as a result
- Using developing technology and services to meet expectations of customers
- Comfort and protection; safety and privacy
- Setting a vision for mobility in the future, and resilience so we can respond appropriately to what is coming
- It is something the government can facilitate more than the private sector.



- Transportation is a fundamental component and determinant of peoples' quality of life
- In cities like Seattle, transportation and housing are at the centre of the affordability and displacement crisis
- Within the new mobility branch of SDOT and the broader mobility options and transit team, staff are searching for a way to provide more affordable and sustainable options. This includes improving the transit experience in particular, at the expense of single-occupancy drivers, so people not only have cheap options but also sustainable options.



- Transport is an important factor in social mobility
- Ridesharing can go a long way towards filling gaps: E.g. where there are limited alternatives to driving a private car; people who cannot drive due to various factors
- Affordable, safe and reliable transport is essential for connecting people to community, jobs, services and recreation
- Innovations like pool options can help solve congestion and reduce the number of private cars on the road
- Traditionally, access to transportation is enabled only by the city, or through a specific high-capital investment: E.g. car purchase
- A number of factors go into transport decision-making; Where you live, how you get there, what storage you have and so on
- MaaS enables the consumer to make individual economic choices on a trip-by-trip basis rather than in long-term periods: E.g. buying a car
- Having more options allows consumers to make better economic decisions, which makes them happier with their transport purchases
- Enabling transport (e.g. to bars at night when transit does not provide services), enables people to spend more money on local businesses, explore their city and support economic growth without the major capital investments associated with traditional transportation
- When people buy a private car, due to the high initial costs involved, they are more inclined to use it even when it might not be the most effective option.



### 3. What are the relative value propositions of mobility services and transit (e.g. convenience, price, comfort)?



GOVERNMENT  
Greater  
Perth

#### For private mobility services:

- Revenue
- Using under-utilised assets
- Meeting customers' needs and expectations, even as they change
- After-hours operation
- Social journeys
- Infrequent and special-purpose trips, and responsiveness are important elements of use
- TSP can be agile and dynamic in responding to customers through data analysis. They can collect, analyse and respond very quickly. A good example is when TSP vehicles were banned from Optus Stadium when events were on post-opening. They responded quickly with discounted rates to drop people nearby
- In the past, transport habits were relatively consistent. Today, they are much more dynamic and different journeys are more complex. Simplistic models of bus services and roads doing the same thing every day do not work as well anymore compared to the dynamic nature of private mobility services
- Taxis are often expensive, and drivers do not tend to be the most customer-focused people. Ubers have a good reputation in Perth, a lot of which is due to the star rating.

#### For public transport operators and authorities:

- Ensuring Perth metropolitan area accessibility
- Frequency of services
- Lower cost option
- Meeting equity and safety expectations
- Rigid; habitual; set to schedule
- Price-point
- Satisfaction monitors tell us that people do not like waiting but they are comfortable once they are on-board
- Transperth has been monitoring passenger service for around 27 years, and it involves about 14,000 passengers. Ratings have never been better.



GOVERNMENT  
King  
County

- The two primary decision factors are cost and convenience
- Companies are doing their best to sell additional determinants and value-adds, which will probably play into some peoples' decisions, like airlines providing food choices and entertainment options, but in reality, people just want to get where they want to go as cheaply, safely and quickly as possible.



- Ride-sharing complements other transport options
  - Ride-sharing services can extend the reach of public transport and help connect individuals and opportunities, people and jobs within and across metropolitan areas
  - Ride-sharing can provide an affordable, reliable and safe alternative travel option, which can address the last-mile problem when people choose to drive if any part of their trip is outside walking distance
  - For TSP, the two critical components of transportation are price and time: What is the opportunity cost of this trip? How much is my time worth to me?
  - With other modes it can be a less obvious decision about the opportunity costs
  - The questions TSP anticipates are: How much is time worth? How long will it take? How much will it cost? How are these costs relative to each other? How many legs of the journey are there on different modes, and do I have control over some or any of these?
  - There is likely an economic value people assign to ideas of control or certainty, too
- In transit-rich areas, TSP can yield value by providing service where transit is not operating at optimal levels: E.g. BART in the San Francisco Bay Area, where customers buy a ticket, move downstairs to the platform, and have already spent their money without knowing when the next train will arrive since there is no sign outside telling you when the next one is
  - TSP other value proposition is filling the gaps in the transit network when there are critical changes. In Seattle, there are some specific geographic constraints and a lot of construction: E.g. the Seattle viaduct will close in January for six weeks. TSP can respond more dynamically and flexibly than traditional transportation networks
  - Seattle is very hilly. If a customer has mobility issues or is toting goods or children, they can be serviced in a way that may be more convenient than transit
  - Transit still wins in Seattle when it is done well - highly useable service - and invested in.
-

## 4. What data sources provide insights into ride-hailing and transit use, and what are their limitations?

### GOVERNMENT Greater Perth

#### Data sources:

- Roy Morgan study; public data; market research
- Meeting with Uber; info straight from the vendor
- Shofer has services aligned with coming transport on-demand Bill and indicated willingness to share data
- Transit – SmartRider via Planning and Transport Research Centre assessment; Census Journey-to-Work; published articles; published Public Transport Authority data
- Auditor General report – planning and management of bus services
- Not aware of any publicly-available TSP datasets
- The Transperth view is that TSP are not drawing away a lot of ridership from transit. Rather, more taxi or car trips are being reassigned, or journeys are being completed that would not have been taken, otherwise

- The recent drop-offs in public transport ridership have related to full-fare journeys. CBD trips are right down, which can skew results. This actually relates to a positive passenger experience, as it creates additional capacity and therefore ability to sit down
- SmartRider is a very rich dataset but requires specific queries. About 10 years of data is now available.

#### Limitations:

- Do not know demographics or trip purposes among travelling public
- No bike share data
- Cash trips on public transport – we do not get information from these
- Where trends are evident, there has not been enough deep digging to fully understand what they – the trends – mean: E.g. young people driving less – what are young people doing if they are driving less, but also taking public transport less?
- Factors influencing public transport trends.

### GOVERNMENT King County

- Data sources and limitations are interrelated, as the TSP have not proved to be the best partners
- Seattle has been sued and identified as a target of the companies
- Washington State has the most conservative public disclosure laws in the country, so mobility companies have fought the City to keep private the data submitted to the City as a condition of operation
- The legislation in Seattle requesting this data was enacted in 2014, so what it asks for (Zip-code-level data) is not very useful from a planning and operations perspective, but it is from a legislative perspective
- Useful data sources depend on what problem you are trying to solve and what your metrics are. They vary on place, type of issue, what the city's values are and so on

- A particular issue Seattle is trying to solve is that of downtown congestion directly related to TSP trips and circling in the peak. Data about deadhead time would be extremely useful, as would trip origin and destination with timestamps (at a distance level that protects privacy) but the companies are not willing to provide it
- Seattle is the first city in the country to establish regulations and a permit for free-floating bike share. The goal was to collect data on the services to understand how they operate, what the challenges are, and then to evaluate how it works how to use it. Seattle collected a robust dataset exploring service usage, how maintenance is invested, safety records, service area gaps and so on. SDOT produced [a report](#), which delves into the metrics of success and what the data revealed. As supply increased, demand increased, and areas formerly unserved by the dock-based system had an explosion in demand.





- TSP operates in more than 600 cities in 65 countries. Around 15 million trips happen each day. This scale provides a rich source of data and information, some of which is publicly available via the Movement tool, which uses anonymised, aggregated data from trips to measure changes in travel times
  - The tool was developed to assist policymakers and is free to the public. It can be used to compare travel times between locations on any number of parameters like the date, time of day, and whether it is a weekend or peak hour
  - TSP is open to working in partnership with governments and the private sector to improve access to transport and mobility services for the community
  - Demand – TSP can see when a person opens the app and considers hailing via TSP, and what percentage of people follow through
  - TSP can analyse issues in real-time– what was the trip ETA? Was there a functionality issue? What was the cost of the trip? This enables a better understanding of what causes people to take/ not take the ride
  - User experience (UX) is key to survival of TSP as a private company. Therefore, any and all information is needed regarding why customers do or do not use the service
  - It is equally important to address issues for drivers
  - TSP has strategically only rolled out MaaS in Santa Monica, where the GTFS feed is extremely reliable – the transit connections appear within the app, so they are tied closely to the brand. The accuracy of transit data is crucial to protect the brand’s reputation for accurate timing
  - TSP wants to be a transportation portal, as it expands into other modes via rolling out MaaS elsewhere. They want to be a tool for users to take other modes
  - TSP had a major product redesign to this end, so the app now asks where the user is going rather than where they are as a first step. Including real-time transit information in the app is a big part of offering the right service
  - A reporting bias exists where people are more likely to reach out when they have had a negative experience and less likely when the experience was positive
  - TSP employs sentiment analysis to determine what aspects of service people comment on whether it is positive or negative feedback, which feeds into product development through the Net Promoter Score (NPS) survey. Sentiment is a metric that a large section of the company is focused on enhancing and is gauged regularly.
-

## 5. What do you think are the most important things to improve with regard for user experience?

GOVERNMENT  
Greater  
Perth

- Seamless journeys, managing transfer penalties and so on
  - Multi-modal journey planning and consumer awareness of the options available
  - Having good information: E.g. how to catch public transport. Making the simple things easy and accessible
  - Transperth gives a good user experience based on purpose, cost structure and other attractions whereas TSP have a different value proposition so the user experience needs to differ: E.g. in terms of personalisation
  - In Perth, Transperth leverages convenience factor for a lot of trips – especially to low-parked areas
  - Messaging from customers is taken seriously but tends to be different every year: Cost of fares, on-board experience and so on. Transperth develops a business plan around these and requires responses from contractors.
- 

GOVERNMENT  
King  
County

- Investing in transit via additional service and transit priority. Growing transit ridership in Seattle is the result of significant investment in getting frequent transit to as many people as possible and in transit priority at intersections and along corridors
  - Many cities make the mistake of focusing on ‘customer experience noise’, such as public WiFi, and better branding and wayfinding. These are important components to the overall transit experience but the primary service investments should be in service and in capital. The number one determinant of peoples’ willingness to take transit more often is service
  - Organisation of options is crucial amidst the growing number of options. Private shared mobility options do not self-organise. Seattle has spent a lot of time, energy and resources on advancing the concept of mobility hubs, both as an organising framework for mobility, for automating the future, and to elevate the customer experience on transit. Access to and between modes for people in and around the hub area is not often considered. Much money and energy is spent on getting the mobility options to a city but not how those options interact and connect.
-



- Safety – of both passengers and drivers. Current techniques include:
    - Giving riders their driver's information in advance to help them verify the right car
    - Tracking each trip with GPS technology
    - Providing 24/7 support and investigating issues
    - Features like Real-Time ID Check, which help verify that the right driver is behind the wheel on an ongoing basis
    - Safety Centre for riders and drivers to access safety tips and learn about driver screening processes, insurance, and community guidelines
    - An emergency assistance button.
  - Delivering safety within a quality product is crucial. This is the motivation behind the two-way ratings system, the ability to share ETA and location with anyone at any point
  - Customer satisfaction is also served by the customer service hotline and the ability to communicate with the driver if a customer has lost something. Enabling people without smartphones to request rides is important. TSP has a web- and dispatcher-based service called Concierge, which mediates between TSP and non-smartphone-users, often at facilities like senior homes. There are different ways to pay for these rides, depending on the specific host facility. Service like Concierge opens this level of mobility to a whole new class of people. It is a similar scenario with people with hearing or vision impairments, and for non-emergency medical transit
  - TSP also pursues partnerships with transit. There are over 25 partnerships with cities or counties across the county to operate a variety of services: first/ last mile, job access/reverse commute, paratransit, supplemental suburban transit, late night services. In Seattle, they partner with Pierce County to provide first/last mile services to transit hubs. Each partnership is a little different based on needs.
-

## 6. What are barriers to achieving better user experience right now?

### GOVERNMENT Greater Perth

- Reliability – real-time tracking and being able to reflect service appropriately via the correct applications. User experience is extremely important; particularly, as customer expectations grow and change
- Personal understanding of the costs of choices
- Apps – these can overwhelm users by volume and complexity
- Lack of transparency relating to data and operations.

### GOVERNMENT King County

- Data is also very important. In Seattle, the city has an open data feed based on GTFS that indicates when the next bus will come. An app indicates what the routes are and offers trip planning. The problem is that the underlying data in the data feed is fundamentally flawed. If there is a disruption in the transportation network, the data offer as ‘real-time information’ can be anywhere from 25% to 75% off. Real-time information must be correct
- As people rely more on app-enabled mobility, TSP use of extremely accurate real-time data can sway people towards private mobility services and away from transit. Investing in true real-time data and piping it out to mapmakers is critical
- Seattle is at the leading edge of proving real-time data feeds on real-time conditions in the transportation network. Around 75% of the transit fleet is equipped with real-time tracking. There is some prediction that is not direct information, so it is incorrect when there are disruptions and Seattle is notorious for disruptions.

### Mobility Service Vendor

- Challenges surrounding such partnerships [with public transport operators] include integrating with real-time transit information, ascertaining how customers pay, and ensuring there’s a TSP vehicle waiting for customers at the other end of their trips.



## 7. How can industry build a better collective knowledge-base regarding mode choice?

### GOVERNMENT Greater Perth

- Industry cannot do it by itself. It must be by collaboration/ communication and partnership
- Open data for everyone. Not just government, but private companies need to share data, too. Uber and Lyft have recently committed to the SharedStreets platform, which is a great start.
- Events like Hackathons to promote awareness and use of open data
- Government needs to be sceptical with big data. What can we make it tell us? What is the question that needs answering? Where is the data coming from? Who is it helping? Also, who benefits from open data? The risk is nefarious use - so what are the motives?
- One sector cannot do it all
- Not playing catch-up
- Disruption comes quickly. We need to be ready for it and anticipate it. It is hard to innovate when you are playing catch-up and being challenged in lots of ways
- Relinquishing control
- Society demanding what we need. If the industry wants the government to do something different, we have to stand up and demand it. If we want better outcomes, we have to demand them
- Being vocal. The public need to stand up and demand action from the politicians, and industry needs to be vocal with government, the public stakeholders, other industry bodies, service providers or whoever it might be
- Academia could act as the middle man, brokering deals between the private sector and government. Universities are trusted intermediaries, and we should utilise their abilities to that effect if we can
- Different businesses should 'tend their own knitting' so there is not so much need to worry about bleed between public and private
- Government should regulate mass transit and leave private to fill the gaps. Government benefits from economies of scale and subsidy. It also has social KPI that the private sector struggles to replicate.

### GOVERNMENT King County

- Understanding the customer and how the customers' changing requires ongoing assessment of customer trends, customer attitudes, and global technology trends and how these affect peoples' decision-making
- Another factor is looking at the overarching investment portfolio for a City. How much money is being spent on sustainable transport modes in the city? How much is being spent by the City? How much is being invested in the private companies? How much money is being expended to grab more customers? There is a direct correlation between an increase in spending in the private sector, a decrease in spending in the public sector and transit ridership loss



- TSP technology is part of a broader ecosystem and it is always looking for ways to have a positive impact in communities
  - TSP is open to and is actively pursuing conversations and partnerships with government and industry on together making our cities stronger
  - TSP is concerned with whether they can start managing kerb space to improve the user experience and at what point ridesharing will be accepted as part of a transportation network, and something that can affect land use changes and optimisation
  - The Shared Streets platform is about taking a collaborative approach. It is important that participation extends beyond Lyft, Ford and Uber. Private competitors need to sign on as well and come up with something everyone can agree on, share, access, and use together
  - Transit agencies cannot share data on demand, in part because there are other responsibilities for them to manage. Shared Streets lets transit agencies make their own decisions, lets TSP store data in a way that is safe for them, and helps TSP make requests of agencies as a coalition and not a single company
  - Freight and commercial loading is a huge challenge when looking at kerb use and congestion. Those companies are difficult to bring to the table, but Shared Streets hopes to include FedEx and UPS.
-





