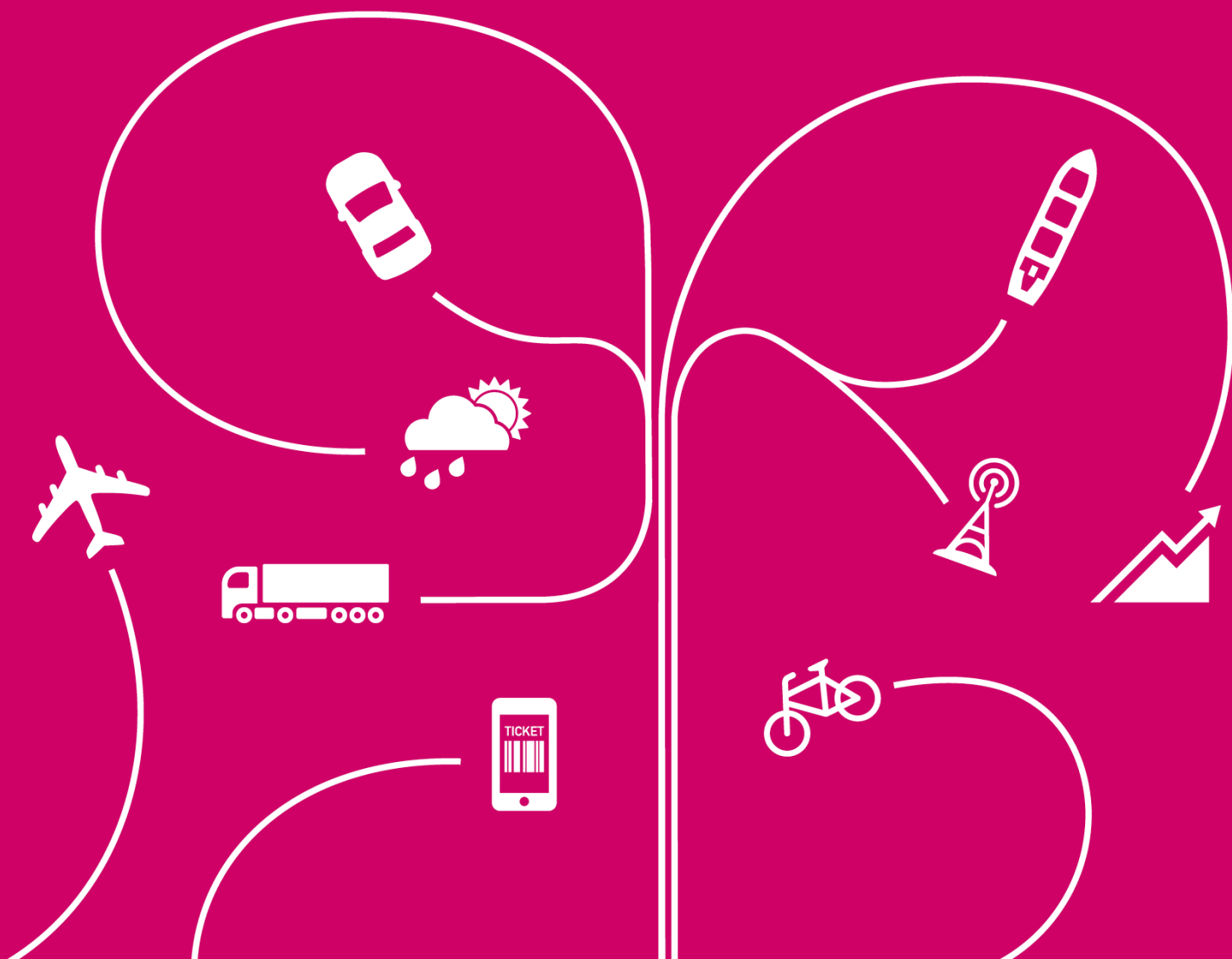


The Transport Data Revolution

Investigation into the data required to support
and drive intelligent mobility

March 2015



Acknowledgements

This report has been prepared for the Transport Systems Catapult by Integrated Transport Planning Ltd, together with the expertise of White Willow Consulting, Advancing Sustainability, and the Horizon Digital Economy Research Institute at the University of Nottingham. The Catapult would like to express their gratitude to the project team for undertaking this work, which consisted of:

- Neil Taylor, Ian Stott, Jon Parker, Jim Bradley – Integrated Transport Planning Ltd.
- Andy Graham – White Willow Consulting
- Chris Tuppen – Advancing Sustainability
- Jeremy Morley – Horizon Digital Economy Research Institute

This research involving staging a number of workshops exploring the data requirements for intelligent mobility. The Catapult would like to express their gratitude to following persons who kindly spared their time for the purposes of this study:

Frank Baxter – Southampton City Council

Roger Beecham – City University

Keith Bevis – Hertfordshire University

Nick Bromley – Greater London Authority

Neil Brown – Sustainable Environment

Dan Clarke – Connecting Cambridgeshire

Brad Cooper – DARTT

Mark Crosier – Prophecy Partners

Charlie Davies – Igeolise

Alistair Duke – BT

Mark Halsall – Northamptonshire County Council

Lee Hannis – Science and Technology Facilities Council

James Harris – Elgin

Neil Herron – Activ8 Virtual Parking Solutions

Sarah Gonslaves – Milton Keynes Council

Philip Kirk – Oxford Bus Company

Robery Lee – BAE Systems

Peter Lilley – Igeolise

Peter Miller – ITO! World

Vishnu Muralidharan – SBD

Lee Omar – Red Ninja Studios

Shane O'Neill – Elgin

Carl Partridge – Fat Attitude

Maurizio Pilu – Connected Digital Economy Catapult

Jonathan Raper – transportAPI

Jonathan Shewell-Cooper – ATOS

Phil South – Oxford Bus Company

Michael Walker – AIMES Grid Services

Simon William – Quantum Black

Andreas Zachariah – Carbon Diem

Executive Summary

The UK needs to move much faster in plugging various “open data” gaps, if the country is to realise its ambitions of being a global leader in the emerging Intelligent Mobility market. Defined as the smarter, greener and more efficient movement of people and goods around the world, Intelligent Mobility is a sector of the wider transport industry which is predicted to be worth around £900 billion a year by 2025. In order to play a leading role in this sector, however, the UK will have to address a number of issues in regards to data access, data handling and data skills – with data having been identified as “a new form of oil” for future transport systems.

The key findings of this report can be summarised as follows:

- The transport industry will be confronted with ever vaster amounts of data in the coming years as developments such as the Internet of Things rapidly expand the volume, velocity and variety of data related to transport and mobility;
- This report has identified 11 obvious transport-related data gaps – in some cases referring to datasets which do not yet exist at all in the UK, in other cases to datasets which exist only in ‘silos’ or which are not yet open or freely available;
- Standardisation of data remains an issue, with few private sector providers currently adhering to globally common data formats, and some forms of real-time data standards still to be defined;
- More has to be done to address public concerns over personal data privacy – identified as one of the key long-term threats to creating successful crowd-sourced products and services;
- UK investment is urgently needed to ensure sufficient levels of skilled technical talent, organisational capability and technology in the transport sector to handle the imminent surge in transport-related data.

The report finishes with 20 recommended actions that the UK can take in both the short- and long-term to address the issues and concerns that have been raised.

Methodology and key findings

This report sets out the findings from a practically-focused four-month review of the data required to support UK growth in the emerging Intelligent Mobility market. The report was commissioned by the Transport Systems Catapult which was created to drive and promote Intelligent Mobility – defined as the smarter, greener and more efficient movement of people and goods around the world.

The study was led by independent researchers at Integrated Transport Planning Ltd (ITP) in collaboration with experts from White Willow Consulting, Advancing Sustainability and the Horizon Digital Economy Research Institute. The research combined desk-based meta-analyses and workshops with organisations involved in the delivery of transport and data services. We consciously considered ‘more than just data’; since the capabilities associated with the sourcing, aggregation, fusion, and manipulation of complex, multi-source data are highly significant.

The investigation was framed around eight key questions:

- What mega-trends are shaping transport and mobility?

- What datasets are relevant to transport?
- How is transport-related data created and shared?
- How is transport data being exploited, and by who?
- Where is the future value in transport data exploitation?
- What data related capabilities are needed?
- What do roadmaps for intelligent mobility look like?
- What can the UK do to catalyse transport data exploitation?

The findings for each of these questions will now be summarised.

What mega-trends are shaping transport and mobility?

The Internet of Things¹ is predicted to contain 50 billion devices in 2020, rapidly expanding the volume, velocity and variety of data related to transport and mobility. This data is expected to become a ‘new form of oil’ for transport systems – charting demand for mobility, and enabling smarter coordination of increasingly ‘on-demand’ service that meet these demands.

Consequently the range of Open Data – public datasets free for re-use – is expanding, improving in quality, and becoming more dynamic through live feeds that are constantly updated. The increased range and scope of the datasets becoming available makes their coordination more challenging, and can even make it harder to discover and exploit high-value transport data feeds.

This is significant because in the future, all transport companies are also expected to be data companies – exploiting the ‘digital exhaust’ from their operations and customer interactions. Many of the major global data and technology companies are already investing in transport systems to explore whether they can provide enhanced mobility services. These will likely draw on the deep insights their analysis of customer data reveal on individual-level and aggregated travel intentions, actions, and purposes. The dominance of major technology companies and vehicle manufacturers means they are best-placed to crowd-source and exploit large datasets, and drive the globalisation of standard data formats that will support more user-focused and integrated transport systems.

Intelligent mobility itself is expected to become a mega trend. We identified a total of 20 services that may emerge over the next 10 years, and their emergence is anticipated in three stages:

- **2014-17:** Better integration of existing transport systems using sources of data that already exist, and becoming increasingly open or available at low-cost.
- **2018-21:** Optimisation across multiple transport networks through exploitation of archived data collected in real-time, to which predictive analytics are applied.

¹ The Internet of Things refers to the interconnection of uniquely identifiable (and addressable) embedded computing-like devices within the existing infrastructure of the internet.

- 2022-24: New forms of on-demand mobility for people and things requiring real-time data-sharing and interpretation, technological innovation, and legislative/public attitude shifts.

What datasets are relevant to transport?

We established a catalogue of over 200 transport-related datasets, structured around nine defined themes: Place & Space, Environment, People, Things & Movement, Disruption and event-related data, Public Transport Services, Personal Automobility, Freight connections, International Connections, and Consumption & transaction data.

A total of 19 datasets will likely drive the emergence of intelligent mobility: map data; weather; personal location data; network disruptions; planned events; real-time network capacity for people, vehicles & goods; public transport schedules; vehicle location data; fare and pricing data; sentiment data from service users and non-users; third party service usage data; and payment/transaction data.

Within these core datasets we identified 11 obvious transport-related data gaps which could be filled with varying degrees of effort. In some cases these datasets do not exist at all (e.g. on street parking bay availability), in others the data exist in silos (e.g. automated cycle count data, Urban Traffic Management and Control traffic flow data) or are not open/available (historic passenger ticketing data). Most commonly the datasets are only localised, rather than consolidated at a national level (e.g. databases of major events held by local authorities). In almost all cases the technical challenges to making these data available are secondary to data owners' attitudes, costs of establishing and maintaining sensor networks, in-house skills needed to support data sharing, and data privacy concerns.

How is transport-related data created and shared?

There are five primary mechanisms for data creation (manual collection, overt crowd-sourcing, covert crowd-sourcing, sensor-derived, service provider generated). Web-connected fixed and mobile sensors, plus crowd-sourcing are the emerging transport data collection mechanisms. In particular, personal/vehicle location data are the likely 'game-changers' for intelligent mobility.

The availability of personal/vehicle location-based datasets is being driven by the private sector, which plays a growing role in collecting and aggregating transport-related data. The increasing number of partial datasets creates a 'signal problem' – because we know less about what data being collected actually represent (e.g. in terms of sampled population, demographics). Data validation and curation therefore become very important, but few public sector data owners have resources to do this well. Key players such as Transport for London are leading the way, while many local authorities are still developing digital strategies. Some are replicating work being done by their neighbours, and might benefit from advice and guidance on maximising their investment in open data.

Transport data discovery and its uneven availability is becoming a major challenge:

- The multitude of catalogues and platforms needs to be embraced and documented.
- The best transport data feeds are concentrated at strategic national/major city levels.
- Almost no private sector data (even that you can buy) is documented in any form of catalogue.

Very little real-time data feeds are openly available, and that is what most app developers and intelligent mobility service providers are seeking. There is even less archived data that have been collected in real-time. This will be critical for predictive analytics based on past trends/events and scenarios (2014-2021,

and will need to become available if intelligent mobility is to evolve through to real-time service automation and optimisation (2021+).

Developers and service providers are increasingly interested in globally common data formats (e.g. GTFS, DATEX II in Europe), which make it easier to scale products and services internationally, yet few private sector providers adhere to them. Significantly, some forms of real-time data standard are yet to be defined (e.g. personal location-based data feeds).

How is transport data being exploited, and by who?

There are five types of transport data service: suppliers, aggregators, developers, enrichers, enablers). These will increasingly interact with each other, with intelligence/information & insight provision, and with tangible mobility service provision.

The aggregation and sale of raw open data is becoming a viable business, but the market for this data is still immature in relation to public transport data, with innovators essentially betting on future markets for these services. The automobile industry is very different, with more established end-to-end data value chains for traffic/disruption/roadworks/parking data. These are predicated on continued vehicle sales, into which these data services are bundled, and currently receiving investment through connected and autonomous vehicle innovation.

Personal information services are increasingly reliant on cross-subsidised business models driven by major technology companies. This is making the marketplace challenging for smaller providers, because users have a high 'expectation of free'. Partly as a result, half (50%) of iOS developers and even more (64%) Android developers, are operating below an "app poverty line" of \$500 per app per month. Public attitudes to personal data privacy and concentrations of market power, and legislation to curb any abuses, present long-term threats to crowd-sourcing business models that major tech firms have deployed successfully to-date.

Major players within the transport and logistics sectors are using data analytics to optimise their operations, but often within their vertically integrated service offerings (e.g. DHL, fleet operators, UTMC providers, INRIX), rather than across transport networks. A lot of the analytical insight work being done currently is at academic research level, rather than as a service to transport operators and local authorities. This suggests potential buyers for data-led learning, analytics, and optimisation services are yet to recognise the scope of these services; or are buying them through conventional consultancy service arrangements without even realising it.

Demand-driven services like Lyft/Uber don't actually require much transport data currently, but if they want to integrate with other existing forms of transport (e.g. rail) or embed into lifestyle services (e.g. retail/dining out) then it will become important to develop mechanisms for exchanging their data with third parties in real-time. There is a significant opportunity around seamlessly integrated cashless, payment systems that organisations like Droplet are pursuing. These will significantly change the way people travel and offer scope to build very powerful consumer-behaviour datasets that exploit demand for the movement of people and things.

How is transport-related data created and shared?

There are five primary mechanisms for data creation (manual collection, overt crowd-sourcing, covert crowd-sourcing, sensor-derived, service provider generated). Web-connected fixed and mobile sensors,

plus crowd-sourcing are the emerging transport data collection mechanisms. In particular, personal/vehicle location data are the likely ‘game-changers’ for intelligent mobility.

The availability of personal/vehicle location-based datasets is being driven by the private sector, which plays a growing role in collecting and aggregating transport-related data. The increasing number of partial datasets creates a ‘signal problem’ – because we know less about what data being collected actually represent (e.g. in terms of sampled population, demographics). Data validation and curation therefore become very important, but few public sector data owners have resources to do this well. Key players such as Transport for London are leading the way, while many local authorities are still developing digital strategies. Some are replicating work being done by their neighbours, and might benefit from advice and guidance on maximising their investment in open data.

Transport data discovery and its uneven availability is becoming a major challenge:

- The multitude of catalogues and platforms needs to be embraced and documented.
- The best transport data feeds are concentrated at strategic national/major city levels.
- Almost no private sector data (even that you can buy) is documented in any form of catalogue.

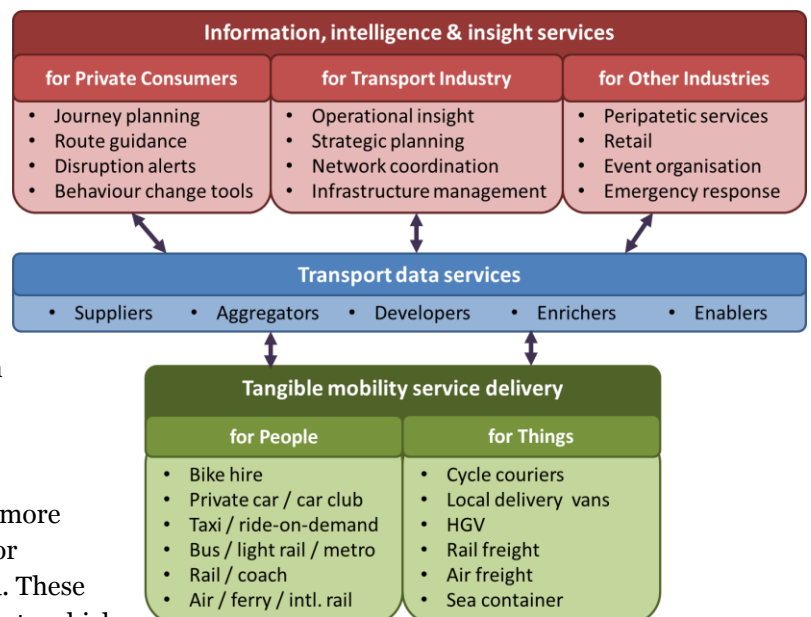
Very little real-time data feeds are openly available, and that is what most app developers and intelligent mobility service providers are seeking. There is even less archived data that have been collected in real-time. This will be critical for predictive analytics based on past trends/events and scenarios (2014-2021, and will need to become available if intelligent mobility is to evolve through to real-time service automation and optimisation (2021+).

Developers and service providers are increasingly interested in globally common data formats (e.g. GTFS, DATEX II in Europe), which make it easier to scale products and services internationally, yet few private sector providers adhere to them. Significantly, some forms of real-time data standard are yet to be defined (e.g. personal location-based data feeds).

How is transport data being exploited, and by who?

There are five types of transport data service: suppliers, aggregators, developers, enrichers, enablers). These will increasingly interact with each other, with intelligence/information & insight provision, and with tangible mobility service provision.

The aggregation and sale of raw open data is becoming a viable business, but the market for this data is still immature in relation to public transport data, with innovators essentially betting on future markets for these services. The automobile industry is very different, with more established end-to-end data value chains for traffic/disruption/roadworks/parking data. These are predicated on continued vehicle sales, into which



these data services are bundled, and currently receiving investment through connected and autonomous vehicle innovation.

Personal information services are increasingly reliant on cross-subsidised business models driven by major technology companies. This is making the marketplace challenging for smaller providers, because users have a high 'expectation of free'. Partly as a result, half (50%) of iOS developers and even more (64%) Android developers, are operating below an "app poverty line" of \$500 per app per month. Public attitudes to personal data privacy and concentrations of market power, and legislation to curb any abuses, present long-term threats to crowd-sourcing business models that major tech firms have deployed successfully to-date.

Major players within the transport and logistics sectors are using data analytics to optimise their operations, but often within their vertically integrated service offerings (e.g. DHL, fleet operators, UTMC providers, INRIX), rather than across transport networks. A lot of the analytical insight work being done currently is at academic research level, rather than as a service to transport operators and local authorities. This suggests potential buyers for data-led learning, analytics, and optimisation services are yet to recognise the scope of these services; or are buying them through conventional consultancy service arrangements without even realising it.

Demand-driven services like Lyft/Uber don't actually require much transport data currently, but if they want to integrate with other existing forms of transport (e.g. rail) or embed into lifestyle services (e.g. retail/dining out) then it will become important to develop mechanisms for exchanging their data with third parties in real-time. There is a significant opportunity around seamlessly integrated cashless, payment systems that organisations like Droplet are pursuing. These will significantly change the way people travel and offer scope to build very powerful consumer-behaviour datasets that exploit demand for the movement of people and things.

Where is the future value in transport data exploitation?

A total of ten different business models for exploiting transport-related data emerged through our dialogue with transport operators, practitioners, and data specialists. For UK businesses, the most exploitable future value for intelligent mobility services in transport data terms most likely resides in components of value chains that major global technology companies are currently overlooking or failing to exploit. These are anticipated to be:

- Opening-up and releasing unexploited datasets currently held by the public sector and its contracted service providers as open data.
- Autonomous sensor-based monitoring and data capture for transport networks and services via the growing Internet of Things.
- Aggregating and collating multiple transport data feeds so they can be readily combined and analysed for patterns, deployed in models, and up-scaled from representative population samples to inform strategic transport planning and analyses.
- Statistical and computational analyses/modelling of transport datasets, with consumer and socio-economic datasets, to create market intelligence and actionable insights.
- Predictive and real-time analytics, and automated control systems, that optimise capacity in relation to demand across all transport networks.

- Products and information services that are compelling enough to be ‘baked-in’ to the next generation of intelligent mobility services (e.g. connected/autonomous vehicles).
- Developing ‘intermediate technology’ transport applications, tools and open data standards relevant to emerging economies’ transport systems.
- Establishing ethically-focused tools that fairly and transparently trade user’s digital exhaust data in return for contextually relevant information, insight, and/or incentives.

We identified 20 possible transport data exploitation barriers framed around the themes of data availability, data usability, data relevance, structural impediments, and market forces. Most of these challenges are considered solvable; and are likely to be dealt with organically as the Internet of Things develops, and service provider’s demands for existing public datasets grows. Those which relate to market forces are less easy to address, and may instead represent unavoidable ‘features’ of transport data markets around which intelligent mobility services will need to work.

What data related capabilities are needed?

We defined five key data capability groups for intelligent mobility, focused on:

- Raw data creation, collection and curation
- Dataset handling and manipulation
- Computational and statistical analyses
- Human intelligence and use of data insights
- Software and technology development.

These led us to three key requirements for building capability and capacity for intelligent mobility:

- Skilled technical talent capable of handling and analysing very large datasets compiled from multiple sources.
- Organisational capability that ensures business leaders’ understand new analytical processes and business models in outline and can use actionable insights for strategic decision-making.
- Technological investment to ensure access to requisite data storage capacity and computational processing power, for example through the use of cloud-hosted servers.

Based on the CEBR’s forecasts on the number of jobs that will be created by Open Data, we estimate that by 2017, as many as 3,000 data specialists will be needed to support the UK transport industry’s drive to achieve data-driven efficiencies and optimisation. Additional skilled analytical jobs are also likely to be created as new intelligent mobility services emerge.

What do roadmaps for intelligent mobility look like?

We defined five roadmaps that connect with the 20 intelligent mobility services and 19 core intelligent mobility datasets identified through this review. These are available in Chapter 8 of this report, and we recommend they are maintained and updated.

What can the UK do to catalyse transport data exploitation?

We identified 20 possible actions that could be undertaken by various organisations to accelerate the adoption of intelligent mobility services, deepen integration between existing transport networks, and establish data-driven analytics that support more intelligent and optimised mobility services for people and things:

Quick Wins	
1	Establish an advice service, with accompanying data collection and publishing tools, to help UK public sector open-up, maintain, and exploit the core transport datasets identified by this review. Promote linked open data in line with ODI good practice.
2	Incubate, publicise, and maintain an online transport data ‘catalogue of catalogues’ that becomes the ‘Argos’ or ‘Amazon’ for UK transport data. Requirements include: * Focus on core datasets for intelligent mobility identified in Table 3-3 of this report. * Include both open (free) and available (licensed/paid-for) datasets. * Showcase transport datasets so they are easy to discover and learn about. * Achieve national coverage for each dataset by grouping locally maintained feeds. * Use the IoT Hypercat standard, so data are easy to find and exploit by both humans and machines; and readily related to datasets ‘siloed’ in other data hubs.
3	Raise awareness of this catalogue, and the benefits associated with data-driven mobility services, analytics, and associated data-services; through regular presentations, workshops and best-practice/learning events aimed at UK city/local authorities and transport operators/coordinating agencies.
4	Compile regular (quarterly/annual) market monitoring and intelligence reports to build an evidence base valuing and tracking ‘user needs, drivers and opportunities’ for exploiting transport data, as identified in paragraph 6.3 of this report. Publish and publicise this intelligence to send commercial signals to UK transport industry and data marketplace.
5	Work with UK transport operators, agencies, and local authorities to broker open releases of as many core transport datasets as possible, specifically focusing on data gaps identified in Table 3-2 of this report. Work with public sector transport agencies to access the Release of Data Fund, and Breakthrough Fund, and catalogue new open data sets in-line with recommendation #2.
6	Engage major private sector organisations that hold data (bus operators, retailers, payment system providers) to share their data within ‘sandpit’ environments for innovative feasibility projects. Outcomes will need to showcase the intelligence, efficiencies, optimisations and mobility services that data-driven intelligent mobility providers can deliver if private sector datasets are made available for exploitation. Such projects need to provide direct routes-to-market for innovations delivering tangible value in meeting private sector intelligence/mobility needs.
7	Engage with emerging service providers (e.g. Uber) and existing transport network operators to foster data-driven integration between ‘old’ and ‘new’ mobility services.

8	Collaborate with UKTI, international development agencies, and developing/transitional city authorities to identify overseas data-driven mobility user-needs, and opportunities for UK companies to exploit intelligent mobility expertise.
9	Devise a suite of Transport Data Science courses and modules with UK Universities. Their focus should be on increasing the number of graduates with the skills that meet the core capabilities and skills identified in Chapter 7 of this review.
Strategic longer-term activities	
10	Develop and maintain a publicly accessible cloud-hosted archive of core transport datasets (both ‘static’ and ‘real-time’), so their availability is guaranteed for historic data pattern analyses / predictive analytics in the future. This is considered important for migration to intelligent mobility based on predictive analytics (2014-2021) as a precursor to real-time service automation and optimisation (2021+).
11	Establish a centre of excellence for data-driven business model development to provide guidance and advice for innovative companies seeking to exploit transport-related data. This would draw on market intelligence reports and provide much needed user-needs focus.
12	Incubate, launch and publicise an online marketplace for commoditised and wholesale transport datasets catalogued in #2 above. The TSC can act as honest broker to connect buyers and sellers of transport data together online (including re-sale through existing platforms like transportAPI). Improved transparency and reduced cost of transport data-as-a-service are expected to accelerate the development and commercialisation of user-focused intelligent mobility services.
13	Work strategically with Innovate UK, Transport KTN, and relevant Catapults (Transport Systems, Satellite Applications, Digital, Future Cities) to co-ordinate innovation projects and harness the datasets they create so they are included in the catalogue described in #2.
14	Establish strategic placement opportunities with the UK transport industry to ensure these graduates are retained by UK intelligent mobility service providers.
15	Professional bodies such as the Transport Planning Society (TPS), Chartered Institution of Highways & Transportation (CIHT), and Chartered Institute of Logistics and Transport (CILT) to build organisational & management capabilities through transport data analytics training schemes for professionals in the sector.
16	Support the growth of IoT sensor networks through demonstration/collaborative technology projects with transport industry partners. These could be established through demonstration projects (e.g. Finnish Traffic Lab) to quantify benefits and impacts of integrating IoT sensor data from road, parking, public transport, walk/cycle sources, and to accelerate uptake of demand-responsive mobility networks and services.

17	Transport sector to participate in debate around data privacy issues, and ethical personal data/'digital exhaust data' re-use. Mobility organisations keen to share user's data, and providers looking to consume it, would benefit from clear guidance on legitimate uses and appropriate practices for anonymising and aggregating user-derived data. Consumer guidance on protecting personal (location) data and their rights
18	Transport sector to play an active role in supporting, defining, and maintaining increasingly global common data formats (DATEX II and GTFS) and defining new standards (e.g. for the exchange of personal location data) that will be needed to drive intelligent mobility service delivery over the next 10 years.
19	Engage in/influence government discussion around cashless payment and e-purse legislation. The current £20 threshold will become a barrier to integrated cashless payment systems for future intelligent mobility.
20	Engage with all UK major freight handlers and forwarders, e.g. through industry representative groups, to identify non-competitive/collaborative opportunities for data-driven sharing of logistics networks/operational data so as to maximise optimisation and efficiency across the sector.

Contents

	Page
1 Introduction	
Purpose of this study	4
Structure of this report	5
2 What mega trends are shaping transport and mobility?	
Increasing volumes and accelerating sample rates for data	6
Interconnected and Open (transport) Data	8
Data companies becoming mobility service providers (and vice versa)	10
Pervasive dominance and influence of multinational technology companies	11
The emergence of intelligent mobility	12
3 What datasets are relevant to transport?	
Key themes for transport related data	14
What immediate data gaps exist?	17
Key transport-related datasets for intelligent mobility	19
Challenges to filling intelligent mobility data gaps	21
4 How are transport-related datasets created and shared?	
Mechanisms for data collection	22
Who collects transport-related data?	25
How is transport data validated and curated?	26
How is transport data being shared?	26
Importance of real-time data for intelligent mobility	30
The role of standardised data formats	33
Challenges for evolving transport data creation and sharing	35
5 How are transport datasets being exploited, and by who?	
Defining transport data services in relation to mobility services	37
Transport data services and providers	39
Information, intelligence and insight services	44
Tangible mobility services delivery	58

	Page	
6	Where is the future in transport data exploitation?	
	Intelligent mobility motivations for exploiting transport data	65
	Business models for exploiting transport data	67
	Identified challenges and obstacles to exploiting data	69
7	What data-related capabilities are needed?	
	Key data-related capabilities for intelligent mobility	71
	Requirements for building capability and capacity for intelligent mobility	74
8	What do roadmaps for intelligent mobility look like?	
	Resilience	81
	Autonomous Systems	82
	End to End journeys	84
	Information Exploitation	85
9	What can the UK do to catalyse transport data exploitation?	
	Independent activities	86
	Activities requiring co-ordinated action	87

List of Figures

Figure 4-1: Mechanisms for transport data collection	23
Figure 4-2: Indicative pathway for real-time data collection and sharing	32
Figure 5-1: Typology of mobility service types influenced by transport data	39
Figure 5-2: Data journey through transportAPI's platform	43
Figure 5-3: Data journey through Parkopedia's platform	48
Figure 5-4: Data journey through Elgin's platform	53
Figure 5-5: Data journey through TrafficMaster's platform	57
Figure 5-6: Data journey through Uber's platform	59
Figure 5-7: Data journey through the FreightArranger platform	64
Figure 7-1: The (un)sustainability of global app-based businesses	77

Page**List of Tables**

Table 2-1: Intelligent mobility services predicted to emerge in the next 10 years	13
Table 3-1: Intelligent Mobility data themes and example datasets	15
Table 3-2: Transport-related data gaps	17
Table 3-3: Key transport related datasets for intelligent mobility	19
Table 4-1: Emerging transport data collection mechanisms	24
Table 4-2: Commonly-used transport-related data portals	27
Table 4-3: Commonly-used national transport data standards and formats	34
Table 5-1: Transport data service types	38
Table 5-2: Examples of transport data services and providers	41
Table 5-3: Example information, intelligence and insight services for private consumers	45
Table 5-4: Example information, intelligence and insight services for the transport industry	50
Table 5-5: Example information, intelligence and insight services for other industries	55
Table 5-6: Examples of tangible mobility services for people	60
Table 5-7: Tangible mobility services for things	62
Table 6-1: Motivations for exploiting transport data	65
Table 6-2: Estimates of GVA per capita from Open Data	66
Table 6-3: Business models for exploiting transport data	68
Table 6-4: Possible transport data exploitation barriers for intelligent mobility	70
Table 7-1: Big data techniques and technologies relevant to intelligent mobility	72
Table 7-2: Data-related capabilities and example activities for intelligent mobility	73
Table 8-1: Commentary on intelligent mobility roadmaps (by TSC theme)	79

Appendices

- Appendix A: Intelligent mobility services identified through this review
- Appendix B: Catalogue of transport-related datasets identified through this review (separate files)
- Append C: Database of service providers identified through this review (separate files)

1 Introduction

1.1 The [Transport Systems Catapult's](#) (TSC) five year delivery plan estimates the global [intelligent mobility](#) market will be worth £900bn by 2025². Opportunities to capture a UK share of this market closely align to the organisation's vision for more responsive and predictive transport systems that:

- Better meet the needs of an inclusive society by efficiently and sustainably connecting goods, services, events and people.
- Optimise the use of available infrastructure capacity to maximise the time, energy and resource efficiency of travel and transportation; thereby supporting continued economic growth and expansion.
- Are more readily connectable and flexible - promoting seamless intermodal journeys that can flex according to disruptions, changes in schedule or priority, and competing demands for other seemingly unrelated services.
- Generate smaller environmental and social impacts than current transport systems

1.2 The Catapult anticipates scope for intelligent mobility advances to emerge around 5 key themes:

- Autonomous systems
- Smart infrastructure
- Information exploitation and customer experience
- End-to-end journeys
- Resilience

1.3 The sourcing, aggregation, fusion, and manipulation of complex and multi-source data are therefore considered central to achieving more intelligent mobility. Included in this grouping will be the ability to manage not only extremely large datasets; but also extract actionable insights, model, compare and predict behaviour and meaning from sparse and incomplete datasets. These may be collected from different sources, along different timescales, and will therefore require flexible techniques that take account of this.

All of the above could conceivably be of use in the future delivery of intelligent mobility services to facilitate the development of new applications, insights and capabilities. They are also relevant in the context of opening-up and exploiting the wealth of existing datasets held by existing transport service providers and coordinating agencies.

Purpose of this study

1.4 The range of transport-related datasets and processes, and speed of change in this sector, makes it challenging to establish a coherent overview of all that is happening in the marketplace. Without this the organisation's efforts to define a strategic plan for influencing the market for intelligent mobility, support capability-building in the UK, and achieve its intelligent mobility vision could be hampered.

² Transport Systems Catapult (2013) *Five-Year Delivery Plan to March 2018*. Available [online here](#), last accessed on 28/04/14.

- 1.5 In March 2014 the Catapult commissioned [Integrated Transport Planning Ltd \(ITP\)](#) to lead a structured investigation of this topic area to provide this coherent overview. We associated with experts from [White Willow Consulting](#), [Advancing Sustainability](#) and the [Horizon Digital Economy Research Institute](#) to rapidly gather and report evidence in a practically focused way.
- 1.6 Our review was consciously about ‘more than just data’ since techniques used to capture, share, aggregate, and process relevant datasets are also important. While the primary focus has been on the UK; we brought insights from international open transport data initiatives, research, project stakeholders, and collaborators to the study. We combined desk-based research, stakeholder interviews and technical workshops to draw upon insights and experiences from industry experts and practitioners in order to:
- Cover major studies and projects both completed and currently under-way.
 - Break-down and connect deployed capabilities, functional areas and technologies.
 - Focus specifically on the creation, management and exploitation of complex datasets.
 - Identify the tools and applications that are deriving tangible value; and consider the drivers, key players, investment, trends and gaps that emerge through the research.
 - Relate all of the above back to the 5 themes in a future roadmap for intelligent mobility data and through our study recommendations.
- 1.7 By reviewing and collaboratively defining the current landscape and future horizon for transport data related studies, activities and initiatives; we have effectively been able to work backwards from a 2024 scenario in which an intelligent mobility vision is being realised. The roadmaps described in this report chart the opportunities, gaps, phasing, dependencies, and drivers in respect of intelligent mobility data journeys from 2014 - 2024.

Structure of this report

- 1.8 This report represents the main deliverable from our study, and is structured as follows:
- Chapter 2 documents the key trends that we anticipate will shape the emergence of data-driven intelligent mobility services.
 - Chapter 3 catalogues the transport-related datasets we identified as being relevant in 2014, and are likely to remain pertinent to intelligent mobility in 2024.
 - Chapter 4 documents how these datasets are created and shared, and relates these approaches to key market trends identified in Chapter 2.
 - Chapter 5 explores current and future mobility services that are exploiting transport-related datasets, and the organisations behind these innovations.
 - Chapter 6 identifies the exploitable value we have identified in relation to transport-related data, the different business models established and emerging in the sector, and the obstacles to exploiting transport-related data.
 - Chapter 7 identifies the capabilities needed to exploit transport-related data to 2024.
 - Chapter 8 presents a series of themed intelligent mobility data roadmaps to 2024.
 - Chapter 9 summarises opportunities and ideas for how the UK’s exploitation of transport-related data might be accelerated.

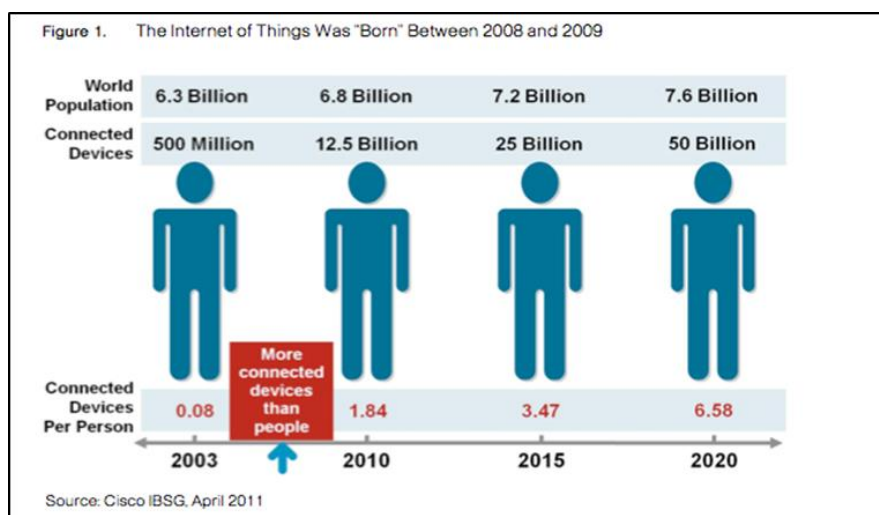
2 What mega trends are shaping transport and mobility?

2.1 This study identified a number of ‘mega trends’ that are expected to shape the way transport related data are collected, processed, and used over the course of the next 10 years. In turn, they are expected to change the way transport services are delivered, with a greater emphasis on enhanced mobility for people and things. These trends have been summarised and reported up-front since their influence is expected to be widespread, touching all of the discussion points in following chapters of this report.

Increasing volumes and accelerating sample rates for data

2.2 Increasing rates of transport data collection will continue to be facilitated by anticipated growth in the speed and coverage of wireless communications technologies, computer processing power, and data handling capacity. The key driving forces are:

- The proliferation of devices and ICT applications connected to the internet (such as smart meters, smart grids, and smart transport services) based on sensor networks and machine-to-machine communication³. The number of networked sensors and information generators is growing at over 30% per annum, creating a rapidly expanding ‘Internet of Things’ (IoT)⁴ that is projected to contain 50 billion devices in 2020.



³ OECD (2012) Exploring Data-driven Innovation as a New Source of Growth – Mapping the Policy Issues Raised by ‘Big Data’. Available online at:

[http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/ICCP\(2012\)9/FINAL&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/ICCP(2012)9/FINAL&docLanguage=En), last accessed 02/07/14.

⁴ Estimates from [Gartner Group](#), [ABI Research](#) and [Cisco IBSG](#) suggest between 20, 30 and 50 billion devices will be connected to the Internet of Things by 2020.

- Globally ubiquitous broadband access and low cost wireless connectivity (WiFi, 4G and 5G mobile comms, and White Space communications⁵) will enable the transmission and association of increasingly larger quantities of digitised data, at faster speeds. The faster these technologies roll-out at scale to fill key ‘notspots’ (road/rail networks), the greater the pace and volume of transport data sampling is likely to accelerate.
- The technological changes described above mean that people’s interactions with the growing Internet of Things surrounding them can be automatically tracked and stored more comprehensively and accurately than at any point in time previously. Individual’s digital exhaust⁶ data, generated by their interactions and transactions with both private and public sector organisations, are already rapidly increasing in volume and have value to marketers and service providers. [Reputation](#) has calculated that free (at point of use) access to Google’s web-based services is effectively traded with each user for up to \$5,000 worth of their personal data per-annum, in terms of its value to Google⁷.
- The combined effect of the three trends described above is that the size of transport-related datasets – in line with those held in other industry sectors – is set to grow significantly over the next 10 years. The volume of all digital data being created and stored on servers and in the cloud is doubling almost every 1.2 years, and the volume of newly created data was estimated to have exceeded 1,000 exabytes (1 exabyte = 1,000 petabytes, 1 petabyte = 1,000 terabytes) in 2013 alone⁸.
- These ‘Big’ datasets are increasingly beyond the capacity of conventional databases; requiring new techniques, tools and computing systems to store, manage and utilise them. In transport-related data terms, a key trend will be the adaptation of existing non data-driven processes, and currently closed operational datasets; to capture, curate, store, search, share, transfer, analyse and visualise datasets so they are converted into actionable insight and usable information. Such data-driven insights will also require new human intelligence and understanding to ensure the wisdom is applied in an informed manner. The capabilities for exploiting ‘big transport data’ will therefore be needed at both managerial/leadership levels as well as analytical and statistical roles.

⁵ Centre for White Space Communications (2014) *What is White Space*. Available online at: <http://www.wirelesswhitespace.org/about-us/what-is-white-space.aspx>, last accessed 02/07/14.

⁶ Dale Neef (Forthcoming) *Digital Exhaust: What Everyone Should Know About Big Data, Digitization and Digitally Driven Innovation*.

⁷ MarketWatch (2012) *Who Would Pay \$5,000 to Use Google? (You)*. Available online at: <http://blogs.marketwatch.com/realtimeadvice/2012/01/25/who-would-pay-5000-to-use-google-you/>, last accessed on 02/07/14.

⁸ MIT Center for Digital Business (2012), quoted in Reagan JR (2012) *The Trail of Digital Exhaust*. Available at: http://www.deloitte.com/view/en_US/us/Services/additional-services/deloitte-analytics-service/short-takes/d1c68e6b7b407310VgnVCM1000001956fo0aRCRD.htm, last accessed on 03/07/14.

Big Data – when does a dataset become ‘Big’?

The trend to larger data sets is due to the additional information derivable from analysis of a single large and complex set of related data, as compared to separate smaller sets with the same total amount of data. These larger datasets allow new (or previously inaccessible) correlations and insights to be identified⁹.

- 2.3 The ‘perfect storm’ created by the combination of technological advances described above has led commentators to describe data as “the new oil” that can power the economies and transport systems of the future. It is clear that the wealth of data set to become available over the next 10 years provides unprecedented scope to (re)organise, rationalise, and affect the way the world works.

Interconnected and Open (transport) Data

- 2.4 The Open Data movement has its conceptual roots in the international scientific community’s World Data Centres which have provided open access to scientific data since the 1950’s¹⁰. More recently, and particularly in the last 5 years, a growing amount of data has been made openly available by governments around the world.

Open data – how does it differ from other data?

As defined by the Open Data Institute (ODI), Open Data are available for anyone to use, for any purpose, at no cost. The degree of openness of a dataset can be rated based on the 5 star scale proposed by Tim Berners-Lee, the inventor of the Web and Linked Data initiator. The characteristics that differentiate Open Data from other available data are:

Open data are typically ‘cost free’ to consume and re-use.

Open data are readily discoverable, and supported by meta-data that make it easier for machines (e.g. search engines) and humans to find.

Open datasets are free from licensing restrictions, and typically the only requirement for re-use is attribution and commitment to onward sharing.

Open datasets are free from licensing restrictions, and typically the only requirement for re-use is attribution and commitment to onward sharing.

⁹ The Economist (2010) *Data, Data, Everywhere*. Available at: <http://www.economist.com/node/15557443>, last accessed on 03/07/14.

¹⁰ World Data Center System (2009). *About the World Data System*. Available online at: <http://www.icsu-wds.org/organization>, last accessed on 03/07/14.

- 2.5 In the UK, open transport data have mainly come on-stream since 2009¹¹, when the government made public open data the focus of its transparency commitments¹². Consequently the National Transport Access Nodes (NAPTAN) dataset has been joined by the datasets that power National Rail Enquiries' journey planner, real-time rail departures and arrivals, car park location data, bus & light rail schedules for the whole of the UK, and real-time traffic speeds and disruptions on the UK's strategic road network¹³ (but not all roads).
- 2.6 Although in theory all these datasets should be made available as Linked Open Data, so that their relationships to data from other sources can be defined and referenced using standard web technologies (e.g. HTTP, RDF and URIs), this is not currently happening. Instead they are freely available as 3 star open data (structured data in an open format, such as CSV or XML). They are not served using URIs (Uniform Resource Identifiers), which means that although they are available via the internet they are not technically connected 'in the web', and therefore cannot be readily linked to other similar datasets. This reflects a trend in which the public sector has, quite logically, been working hard to open-up data in its current state, before working to improve the quality and usability of datasets that are well-used.
- 2.7 Open transport data, and in particular public transport data, have been a success story for the UK government in respect of the number of applications and tools created based on these datasets, and subsequently purchased or downloaded by users of UK transport systems. The expectation within the industry is that, over the next 10 years:
- The quality of linked open data will improve, with increasing numbers of public datasets attaining 5 star status.
 - A growing number of public open datasets will be provided as real-time feeds using APIs and/or globally-common standards.
 - The breadth of available public sector data will continue to expand so that datasets not currently available emerge to fill obvious current data gaps (bus fares, urban road traffic conditions, real-time train vehicle locations).
- 2.8 It remains to be seen how freely and openly the data created and stored by the private sector (e.g. technology companies, retailers, transport service providers) will be shared over the next 10 years. Based on current practices it is unlikely to be fully open, but this has not proved to be a barrier to major international data companies such as TomTom and INRIX in the past. They have a history of collecting, curating, and selling traffic data for OEM and consumer Sat-Nav products; having done so since before Open Data were available.

¹¹ Deloitte (2012) *Open Growth – Stimulating demand for open data in the UK*. Available online at: http://www.deloitte.com/view/en_GB/uk/market-insights/deloitte-analytics/bfb570a79416b310VgnVCM1000003256f70aRCRD.htm, last accessed on 17/07/14.

¹² Cabinet Office (2014) *Transparency and Open Data: progress against commitments*. Available online at: <https://www.gov.uk/government/publications/transparency-and-open-data-progress-against-commitments>, last accessed on 31/07/14.

¹³ Nick Illsley (2012) *Department for Transport - Our Open Data/Transparency Story*. Available at: <http://data.gov.uk/blog/department-for-transport-our-open-data-transparency-story>, last accessed on 03/07/14.

Data companies becoming mobility service providers (and vice versa)

2.9 A recent trend (within the last 5 years) is for companies that have traditionally dealt in data and converted it into information that is useful to people, or used it to provide virtual services; to get directly involved in the provision of mobility services for users – particularly those involving personal automobility. Examples include:

- [Parkopedia/Park@MyHouse](#) setting aside £15m to rent physical parking spaces from corporate and Council providers to fill gaps in the marketplace they have identified using their global parking space data. Investment from BMW, which is building its in-vehicle parking service (ParkNow) using data provided as a service by Parkopedia¹⁴ and payment systems provided by Now! Innovations¹⁵, has facilitated this.
- [TomTom](#) has evolved from being a company that only sells Sat Nav hardware products directly to consumers to one that also uses real-time analytics to create and sell data products that power OEM products and services for vehicle and device manufacturers¹⁶.
- [Google](#), originally a data and information company, is investing in numerous technologies in the transport sector (e.g. Uber¹⁷, driverless cars¹⁸, cashless payment systems¹⁹) that indicate it is primed to become a mobility service provider drawing on its vertically-integrated stack of maps, places, and search-driven data.
- Strava's personal running and cycling GPS tracking service has created a 300 billion GPD point database, which is updated by 2.5 million GPS-tracked activities each week. Since May 2014²⁰ users' anonymised data has been aggregated and sold through [the Strava Metro](#) data service.

¹⁴ Roadtraffic-technology.com (2013) *Parkopedia to offer parking information services to BMW*. Available online at: <http://www.roadtraffic-technology.com/news/newsparkopedia-offer-parking-information-services-bmw>, last accessed on 08/07/14.

¹⁵ Smartplanet.com (2014) *BMW expands parking service*. Available online at: <http://www.smartplanet.com/blog/bulletin/bmw-expands-parking-service/>, last accessed on 08/07/14.

¹⁶ Guardian (2011) *Harold Goddijn: TomTom's founder needs his business to turn the corner*. Available online at: <http://www.theguardian.com/business/2011/nov/24/harold-goddijn-tomtom>, last accessed on 08/07/14.

¹⁷ Techcrunch (2013) *Google ventures puts \$258m into Uber, its largest deal ever*. Available online at: <http://techcrunch.com/2013/08/22/google-ventures-puts-258m-into-uber-its-largest-deal-ever/>, last accessed on 08/07/14.

¹⁸ BBC (2014) *Google is to start building its own self-driving cars*. Available online at: <http://www.bbc.co.uk/news/technology-27587558>, last accessed on 08/07/14.

¹⁹ New York Times (2014) *Transit Cards to Replace Cash on Kenyan Minibuses Are a Hard Sell*. Available online at: http://www.nytimes.com/2014/07/16/world/africa/transit-cards-to-replace-cash-on-kenyan-minibuses-a-hard-sell.html?_r=0, last accessed 23/07/14.

²⁰ Outside (2014) *Strava's plan to Revolutionise Commuting*. Available online at: <http://www.outsideonline.com/outdoor-gear/gear-shed/cycle-life/Stravas-Plan-to-Revolutionize-Commuting.html>, last accessed on 08/07/14.

Beyond the release of open data feeds covering public transport services that are operated on behalf of the public sector (e.g. TfL's open data feeds²¹), there are almost no examples of private sector operators of bus and train services seeking to exploit their operational and patronage data in the same way. We anticipate this may change over the next 5-10 years as the operators of existing transport services (bus, rail, air, freight) respond to the emergence of new data-driven mobility companies and seek to commercially exploit their operational and passenger-related data. As noted earlier, while we anticipate these datasets will become more available, it currently appears unlikely they will be released as Open Data.

Pervasive dominance and influence of multinational technology companies

2.10 A final 'hard to ignore' trend is the pervasive dominance of major technology companies and their interest in the field of personal mobility and transport-related data. Organisations such as Google, Microsoft, Ebay, Apple, Nokia, Yahoo, Samsung, Facebook, Twitter, and Amazon increasingly control, or are driving change, in relation to:

- Effective systems (both hardware and software) for crowd-sourcing and aggregating very large sets of personal movement and location data that are fundamental for intelligent real-time transport system planning and operations.
- The development of very large map-based and place-related platforms that are being maintained through a combination of crowd-sourced sensing, 'ground-truth' auditing, remote sensing, and open data consumption.
- Defining international data standards for emerging communications technologies and de-facto common global formats (e.g. the [General Transit Feed Specification](#)) for storing transport-related data so it can be converted into user-facing information.
- New forms of mobility for people (e.g. Google's driverless car programme) and things (e.g. Amazon's drones for direct delivery of goods to people).

This trend poses questions about the likely role of the public sector in respect of the data-driven underpinnings anticipated to power more intelligent mobility services. Because of their scale, dominant technology companies are likely to be enablers of intelligent mobility services – moving transport markets into new areas and integrating their existing products and services. However, the vertically integrated and closed nature of the data services that power these major tech companies also significantly threaten UK innovation companies, and existing mobility service providers that are not already operating at scale.

²¹ Transport for London (2014) *Open Data Users*. Available online at: <https://www.tfl.gov.uk/info-for/open-data-users/>, last accessed on 08/07/14.

The emergence of intelligent mobility

- 2.11** The final trend we explored was the anticipated emergence of disruptive intelligent mobility services. These are expected to both exploit, and create, new datasets focused on individual and organisational mobility needs; as well as their combined impact on the movement of ‘things’ (raw materials, manufactured goods, personal services).
- 2.12** Our first workshop explored what transport and data specialists understood to represent intelligent mobility services, and the timescales in which they expect these services to emerge. This activity was backed-up by desk-based meta-analyses to validate this consensus (to an [early adoption stage](#) in an average-sized UK city, such as Coventry) against other published industry ‘expert’ forecasts on the emergence of new services.
- 2.13** The intelligent mobility services identified by this process are summarised in Table 2-1 overleaf, and further detailed in Appendix A to this report. Roadmaps that group the services based on which of the 5 intelligent mobility themes they relate to, and their enabling transport-related datasets, have been presented in Chapter 9 of this report.
- 2.14** In preparing this table, and those in Appendix A, we note that predicting future trends in both mobility service provision and technological development is inherently uncertain. While many of the emerging innovations identified are re-imagined versions of existing forms of transportation (or disconnected ancillary services), some represent entirely new forms of mobility that take advantage of technological advances and the accompanying data flows they create. We deliberately made no value judgement in favour of either group, but note it is typically easier for people to think of innovation in terms of evolution from existing forms of mobility.
- 2.15** As such, the list of intelligent mobility services presented in Table 2-1, and in Appendix A, should be treated as both formative and indicative of what may emerge to 2024. The precise timing or ordering of progress is unlikely to be accurate, but represents a set of informed assumptions based on published research, pilot projects and industry opinion. Finally, given the data-focused topic of this study, the list of intelligent mobility services developed through this review is assumed to represent a primarily data-driven subset of all the innovative transport technologies that are currently being developed.

Table 2-1: Intelligent mobility services predicted to emerge in the next 10 years

Timing	Intelligent Mobility Services identified
Near-term 2014 to 2017	Route mining from ‘big’ origin and destination datasets for improved transport modelling and personal decision-making.
	National real-time parking space information service integrated with route guidance and payment mechanisms.
	Crowd-sourced transport on-demand (e.g. Uber/Lyft).
	Travel Smart Meter to enable objective travel impact monitoring.
	Personalised, context-specific, multi-modal journey planning with real-time guidance and updates.
	Smarter logistics to directly connect customer & goods, reducing need to travel.
Med-term 2018 to 2021	Transport provided as a consumer service, rather than as a dedicated journey service in its own right.
	Personalised mapping services that cover indoor and outdoor places.
	Using historic public transport usage data to inform service planning and smooth demand across transport networks.
	Bus services that are truly demand responsive along key corridors / zones.
	Driverless vehicles using dedicated routes or tracks in urban areas
Long-term 2022 to 2024	National real-time capacity indicators for all modes of public transport and roads with historic data available for predictive analytics.
	New airborne freight transportation e.g. airships / drones
	Autonomous vehicles running on-street alongside non-autonomous vehicles.
	Contactless/mobile ticketing for all UK transport services & operations.
	Demand responsive pricing across all transport services and networks (including roads)
	Consolidation of logistics operations across modes, providers and networks

3 What datasets are relevant to transport?

3.1 Through this study we reviewed available UK transport-related datasets relevant to intelligent mobility. This meta-analysis drew on the work of a number of major studies and demonstration projects including the data audit for the Glasgow Future Cities demonstrator project²² and datasets contained in the [Stride](#) and [i-Move](#) Internet of Things demonstrator projects funded by the Technology Strategy Board. We also explored other catalogues for transport data including [data.gov.uk](#), [TfL's open data portal](#), the [Open Knowledge Foundation's datahub](#) various city authority open data portals, and map-based information portals.

Key themes for transport related data

3.2 The catalogue of datasets established through our meta-analyses is included as **Appendix B** to this report. Each dataset we identified has been grouped into one of 9 key themes. These themes were defined by our research team based on learning from our meta-analyses, with further refinement achieved through workshop discussions with transport data experts.

3.3 Table 3-1, overleaf, defines these 9 themes for transport-related data and includes examples of the types of data that can currently be found in relation to each one. Where available, the appended data catalogue contains further information and meta-data about each of the datasets we found including:

- The URL where the dataset can be found.
- Details of the dataset owners and maintainers.
- A brief description of the dataset's contents.
- Which of the 9 intelligent mobility data themes the dataset is relevant to.
- The way the dataset is assembled, formatted and released.
- The coverage and update frequency of the dataset.
- Whether it is part of the UK's National Information Infrastructure²³.
- Classification of the degree of fluidity in the dataset (static, dynamic, or real-time).
- Classification of the degree of openness and availability of the dataset.
- Applications and services that are using the dataset.

²² Glasgow City Council & University of Strathclyde Glasgow (2013) *TSB Future Cities Demonstrator Feasibility Study Data Audit*. Summarised in Feasibility Study report available online at: <https://connect.innovateuk.org/documents/3130726/3794125/Feasibility+Study+-+Glasgow+City+Council.pdf>, last accessed on 09/07/14

²³ Cabinet Office (2013) *National Information Infrastructure*. Available online at: <https://www.gov.uk/government/publications/national-information-infrastructure/national-information-infrastructure-narrative>, last accessed on 09/07/14.

Table 3-1: Intelligent mobility data themes and example datasets

Theme	Definition	Example Datasets	
Place & Space	Anything tangible that can be seen, touched or found.	TfL’s digital speed limit dataset Rail station / bus stop locations	OSM/OS/Google/Apple/Nokia/Bing maps Locations of on and off-street car parking
Environment	Data relating to environmental trends and natural occurrences.	Real-time weather sunrise/sunset times Rainfall, tides and marine conditions Floods	Earthquake monitoring Air quality Droughts
People, Things & Movement	Data generated by individuals and things as they move around.	Location history from cell/smartphones Employment / health / education data	Stats19 road incident data Bike journey counters /O-D data
Disruption and event-related data	Dynamic datasets related to physical events that impact transport networks.	Sporting fixtures Live and planned road closures	Real-time traffic incident reports Real-time road gritting in winter Roadwork locations
Public Transport Services	Scheduled and real-time data relating to the movement of public transport vehicles, and their characteristics.	Traveline National Dataset – public transport routes and schedules Traveline NextBus Real-time data	ATOC/National Rail Enquiries live departure boards TfL Journey planner
Personal Automobility	The spatial movement of powered personal vehicles (e.g. cars, motorcycles, taxis).	Highways Agency live traffic information DATEX II XML Waze/Google traffic speed data	TfL live traffic camera images Glasgow on-street parking data Vancouver EV charging stations
Freight connections	Data related to the movement of goods by road, rail, sea, air.	Shipfinder Road and port freight statistics	Track-Trace API Freight train movements

Theme	Definition	Example Datasets	
International Connections	International travel outside of the UK by Air, Rail, Sea.	Flight Radar 24 Eurostar scheduled departures	Port schedules
Consumption & transaction data	Individual preferences and retail choices – both directly and indirectly related to transport.	Oyster card derived travel data Credit card spend-data Domestic energy consumption	Petrol prices & Rail fares Loyalty card purchases Restaurant bookings

3.4 The nine themes identified through this review have the potential to be used as the basis for distinguishing between:

- Core transport datasets (bus stops, service schedules and real-time locations).
- Foundational data (place and space, and environment data).
- Valuable non-transport data (consumption/transaction data, and disruption/event data).

Such distinctions could assist public and private sector organisations looking to open-up and share data they already have, whilst also helping service providers, researchers, and developers to discover datasets that are critically important to intelligent mobility.

What immediate data gaps exist?

3.5 The process of establishing the data catalogue described above, and contained in **Appendix B**, highlighted some of the gaps in UK transport-related datasets. These datasets, and the services they potentially enable, are summarised in Table 3-2, below:

Table 3-2: Transport-related data gaps

Theme	Dataset	Services enabled
People, Things & Movement	Automated cycle count data from National Cycle Route network	Infrastructure usage, further investment business cases, maintenance prioritisation
Disruption and event-related data	Database of all UK major events compiled from local authority event licensing and sporting schedules	Transport capacity service planning, smarter traveller information services
Personal Automobility	Traffic speed and count data from urban traffic management control systems and non-strategic roads	Real-time journey planning and updates across all roads
	Real time parking space availability for on and off-street car parks	Real-time journey planning and updates for drivers
	Real-time parking charge data for off-street and on-street car parks	Enhanced journey planning and true journey price comparisons across modes
Freight connections	Real-time vehicle (road, rail, air, sea) locations and capacity information	Optimised fleet utilisation across logistics providers
International Connections	Real-time air and Eurostar departure and arrival time information data feeds	Ground transport and handling optimisation, reduced passenger waiting

Theme	Dataset	Services enabled
Public Transport Services	UK-wide local bus fare data	True journey price comparisons across modes
	Bus and rail vehicle occupancy data from passenger ticket purchase, barriers, electronic ticket machines.	Smarter traveller information services to reduce passenger waiting and avoid overcrowding
	Real-time bus and rail vehicle locations across the UK	Smarter traveller information services to reduce passenger waiting
Consumption & transaction data	Aggregated historic passenger ticketing data for bus, rail, air, sea modes of travel	Optimised service and capacity planning across networks, smarter journey planning for travellers

- 3.6 The ‘missing’ datasets identified in Table 3-2 occur at different levels. In some cases the datasets do not exist at all (e.g. on street parking bay availability), in others the data exist in silos (e.g. automated cycle count data, UTMC traffic flow data) or are not open/available (historic passenger ticketing data). Most commonly the datasets are only localised, rather than consolidated at a national level (e.g. databases of major events held by local authorities).
- 3.7 The technical challenges to opening-up these datasets and making them available often prove to be relatively trivial. These kinds of challenges typically relate to the use of common formats for sharing the data; and introducing accurate time, location, and date-stamping at the point of collection. Instead the main barriers identified through this review are primarily institutional and commercial, with organisations variably:
- Reluctant to make data available that may be commercially valuable at a later date.
 - Lacking, or not recognising, the rationale or reason for opening-up data they hold,
 - Lacking the in-house skills they need to create data feeds using common data formats, which often require data conversion and pre-processing effort.
 - Lacking the financial resources to maintain and curate a reliable data feed (e.g. on an open, or commoditised basis) due to uncertainty over the business model/value created for the data owner.
 - Concerned that making data available – even in anonymised and aggregated form – may contravene Data Protection Act legislation, on the basis it could be interpreted as misuse of individual’s sensitive personal data when not covered by existing terms and conditions.

Working with government agencies, private sector transport operators, and the logistics sector to ensure these datasets become widely available (ideally as linked open data) should be an immediate priority. Ensuring the available datasets are easy to discover, and well promoted, in-turn helps ensure that added-value service providers can readily find and exploit data to enhance intelligent mobility and create value through their activities.

A key issue here are the conflicting priorities that exist within local and central government agencies. The value of acting as a data service provider, and making reliable public datasets openly available, is a less well understood role than traditional public sector transport activities such as filling potholes and subsidising bus services.

Key transport-related datasets for intelligent mobility

3.8 Reflecting on the mega trends identified in Chapter 2, and the series of roadmaps developed for each intelligent mobility theme (Chapter 9), a total of 19 core datasets for intelligent mobility emerge as key to the evolution of intelligent mobility. They are listed in Table 3-3, and also appear in at least 3 of the 5 intelligent mobility roadmaps presented in Chapter 9.

Table 3-3: Key transport-related datasets for intelligent mobility

Theme	Dataset	Temporality
Place & Space	Internationally consistent map data that geospatially locates: Points of interest and places. Public transport (bus, rail, air, sea) stop and interchange locations. Vehicle parking locations. Vehicle fuelling/charging points. All roads and railway lines used to move people and goods.	Updated whenever the physical environment changes.
Environment	Weather data.	Historic, current, and predicted.
	Air quality	
People, Things & Movement	Personal location data (both aggregated anonymously and person-specific).	Historic, current, scheduled, and predicted.
Disruption and event-related data	Disruption/incident data relating to all transport networks (road, rail, air, sea) - including roadworks and incidents.	Historic, current, and scheduled.
	Planned cultural and sporting events / mass gatherings (e.g. demonstrations).	
Public Transport Services	Transport network capacity data for vehicles and people, covering all modes of transport.	Historic, current, scheduled, and predicted.
	Transport service operating schedules across all modes and networks	Historic, current, and scheduled.

Theme	Dataset	Temporality
Personal Automobility	Vehicle location data.	Historic, current, and predicted.
Freight connections	Network-wide capacity data for goods, covering all modes of transport.	Historic, current, and planned.
International Connections	Network capacity data for vehicles and people, covering all modes of transport.	Historic, current, scheduled, and predicted.
	Service operating schedules across all international modes and networks	Historic, current, and scheduled.
Consumption & transaction data	Fare and pricing data for all modes of transport (including roads and cars).	Historic, current and future.
	Sentiment data from social networks.	Historic & current.
	Service usage data from third parties and mobility service providers (e.g. car share / car club providers).	Historic, current, scheduled & predicted.
	Payments and purchases for goods and services, including mobility services.	Historic, current, and scheduled.

3.9 The most challenging gaps to fill from Table 3-3 are expected to be those datasets which are currently not available publicly, or openly, in any form. These include:

- Personal data relating to an individual’s locations, purchases (including mobility services), and schedule of appointments – which indicate intention and demand for travel, as well as demand for the movement of goods and services to them. These data typically all exist already, but are held in multiple scattered databases and social networks. In this context ‘Big Data’ have been described as being ‘in the dark’²⁴. [Google Now](http://www.google.com), based on Google’s closed datasets, is tipped to provide the first travel-related service that integrates individual’s lifestyle needs with mobility information.
- Planned cultural and sporting event schedules – which are also indicators of demand for travel, and potential sources of network disruption.

²⁴ John Balen, in TechCrunch (2014) *We’re still travelling like it’s 1996*. Available online at: <http://techcrunch.com/2014/05/10/why-are-we-still-traveling-like-its-1996/>, last accessed on 14/05/2014.

- Transport network capacity data for vehicles, people and goods; covering all modes of transport – current datasets define disruptions, speeds and journey times across individual networks; rather than capacity on a network-wide/geo-located basis.
- Fare and pricing data for all modes of transport – not currently available for buses, cars (e.g. running costs, tolls, and parking), air, or sea travel, as a consumable data feed.
- Service usage data from third parties – which are also an indicator of travel demand.

Challenges to filling intelligent mobility data gaps

- 3.10** In some cases the sensor networks needed to capture these data are not yet present (e.g. accurate capacity data from public transport services that do not require pre-booking such as bus and rail services). While this could currently be inferred from ticket sales data or anonymised cell phone location data; more accurate future measures are possible from aggregated flows of data derived from weight-based sensors mounted in-vehicle or in-seat, software-enabled passenger counts from on-board CCTV, and counting devices that are connected to Bluetooth/WiFi networks on board vehicles in service.
- 3.11** A further barrier to filling intelligent mobility data gaps identified above relates to the temporality with which transport-related datasets are collected, maintained and shared. Most of the practitioners engaged through this study observed that data collected in real-time are not currently being stored by the data owners for time-series analyses of historic trends. This is something that is yet to be addressed by demonstration projects such as Stride and i-Move, but which will be critical to achieving connected and user-focused mobility services.
- 3.12** In many cases the datasets considered key to intelligent mobility are owned by private sector enterprises. They either:
- Use the data being generated to operate their business (e.g. [Uber's](#) virtual car service is dependent upon user interactions with its smartphone app)
 - Create it as part of their organisation's 'digital exhaust' of interactions with their customers (e.g. rail and bus operator's ticket sales).
 - Exist specifically to collect and collate datasets they can then sell (e.g. INRIX, TrafficMaster).
- 3.13** This is a potential barrier to the data ever being shared openly. Given private sector profit maximisation imperatives, it will rely on the organisation (or third parties seeking to use the data) identifying a compelling business model that will see them remunerated for third party uses of their data that deliver network-wide benefit.

A total of 19 key transport datasets for intelligent mobility have been identified in Table 3-3. They are considered critical to the emergence of user-focused services that more intelligently connect people with the places, services and goods they demand. Around half of these datasets identified do not currently exist publicly in any form, and none yet demonstrate the scale or international-consistency desired by application developers and new mobility service market entrants.

Establishing sensor networks, crowd-sourcing mechanisms and reporting tools that ensure these key datasets are captured and made available in consistent formats will be key requirements for intelligent mobility. As well as being available in real-time (or near real-time), it will also be necessary to store these data in time-series. This creates scope for predictive analyses, and network-wide optimisation, based on historic operational insights and usage patterns.

Some of the most valuable datasets identified for intelligent mobility are those which relate to the locations, consumption patterns, and schedules of individuals. When aggregated these datasets offer potential to build a comprehensive picture of the demand for movement of people and goods. Currently these datasets are owned primarily by the private sector, and in only a few examples are the data being made available as-a-service for third parties to use (e.g. INRIX cell phone movement data). The political and ethical challenges the public sector faces in assembling these types of datasets, and the need for Key Performance Indicators on the quality of such data, are significant barriers to public open data of this nature ever being released.

This emphasises the need for dialogue between the major public sector (transport coordinating authorities and agencies) and private sector (major banks, retailers, transport operators) data owners around the value and business cases for opening up, trading, and sharing data for mutual benefit.

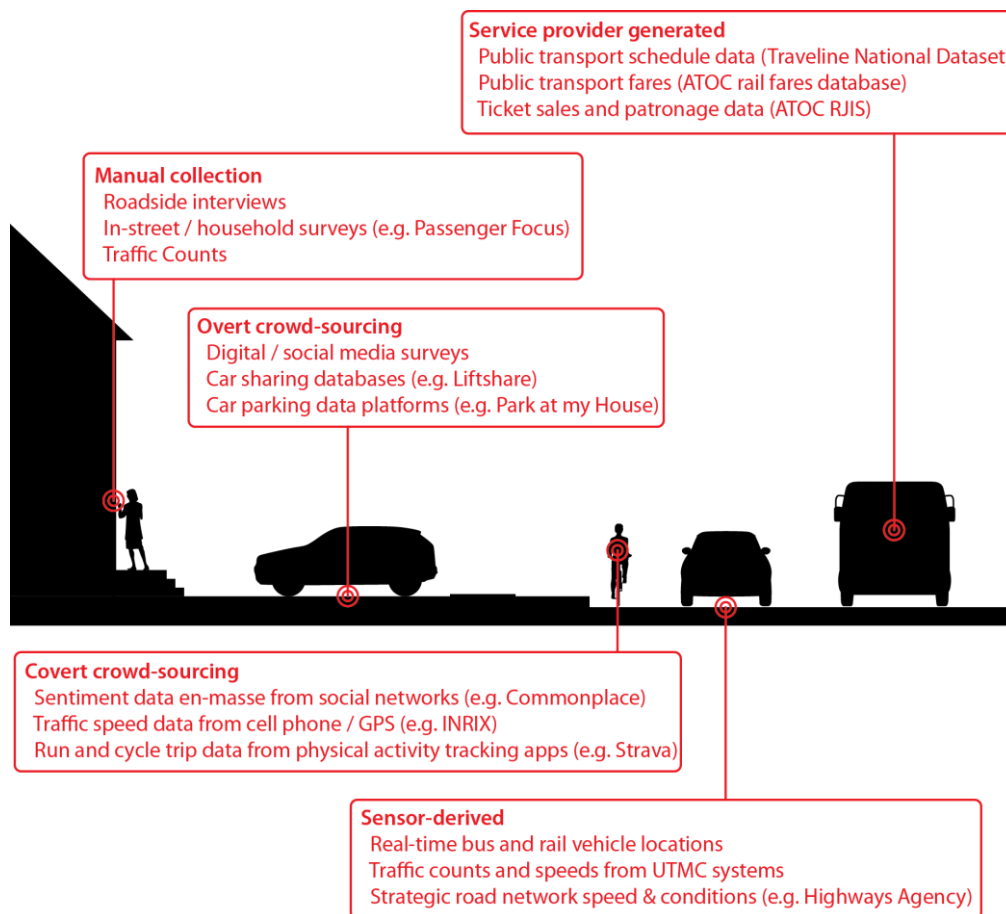
4 How are transport-related datasets created and shared?

- 4.1 The way transport-related data are created and shared is evolving. A growing network of web-connected sensors and opportunities to crowd-source data, in particular from user-based applications and personal communications devices, are resulting in large and constantly updated datasets that relate to transport and personal mobility.
- 4.2 These datasets exist alongside conventional data collection techniques, such as surveys and offline sensors. The data often sit in organisational silos and generate operational data to power discrete critical transport systems like Urban Traffic Management Control and railway network signalling, or deliver insight for strategic planning. This chapter summarises our understanding of the evolution of transport data collection and distribution.

Mechanisms for data collection

- 4.3 Figure 4-1 illustrates the key mechanisms through which transport data are collected.

Figure 4-1: Mechanisms for transport data collection



- 4.4 Historically, transport data have been collected manually from fixed points in space - often by counting ‘things’ (e.g. passengers, vehicles) for the purpose of keeping financial records, documenting usage and service reliability, and planning new infrastructure. More recent state-of-the-art has been for system control infrastructure built-in, or retro-fitted, to the network (e.g. rail-signalling, urban traffic management and control systems, ticket barriers) in order to gather operational data. These data enable increasingly automated decision-making, electronic payment systems, improved capacity, reduced costs, and more intelligent transport planning. They have typically been made available retrospectively or in ‘slow time’ and, until around 10 years ago, largely remained geospatially static.
- 4.5 GPS satellite technology, digital forms of data collection, faster communications networks, and increasing computer processing power have accelerated data collection velocity, volume, and variety²⁵. The time, date, and place accuracy of such data has also increased as a result of these technological advances. Taken together, they are enabling three important new mechanisms for data collection (see Table 4-1) that offer scope to provide deeper insights into how transport systems work, and people interact with them:

Table 4-1: Emerging transport data collection mechanisms

Web connected fixed sensors (individual IP addresses, or web-addressable)	
Track the interactions of people and ‘things’ (e.g. vehicles) with static pieces of transport network infrastructure by measuring the status of the infrastructure.	Smart parking RFID2 sensors such as those being rolled-out in Westminster.
	WiFi and Bluetooth networks for tracking people’s movement (by device MAC address) across transport networks, such as those TfL has installed at tube stations and bus interchanges in London ²⁶ .
Web connected mobile sensors	
Track the movement and status of ‘things’ (e.g. buses, trains, planes) through space as they interact with pieces of transport network infrastructure.	Status, location and behaviour of fleet vehicles, e.g. as made available by DARTT for the Stride demonstration project .
	Real-time bus vehicle locations for Real Time Passenger Information systems
	Connected vehicles, which are being introduced through the EU’s eCall legislation in 2017/18.

²⁵ International Transport Forum (2013) *Big Data and Transport – Overview*. Available online at: <http://www.slideshare.net/OpenSkyData/big-data-in-transport-an-international-transport-forum-overview-oct-2013>, last accessed on 15/07/14.

²⁶ New Scientist (2012) *London’s Tube and Big Data: Underground Movement*. Available online at: <http://www.newscientist.com/cloudup/article/in415>, last accessed on 15/07/14.

Crowd-sourcing from personal communications devices	
Track the movement, status and sentiment of people as they move through space and interact with pieces of transport network infrastructure, the built environment, and services.	INRIX cell phone, Sat-Nav and GPS datasets, which include cell tower and on-call data brokered from two of the UK's major telecoms providers
	Google / WAZE crowd-sourcing user's movement data from Android devices and via the WAZE community Sat-Nav app
	STRAVA metro crowd-sourcing user's running and cycling activity from the STRAVA mobile app
	TravelAI 's software code for detecting and inferring user's travel behaviour in the background of other apps
	The Mobile Millennium real-time traffic information demonstration project run in San Francisco by the University of Berkeley, California

Personal/vehicle location data from crowd-sourced and mobile connected sensor networks are not yet being fully exploited. The greater level of insight they provide into the behaviour of people and movement of things (e.g. vehicles and goods) means they are expected to be a cornerstone on which intelligent mobility services will be developed and delivered in the future. Several potential challenges to storing, sharing, and exploiting these data are discussed throughout this chapter.

Who collects transport-related data?

- 4.6 In the UK the public sector, or the private sector providing services on its behalf, has been the primary source of 'official' transport data in recent times. These datasets have typically been 'complete', in the sense that they provide comprehensive coverage of the transport systems or networks they represent (e.g. national rail timetables, national bus stop locations) and are updated on a periodic basis. As noted earlier in this chapter, and documented in **Appendix B**, this is changing. As people and things become more connected, a wider range of partial datasets is becoming available.
- 4.7 In the UK, car parking provides an example of how multiple sources of partial data now exist. Where once a historic record of the numbers of vehicles parked in specific locations may have been collected and maintained by a local Council, parking data can now additionally be combined from across a variety of sources:
- UTMC car park systems publishing data for use by real time services e.g. in [Norwich](#).
 - 'Smart' parking bays that are networked by sensors embedded in the road surface, such as those being deployed by the [London Borough of Westminster](#) that provide data on the occupancy of on-street bays.
 - Crowd-sourced parking provider data from websites such as [ParkatmyHouse](#).

- Public and private sector parking providers who collect and share car park capacity data in real-time and make it available through services such as [Parkopedia](#).

4.8 The widening variety of data sources makes it harder to know which datasets are valid, and to ascertain what sample of a whole population (e.g. UK rail users) partial datasets represent. This has been referred to as the ‘signal problem’:

The ‘Signal Problem’

“(Big) Data are assumed to accurately reflect the social world, but there are significant gaps, with little or no signal coming from particular communities.”²⁷ Geographic incompleteness of data due to different levels of wireless connectivity, unequal access to technology, and the collection of partial data samples from multiple sources all contribute to this issue.

4.9 A good example of how this signal problem can manifest itself in datasets being used to improve transport services can be found in Boston, USA. The city’s [Streetbump](#) app uses accelerometer and GPS data from user’s smartphones to identify the locations of potholes to assist the City of Boston’s road repair teams to patch 20,000 holes a year. The unintended consequence of relying on this app to make repairs is that roads in lower income areas and in areas with higher concentrations of older people – where smartphone penetration is as low as 16% – do not receive the same degree of reporting as ‘younger’ more affluent neighbourhoods^{27, 28}.

How is transport data validated and curated?

4.10 In 2014 the onus for validation and quality control of publicly-available transport datasets remains with the organisations that created them, and the degree of data curation varies between publishers. The best examples relate to datasets which are being provided by organisations with a clear understanding of the value (monetary, or otherwise) of their data to third parties, and the value to their organisations of third parties using their data.

4.11 [TfL’s open data micro-site](#) provides a good example of how data-publishing organisations can demonstrate quality control and provide assurance of their data when sharing it publicly. As well as describing TfL’s commitment to open data, and the way its data are being presented, each feed contains terms and conditions for its use that protect TfL’s reputation, including downloadable documentation and metadata describing key components of the dataset – such as:

- How often a fresh copy of the feed is published.
- Maximum time permitted between capturing and displaying the feed.
- Maximum time the information can be displayed before being updated.

²⁷ Kate Crawford, Harvard Business Review Blog (2013) *The Hidden Biases in Big Data*. Available online at: <http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/>, last accessed on 16/07/14.

²⁸ Tim Harford, Financial Times (2014) *Big Data: Are we making a big mistake?*. Available online at: <http://www.ft.com/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdco.html#axzz2x9L8avpJ>, last accessed on 23/07/14.

The terms and conditions for several TfL data feeds also include branding restrictions, and style guides clarifying how they can be re-presented. This level of effort demonstrates a degree of care on the part of the organisation opening up their data, and genuine concern that their data are re-used in appropriate contexts. Similar well-curated datasets exist in the data.gov.uk, [Glasgow Open Data store](#) and [Nottingham Open Data](#) websites; but often alongside less well-documented instances of similar data.

How is transport data being shared?

4.12 Transport-related data that have been created and quality-controlled by both public and private sector organisations are increasingly published online through a growing number of portals. A combination of data catalogues and data platforms (sometimes referred to as ‘ecosystems’) are the resources that developers and data-driven intelligent mobility service providers rely on to access complex transport data-sets. We have categorised these into four ‘types’ in Table 4-2:

Table 4-2: Commonly-used transport-related data portals

Type	Description	Examples	Run by
Distributed data catalogues	Available data from multiple organisations are published via a searchable online catalogue. Datasets remain stored with the data owners.	EU Open Data Portal	European Commission
		data.gov.uk	Central Govt. (Cabinet Office)
		London Data Store	Greater London Authority
		Datahub/Transport	Open Knowledge Foundation Transport Open Working Group
Individual data catalogues	A single organisation’s data is published via its own online catalogue. Datasets remain stored by each data owner, and may also be accessible via other data platforms.	Open Data Nottingham	Local Govt.
		Redbridge	
		Glasgow Data	
		TfL open data	Transport Agency
		ATOC	Private sector
		Network Rail Data	
Consolidated ‘Big’ data platforms / ecosystems	Available data are consumed and stored (‘warehoused’) so they can be combined and served-	Met Office	Public Sector
		Stride	Private sector (seed-funded by government)
		i-Move	

	up on-demand from a single interface.	Google Maps APIs Places, Directions Distance Matrix, Geocoding, Weather	Private sector
		INRIX Traffic	Private sector
		TomTom	Private sector
Distributed data aggregation platforms / ecosystems	Provide a gateway into available data from multiple sources using APIs in order to allow for real-time fusion and consumption of datasets.	smart streets	Private sector (seed-funded by government)
		Transport API	Private sector
		ITO! World's 'Transport DMP'	
		ProgrammableWeb	

4.13 Of these different types of data source:

- Catalogues maintain an online record of transport-related datasets which are published to them, and do this using metadata that define key characteristics of the datasets (as described previously). Catalogues are typically cost-free to use, and often publicly supported to provide a searchable way of discovering open datasets that are currently available, or scheduled to be available in the near future.
- Platforms / Ecosystems make transport-related data available in machine readable formats – typically via Application Programming Interfaces (or APIs) – which can be queried on-demand to provide Data-as-a-Service. Some platforms (such as those supported through the [Technology Strategy Board's \(now Innovate UK's\) Internet of Things Ecosystem demonstrator cluster](#)) make machine-readable datasets available free of charge. Other private sector data service providers are establishing business models that remunerate the platform operator/data owner for opening up their datasets on a per-inquiry basis (e.g. Google), or for aggregating open/brokered datasets and making them available in consistent machine-readable formats (e.g. transportAPI and ITO! World). These platforms typically involve aggregation and manipulation of raw data feeds, which are techniques discussed under Stage 4 of the data journey, covered in the next section.

4.14 Identifiable trends that emerged through this study's workshops centred on the way data are currently being made available through catalogues and platforms, and the extent to which existing data-sharing arrangements hamper efforts to innovate. Key topics include:

- Data that are being published openly in the UK (i.e. for free) are generally at a national, or major city level. Little open transport-related data is currently available for smaller cities and towns, or the mobility connections between them. This is concentrating innovations in transport and intelligent mobility service development in major cities such as London, and preventing them from scaling to other parts of the country.

- Efforts to open-up public sector data are widely welcomed; but the wisdom of different arms of government (central and local) developing their own open data catalogues was questioned – e.g. there are already open data stores or catalogues for the [UK](#), [London](#), [Leeds](#), [Manchester](#), [Nottingham](#), [Glasgow](#), [TfL](#). Although in theory all these data should be made available as [Linked Open Data](#), so that their relationships to data from other sources can be defined and referenced using standard web technologies (e.g. [HTTP](#), [RDF](#) and [URIs](#)), this is not happening in practice.
- Raw transport network data feeds are being presented alongside PDF documents containing data and other reports that may be useful for oversight (e.g. pay of civil servants), but which have no direct value in an intelligent mobility context. This emphasises the fact that no comprehensive online transport-related data catalogue currently exists for the UK.
- This lack of a comprehensive transport data catalogue, and the publication of some data using inconsistent/non machine-readable formats, also means time is increasingly spent hunting around for open data from multiple sources and finessing it into shape rather than exploiting it. Considerable focus to-date has been expended on ensuring data are machine-readable, but it is equally important to ensure that people can easily discover and understand what datasets exist.
- Consequently, software developers appear increasingly prepared to pay for access to data aggregation platforms (such as [TransportAPI](#) and [ITO! World's DMP](#)), that provide machine-readable feeds of open data which are available in raw formats. While this adds cost to start-up business models, it avoids duplicative effort by multiple service providers to rationalise the available data into a useable format. It is unclear whether this represents a sustainable long-term role, since better use of standards and data transmission protocols (e.g. APIs) by raw data producers would make these data-service value chains more efficient and reduce costs for developers / intelligent mobility service providers.
- The lack of a comprehensive catalogue covering all types of transport data (free/open/available at cost) also means that valuable items of transport-related data that are not currently openly available (e.g. bus fares, precise real-time train locations, and urban traffic management control system data) are being overlooked rather than prioritised for release. It also means that, apart from directly brokering or purchasing data from its owner, there are currently no mechanisms for accessing transport-related private sector data - for example from transport operators, retailers, event organisers, logistics firms etc. This is already a barrier to developing intelligent mobility solutions that respond to consumer demands. To-date intelligent mobility service providers (e.g. Uber) are overcoming this lack of data by creating their own infrastructure for establishing consumer demand for travel and storing that data to determine historical trends. However, some of the transactional data held by major credit/loyalty card companies and retailers could equally be anonymised and combined to provide deeper insights into the reasons people travel at different times of day/days of the week.
- [Hypercat](#) is emerging from the [Internet of Things ecosystem demonstrator clusters](#) as a hypermedia catalogue format that allows vertically-integrated information hubs (in which data are stored and organised) to expose linked data descriptions of their resources. It is

designed for exposing information about IoT assets over the web, and aims to minimise the need for replicative human intervention behind each new App – Hub, or Hub – Hub, interaction. There is a strong, simple security model that uses widely adopted [HTTPS](#), [REST](#) and [JSON](#) internet programming conventions. This thin meta-data layer makes it easier for apps to either explore what is available on each hub (by crawling about the interface), or search for particular types of resource outright.

The idea of an ‘Argos Catalogue’ or ‘Amazon website’ that acts as the definitive location for people and machines to look for transport-related data products and services was discussed on several occasions during project workshops. The thinking was that such an interface should make it easy for people to discover what data are out there, and then enable them to download, enquire, or consume the data in both manual and programmatic ways. This would effectively act as the ‘shop window’ for transport data that will power intelligent mobility, and act as ‘honest broker’ between data owners and data users.

Importance of real-time data for intelligent mobility

- 4.15 In reality, all transport-related data (including the map-bases onto which transport data are being visualised) are now collected and updated in some form of time-series. For example, rail operating schedules change seasonally, while bus stop locations come into and out-of existence and are updated in the [NAPTAN](#) dataset on a weekly basis. Although some data (e.g. [Highways Agency National Traffic Information](#)) are being collected on a constant-basis for near-term and real-time decision making by transport network managers, such infrastructure-led data collection efforts are primarily focused on major cities and strategic road/rail networks, which have historically benefitted from sensor network investment.
- 4.16 Emerging datasets that have been crowd-sourced from user’s personal communications devices, or mobile web-connected sensors, are not bound in this way. Instead of providing a sample of data that relates to things passing fixed points in space; they track the movement of people and things as they interact with transport networks, the built environment, and services. These data are also collected and aggregated in real-time.
- 4.17 Near/real-time transport data collection and sharing/selling is expected to become the ‘norm’ over the next 10 years²⁹. An example of how this is already happening can be found in the Sat-Nav industry. [TomTom’s](#) map data updates have evolved from being distributed on DVD’s once every 6 months, to making the data available online for on-demand download with a commitment to a minimum of 4 releases per year. This is partially a response to [OpenStreetMap](#) and [Google Map](#) data being updated much more frequently (almost constantly) using crowd-sourcing techniques, partially due to the need for the map data to be current in order to prevent erroneous routing, and partially because technology now allows for these faster update cycles.

²⁹ Aliza Asherman, on Mashable (2011) *Why the Future of Transportation Is All About Real-Time Data*. Available online at: <http://mashable.com/2011/07/27/transportation-future-data/>, last accessed on 16/05/14.

- 4.18 Other transport service providers will likely need to be similarly responsive in the coming years, as the speed with which conventional transport-related datasets can be updated and relayed to travellers is outstripped by the rate and volume of:
- Sentiment data transmitted across social networks, such as that being mined by [CommonPlace](#).
 - Anonymised location data from personal communications devices, connected Sat-Nav devices, and connected vehicles; such as that collated by the [INRIX Traffic](#) app.
- 4.19 However, finding real-time transport data feeds is not currently that easy. Of the 1,095 ‘transport’ data results currently returned by searching on data.gov.uk³⁰:
- 21 are available as XML feeds, which are both human- and machine-readable and have potential to be updated on a very regular basis.
 - 3 of these are updated on at least a monthly basis.
 - 4 are refreshed on a weekly basis, and relate to the [Traveline National Dataset](#).
 - 3 are ‘real-time’ or ‘near real-time’ data feeds: [Gritting decisions by East Sussex County Council](#), the [Norfolk County Council’s live car park data](#) and the [Highways Agency’s live traffic information feeds](#).
 - None represent datasets that have been crowd-sourced from connected mobile sensors or personal communications devices.
- 4.20 As well as emphasising the importance of the points above related to effective cataloguing to allow for efficient data discovery (since other real-time data feeds are available from providers such as Traveline Nextbuses, Network Rail, ATOC, and TfL), this lack of readily-available real-time data, and descriptions of its use, is a key current barrier to intelligent mobility service provision. While not all aspects of transport data need to change in real time, data on current/predicted availability and capacity of transport networks (e.g. roads, parking, bus/rail services) in specific locations will be critical to delivering intelligent mobility services.
- 4.21 Three of the TSB’s Internet of Things demonstrator projects have successfully shown how it is possible for the public sector, and private sector subcontractors, to release relevant real-time data feeds in formats that are useful to developers - albeit only on a demonstration basis to-date:
- [i-Move](#) was successful at extracting data from local authority funded Urban Traffic Management Control (UTMC) systems in Merseyside, which count vehicles using inductive loops and use cameras to manage traffic signals at a network level, and incorporating them alongside other real-time data feeds. A total of 57 datasets were subsequently collated on the [Stride](#) platform, which was built upon technologies developed by BT and Aimes Grid Services and included the following real-time feeds:
 - TfL’s live data feeds for the Tube network in London.

³⁰ As found using the search criteria ‘transport’ on 29/05/14.

- All of the Highways Agency’s available live data feeds.
- Merseyside roadworks, average traffic speeds, car park occupancy from its UTMC system.
- Anonymised fleet-driver behaviour and vehicle data provided by [DARTT](#), with permission from two of their clients.
- Bus and train public transport data provided by [TransportAPI](#).
- Weather observation data from the BBC/Met Office for specific UK-wide locations.
- Road network and public transport event and incident (disruption) data from [Elgin](#).
- GPS derived journey time and event-driven disruption data from [INRIX](#).
- [Smartstreets](#) collated 34 datasets for inclusion in a similar style of data hub. While many of the datasets were also common to the Stride and i-Move platforms described above, it additionally included the following feeds:
 - Real-time UK air quality data from DEFRA.
 - Real-time UK Earthquake data.
 - Historical meteorological data from the Met Office.
 - Real-time road maintenance activities and road gully states provided by Carillion’s fleet of road maintenance vehicles operating in Redcar and Cleveland.

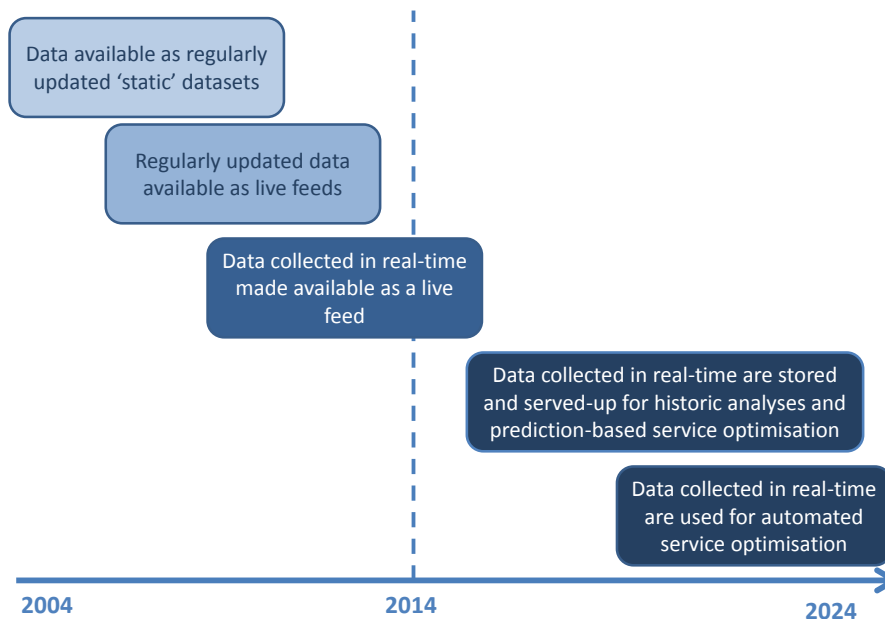
4.22 A common issue across all of these platforms is that, although real-time data are being exposed; they are only being selectively archived (e.g. [historical meteorological data from weather stations](#), maintained by the met office - albeit in text files). The faster sample-rates for data collected in real-time, or near real-time, make their retention for future analyses more challenging, because:

- The constantly increasing scale of these datasets means widely-used relational database management systems (e.g. SQL) for indexing, processing, and sharing the data can be too slow – adding unwanted latency to analysis and re-use tasks.
- The datasets become many orders of magnitude larger than for less dynamic data. As such they require larger data centres/cloud storage, or intelligent archiving techniques which pragmatically balance the storage requirements of data collected in fine-grain (e.g. every 15 second traffic counts by UTMC systems) with practical future uses – such as predictive analytics.
- If data collected in real-time are also intended to be used to power real-time analytical processes combining unstructured data (as they increasingly are in other fields); then scalable, distributed database management systems such as [Apache Hadoop](#) may be increasingly important technologies for intelligent mobility providers and transport network managers in the future. Although unlikely to be an immediate concern, the rising availability and falling costs of data storage are likely to enable the fine-grain storage of data collected in real-time within the next 10 years.

4.23 Each of points above adds cost and complexity to the process of archiving and retrieving real-time data, and currently represents a barrier to its widespread usage in transport contexts. Figure 4-2 provides an indicative pathway for the availability of real-time transport datasets to 2024. It is

based on the rate at which the transport data specialists engaged through this review anticipate that transport data will be collected, and the technologies will emerge to aggregate and process multiple real-time feeds:

Figure 4-2: Indicative pathway for real-time data collection and sharing



- 4.24 It suggests industry expectations are that within the next 2-3 years we will be in a position to use archived real-time data as the basis for network-level predictive transport pattern analysis. Using real-time data for autonomous service optimisation and delivery is expected to be at least 5 years away.

The role of standardised data formats

- 4.25 Formats used to encode transport data at a global level are highly variable and often reflect years of organic evolution in different geographic locations. In the UK they typically reflect *de-facto* standards whose development has been supported by the Department for Transport (DfT), other government agencies (Highways Agency) or transport industry bodies (Association of Train Operating Companies).
- 4.26 The transport-specific data standards currently used in the UK emerged because these different stakeholder groups identified a need for data to be collected and maintained in common formats to allow for interoperability. Much of the open transport data that is currently being published by the public sector, and the private sector on its behalf, would not be possible without these standards.
- 4.27 The main national transport data standards currently in use in the UK are documented in Table 4-3, overleaf, which is a summary of datasets contained in Appendix B. Of these standards, only DATEX II represents an international (EU) standard that allows for interoperable application of traffic data collected in real-time. Its development has been driven by public sector ITS industries of different European countries wanting to collect, maintain and publish data feeds in common file formats and web-service conventions to simplify both:

- The real-time processes associated with exchanging information between traffic management centres, traffic information centres and service providers across the network of strategic roads that make-up the [Trans-European Network, or TEN-T](#).
 - The process of sharing open traffic data/network status feeds with all service providers and actors involved in the traffic and travel information sector. However, many private sector providers do not use it as their client base is reluctant to pay for the overhead of adhering to a standard they don't use themselves.
- 4.28 These same desires across a wider-range of transport and non-transport related fields are anticipated by this study team to become key drivers for data standard development and convergence over the course of the next 10 years.
- 4.29 The summary of national datasets outlined in Table 4-3 is also mirrored for Transport for London (TfL) whose data cover a wider range of transport modes (e.g. London Underground, London Cycle Hire Scheme), and use recognised national standards such as NaPTAN and TransXchange alongside proprietary standards developed for London-specific datasets. In a national public transport context, most of the standards developed represent components of DfT's Public Transport XML Standards, which are based on the TransModel European Reference Data Model for Public Transport. This schema was developed in the late 1990's and early 2000's, and is in the process of being superseded by [NeTEx](#) (with involvement from the DfT) and [SIRI](#) (with involvement from the UK Real Time Interest Group).
- 4.30 Outside of the UK the [General Transit Feed Specification \(GTFS\)](#), originally developed by Google (and still used to power its Transit features) has gained considerable traction in the last 9 years and is being used in over 500 cities worldwide. GTFS is increasingly being used to model transport services operating in developing and transitional cities, as well as more economically developed locations where transport services operate on a scheduled basis. The simplicity of the standard, its community-driven nature, and the way it is maintained; have drawn many transport technology developers towards using GTFS data because they can create one set of tools and then simply manage multiple feeds of data packaged in an identical format. The collection of these data is providing city authorities with the means to better understand and rationalise private sector run collective transport services which operate with limited central coordination³¹

Table 4-3: Commonly-used national transport data standards and formats

Transport mode / data type	Data format / standard	Coverage	Geospatial reference	Maintained by:
Rail - Live Departure Boards	LDB webservice (via SOAP XML)	Great Britain	3 digit rail station codes	Association of Train Operating Companies

³¹ Eros (2014) *Transport Data as a Disruptive Innovation in Mexico City*. MIT Masters Dissertation. Abstract in MIT (2014) *Thesis/Dissertation Abstracts DIUSP Graduates 2013-14* (pg.20), available online at: http://dusp.mit.edu/sites/all/files/attachments/news/Thesis_Booklet_2014.pdf, last accessed on 28/07/14.

Rail Timetable Information Service	Timetable Data Interface (using CIF format)	Great Britain	Timing Point Locations (TIPLOC)	Association of Train Operating Companies
Rail Fares data	Rail Journey Information Service (RJIS)	Great Britain	3 digit rail station codes	Association of Train Operating Companies
Public transport stop locations	National Public Transport Access Nodes (NaPTAN)	Great Britain	OS grid reference	Landmark, on behalf of DfT
Timetable data for bus, light rail, tram and ferry services in Great Britain	TransXchange / Traveline National Data Set (TNDS)	Great Britain	OS grid reference	Traveline
Schedule-based & real-time multi-modal journey planning	Journeyweb	Great Britain	OS grid reference	Trapeze Group on behalf of DfT
Transport mode / data type	Data format / standard	Coverage	Geospatial reference	Maintained by:
Next available bus/tram service from a nearby stop	Nextbuses API	Great Britain	Postcode, street & town or stop name & town	Traveline
National Traffic Information Service and local city data	DATEX II (UML)	England and many UK cities	Lat/Lon	Mouchel & Thales on behalf of Highways Agency, and some local UTMC systems
Topographical database of UK towns and settlements	National Public Transport Gazetteer (NPTG)	Great Britain	OS Grid reference	Traveline
National cycle path network	CycleNetXChange	Great Britain	OS Grid reference	Trapeze Group on behalf of DfT

- 4.31 Irrespective of the merits of the different data standards outlined above, the current challenge they all face stems from the global reach of common software development platforms - particularly iOS, Android and Java web-environments - which are creating:
- Universal consumer demand for transport information presented in a simple multi-modal way that people can understand (e.g. “how do I get home from here?”, rather than “which train do I need to catch?”).
 - Devices (smartphones, tablets, wearable tech.) which provide the means of finding this information wherever it exists, and augmenting it with users’ location-based data.
 - Demand from developers and service providers for more standardised and platform independent transport data feeds so that they can create scalable software tools and mobility services for use in a wide range of geographical markets.

Challenges for evolving transport data creation and sharing

- 4.32 In the context of evolving intelligent mobility practices and services to 2024, the following challenges exist for transport-related data creation and sharing:

Understanding and maintaining open data feeds

- 4.33 Only the best-resourced public authorities and private sector organisations are currently able to take the level of care demonstrated by TfL over data they hold, and invest the effort needed to publish it in a way that assures consumers of its quality. As public sector budgets are reduced, the core business (e.g. for a Local Authority - pothole repairs and basic road maintenance) remains a key part of the local agenda, while the benefits of opening up data are less of a focus for local decision makers.

Developing and adhering to interoperable (increasingly global) data standards

- 4.34 Despite representing a national example of good practice, and having the richest open transport datasets in the UK, TfL itself has acknowledged³² some leading software developer’s concerns³³ that its data are complex to work with. Key reasons behind this are the wide range of systems from which these data are created, and TfL’s partial adherence to existing UK transport data standards alongside their own tailored data protocols.
- 4.35 The issue of transport-related data interoperability – at a global level – represents a significant challenge to the development of data-driven intelligent mobility services over the next 10 years. Due to their scale and product-reach the major global technology firms (e.g. Google, with their [Transit](#) and [Directions](#) products) appear able to drive interoperable transport data standards such as GTFS more effectively, rapidly, and globally than existing industry bodies have to-date. This

³² Techworld (2013) *TfL to release new API for London travel app developers*. Available online at: <http://news.techworld.com/applications/3475779/tfl-to-release-new-api-for-london-travel-app-developers/>, last accessed on 17/07/14.

³³ Techworld (2013) *London’s Citymapper capitalises on “decade of data”*. Available online at: <http://news.techworld.com/sme/3467301/londons-citymapper-capitalises-on-decade-of-data/>, last accessed on 17/07/14.

can challenge transport data providers' desire to remain platform and supplier-agnostic by using common open data standards.

Ease of data discovery

- 4.36 Understanding what datasets are being collected, by whom, what they represent, and if they could potentially be made available for third parties to use will become increasingly important as crowd-sourcing and IoT sensor networks develop – widening the range of available raw transport datasets. This was highlighted by the amount of time our research team expended through the course of this study searching for transport-related datasets across multiple catalogues and data platforms.
- 4.37 A comprehensive catalogue (such as that proposed earlier in this chapter) actively promoting links to all the available (i.e. open and free, open and licensed, available for a fee) transport-related datasets, and processed feeds created from the raw data sources, is something that is likely to accelerate the establishment of a functioning marketplace for transport data (and associated products and services) in the UK.

Maximising the collection and archiving of real-time data

- 4.38 To facilitate the delivery of intelligent mobility services by 2024, there appear to be 3 key stages in respect of the evolution in real-time datasets:
- Increasing the volume of transport-related data that is collected and shared in real-time, or near real-time using existing/emerging common global standards. This will be increasingly important as the number of IoT sensors and connected devices increases.
 - Archiving these data indefinitely, and using pragmatic archiving protocols (e.g. aggregating across increasing time horizons as data gets 'older' and less relevant) that will not preclude offline data mining, historical analyses, and probabilistic modelling based to derive insight from past trends and patterns.
 - Developing industry capabilities and processes that enable transport data, and related data, collected in real-time (related to both demand and supply-side factors) so they can be harnessed in real/near-time to power autonomous systems and network-optimisation.
- 4.39 These three evolutionary stages for real-time data conceivably affect all 5 of the intelligent mobility themes, but are particularly crucial enablers for autonomous systems, information exploitation, and resilience.

The challenges identified above present opportunities to accelerate the uptake of intelligent mobility practices, services, and operational approaches through smarter data collection and sharing. These opportunities are documented and discussed further in Chapters' 6 and 9 of this report.

5 How are transport datasets being exploited, and by who?

- 5.1 The transport datasets discussed in previous chapters are being used in a myriad of ways to provide value-add services to consumer, business, and government sectors. This aligns with the finding that government open transport datasets are the 4th most widely applicable data category (relevant to 12 other sectors) behind only geospatial, environmental and economic data¹¹.
- 5.2 The outputs documented in this chapter represent an aggregation of different examples of current mobility-related services (of varying degrees of ‘intelligence’) identified through our review. We define a high-level set of service-type groupings based on current forms of mobility and use it to identify organisations we learned about through the course of the study. Further details of these transport data service providers, including contact details, are included in **Appendix C** to this report.

Defining transport data services in relation to mobility services

- 5.3 Previous research by [Deloitte](#) and the [Open Data Institute](#) identified a set of common ‘archetypes’ for open data business models, which related to the burgeoning open data marketplace of 2012¹¹. These focused on open data in general, and its applications, rather than all transport data - e.g. including that available, but for a fee/with licensing restrictions.
- 5.4 Based on our learning accumulated through this review, we adapted these archetypes to define five types of transport data service (See Table 5-1), which are relatable to the transport sector in a similar way to the open data marketplace.

Table 5-1: Transport data service types

Type	Description
Suppliers	Organisations that publish their raw data so that others can (re)use it.
Aggregators	Organisations that collect and aggregate data (both open and proprietary), find correlations, identify efficiencies or visualise complex relationships.
Developers	Organisations and software entrepreneurs that design, build and sell web, tablet or smartphone applications for individual consumption.
Enrichers	Organisations that use data to enhance existing products and services, and/or deliver operational efficiencies, through better insight.
Enablers	Organisations that facilitate the supply, publication, sharing, and use of data.

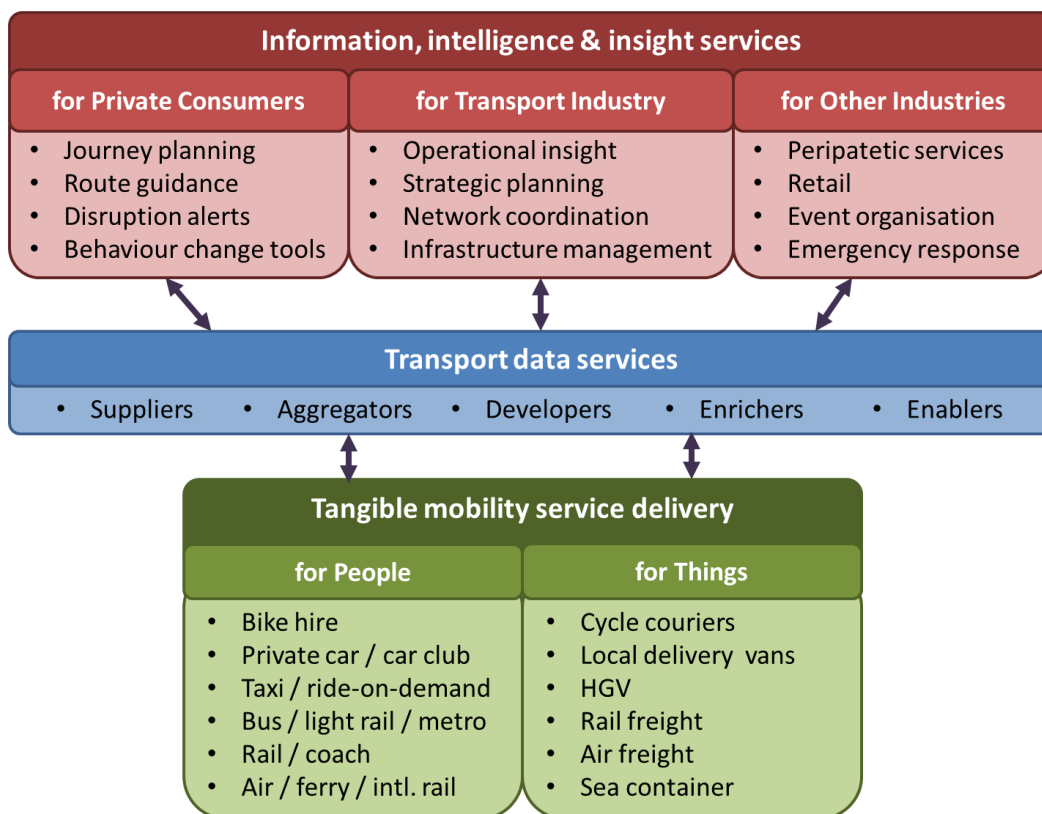
Source: types and descriptions modified from *Deloitte and ODI (2012)*¹¹

5.5 Transport data services touch each of the identified mobility service areas and are expected to provide the basis for deeper integration within, and between different forms of mobility over the next 10 years. In practice, many transport data service providers fulfil more than one of the roles defined in Table 5-1, and are also delivered both:

- Within a transport data ‘service area’ – e.g. organisations refining and adding value to raw data as a service and selling it on to other data companies for further processing, analysis, and/or re-use.
- To other mobility-related ‘service areas’ – e.g. transport data service providers selling data, insight, and analytics to public transport companies, fleet operators, vehicle manufacturers.

5.6 Figure 5-1 illustrates the groupings of mobility service types that relate to current transport systems, but could also apply to intelligent mobility systems that emerge in future:

Figure 5-1: Typology of mobility service types influenced by transport data



5.7 For reasons of presentation, Figure 5-1 is not intended to provide an exhaustive breakdown of every different mobility service component. As such the examples of ‘Information, intelligence and insight services’ listed under the ‘private consumers’, ‘transport industry’, and ‘other industries’ sub-headings represent the most commonly identified services.

5.8 In developing Figure 5-1 we assumed that ancillary services such as ticketing and passenger assistance are included within tangible mobility service delivery for people. In the future it is possible that some components (e.g. payment systems) may break free of conventional transport service delivery mechanisms and become a form of mobility service type in their own right. For

example, the cashless payment system such as that being developed by [Droplet](#)'s smartphone app and E-purse, could become widely adopted by all mobility service providers and stand alone as a form of service used by both private consumers and tangible mobility service providers.

- 5.9 In the remainder of this chapter we discuss each service-type in-turn, identifying who we found is using transport data to deliver them, and providing example 'data journeys' that illustrate how different datasets are being utilised and exploited. We have focused on service providers, rather than data providers, since they are covered in Chapters 3 & 4.

Transport data services and providers

- 5.10 Relatively few services we identified can be considered solely as



transport data services. The majority also provide insight/information or tangible services related to the movement of people and things. They either utilise data they gather about client's movements to add-value to their existing services (with some making this data available to third parties as a secondary business activity), or are data-driven and rely on source data about the movement of people and things (demand) to provide their service.

- 5.11 The market for transport data-as-a-service, particularly in relation to public transport data, is at an early stage of development. The organisations and web-based platforms identified in Table 5-2 are those seeking to become foundations on which data-driven intelligent mobility systems are built. They typically divide into two groups:

- **Public open data catalogues** that maintain online repositories of open data related to a wide range of topics and themes. Transport is just one of many topics on which data are being shared freely and openly. [Data.gov.uk](#) and [Datahub](#) are examples of these, and act as digital 'signposters' – pointing to where data can be found.
- **Specialist transport data platforms/ecosystems/warehouses;** which combine open data alongside those which are available, but only accessible at cost or by accepting licensing obligations. These platforms seek to provide transport data as a service, typically made available via an API or XML feed, and aggregate public open data feeds (such as the Traveline National Dataset, or Highways Agency's National Transport Information Service) alongside datasets being commoditised and sold by the private sector. [Stride](#) / [i-Move](#), [QCumber](#), [traveline's NextBuses](#), and [transportAPI](#) are examples of this type of service.

- 5.12 Among the specialised transport data service platforms transportAPI is unique in re-processing most of the national open public transport datasets³⁴. These are freely available from multiple sources, but can require considerable effort on the part of software developers and mobility service providers to assimilate and maintain due to different data formats and update cycles. This data maintenance task detracts from developing compelling products and services that help people and

³⁴ [ITO! World](#) is also doing this; but additionally provides insight, public information, and analytical services so is covered later in this chapter.

things move around, and is a source of frustration for innovators working to create new sources of transport information and insight. transportAPI's work to aggregate the open datasets using automated data batch processes means it is able to provide a single enquirable API covering bus, tube, light rail, and train services and stops/stations for the whole of Great Britain. This simplification adds value to the raw open data feeds, and enables transportAPI to charge on a per-hit basis for access to the re-processed open datasets consumed via its API.

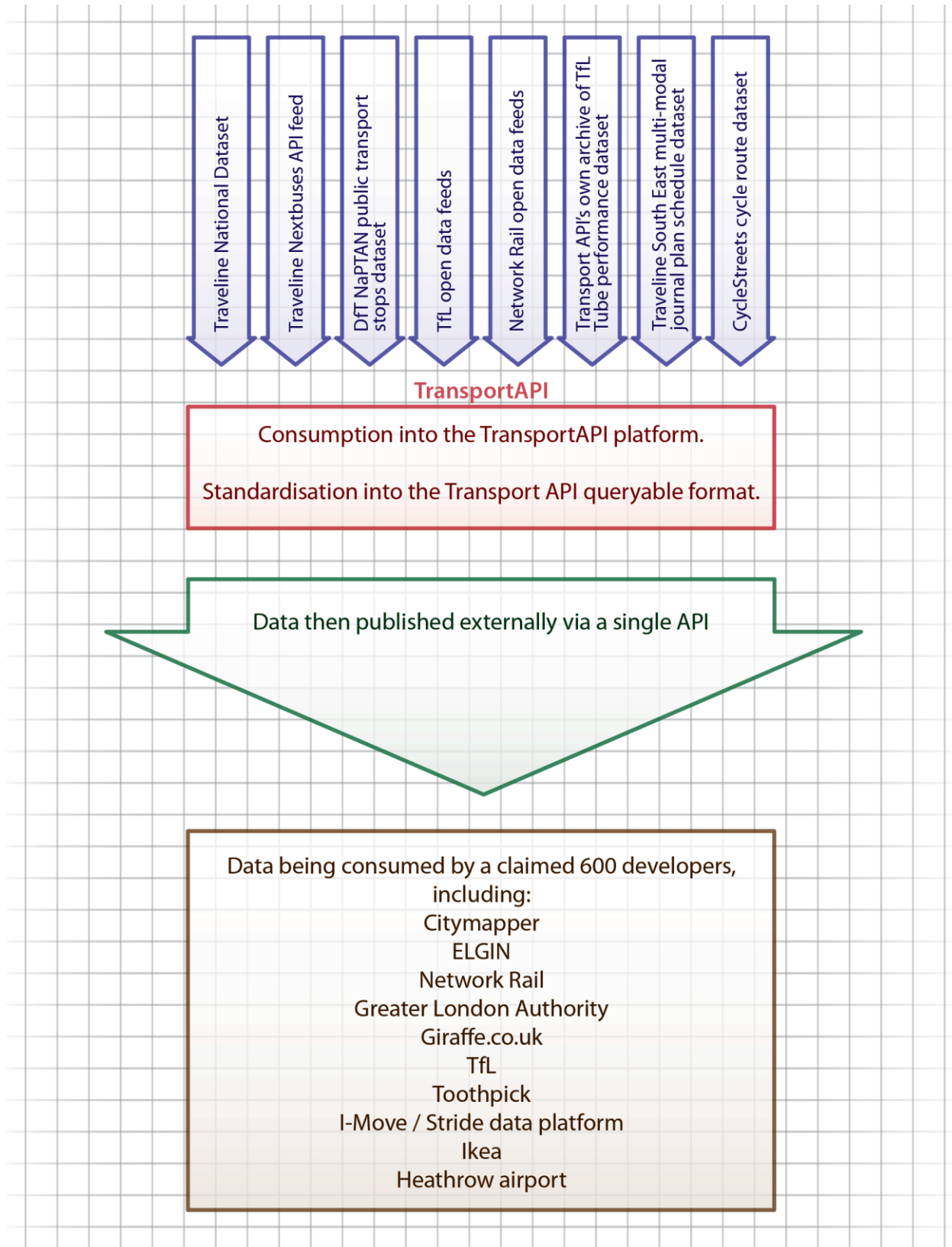
5.13 Figure 5-2 illustrates the 'journey' that open data take from source through transportAPI's platform, and onwards into applications and products.

Table 5-2: Examples of transport data services and providers

Provider	Service description	Use of transport data	Service types
<u>QCumber</u> (CERC)	Prototype smart city data platform for Birmingham, Cambridge, Ipswich and London.	Aggregates road and rail network; traffic counts; road CO ₂ emission; local road, rail, and airport air quality emissions; bus stops; road accidents; cycle routes; cycle parking; gritting routes; walking path data from a combination of local and central government sources.	Aggregate, Enable
<u>Stride / i-Move</u> (TSB funded Internet of Things demonstrators)	Internet of Things (IoT) ecosystem for cloud-hosted transport data and applications.	Aggregates Dartt anonymised driver behaviour data; live national road, rail and bus service data, weather observation data; Liverpool journey time, roadworks, average traffic speed, and car park occupancy data; TfL tube data, TransportAPI data, Dartt client fleet data, UK petrol stations, ELGIN incident alerts, INRIX traffic link road event data. Creates XML/API access to all above datasets through Stride info hub .	Aggregate, Enable
<u>transportAPI</u> (Placr)	Collate and provide UK transport data via an API	Aggregates all UK open public transport data (traveline national dataset , TfL datafeeds , Network Rail , Twitter tweets) Supplies 600 UK developers, including Elgin and Citymapper, via API.	Aggregate, Enable
<u>TravelAI</u>	Automatic inference of individual’s travel patterns based on background app technology.	Uses map and place of interest datasets; open public transport data from traveline national dataset, network rail, TfL and highways agency. Collects and Aggregates user travel patterns from accelerometer, GPS and smartphone sensors to infer travel mode from transport service data. Creates APIs for developers to insert into their apps for context awareness insights and a ‘living’ journey planner.	Aggregate, Enable

Provider	Service description	Use of transport data	Service types
Data.gov.uk	UK open data catalogue	Collates metadata and links to stored datasets for all UK open data	Enable
Datahub	Global open data catalogue	Collates metadata and links to stored datasets for international open data	Enable
traveline, NextBuses	Public transport timetables for bus, light rail, tram and ferry services in Great Britain	Aggregates local transport schedules for UK local public transport services (excluding rail and coach services) in the TransXChange format. Data collected by local authorities and transport operators are compiled and published by 10 Traveline regional organisations in XML format. Real-time and scheduled data are also available via the nextbuses API.	Supply, Aggregate

Figure 5-2: Data journey through transportAPI's platform



Information, intelligence and insight services

5.14 Many of the transport data service providers identified through this review are involved in both the provision of transport data as a service and its exploitation to improve transport information and actionable insight for consumers; the transport industry, and other industrial sectors that interact with transport systems.

For private consumers

5.15 The development of data-driven information, intelligence and insight services for consumers has centred on user-based web and smartphone apps. The examples documented in Table 5-3 typically fulfil one, or both, of the following roles:

for Private Consumers

- Journey planning
- Route guidance
- Disruption alerts
- Behaviour change tools

- Providing personal mobility information as a service that is contextually relevant to consumers, and derived from public open datasets (e.g. train timetables and fares) combined with crowd-sourced datasets (e.g. live traffic conditions) – e.g. [Citymapper](#), [MyCityWay](#), [tubechecker](#).
- Crowd-sourcing data on personal movement and travel behaviour/activity using smartphone sensors (GPS, accelerometer, camera) to establish datasets that can subsequently be leveraged for individual (and aggregate) analyses and visualisation – e.g. [Strava](#) / [Strava Metro](#), [JustPark](#) / [Parkopedia](#), [Life 360](#).

5.16 In both cases, these consumer-oriented services are increasingly reliant on providing web and smartphone applications free to consumers at point-of-use, or using ‘freemium’ models – where premium features are charged-for. The reasons behind this are primarily because:

- Global technology and car manufacturing companies such as Google, Apple and BMW are giving similar consumer services away for free, or integrating them deeply within the software platforms on which their personal communications devices and vehicles function. Consumers therefore expect mobility information services and lifestyle tool to be free.
- Many personal information, insight, and intelligence services are being provided by relatively new companies competing to achieve ‘scale’ and grow a viable user base. Giving a product or service away for free, and funding its sustained development through advertising revenue or data sales (or venture capital funds, betting on a future market for the data/sale to a major technology company), is now common for data-driven personal mobility information and insight services.
- The crowd-sourced/online community-driven nature of some tools (e.g. [Waze](#)) requires that a sufficiently large and varied sample of people use them at least occasionally for the core service to be valid. Such services typically need to be free in order to attract and maintain users (data probes), and to compensate for the fact the service may not initially be as powerful or useful as existing paid-for services (in the case of Waze - commercial Sat-Nav products from the likes of TomTom and Garmin). Vehicle makers are not adopting crowd sourced traffic information as they have issues with coverage and content reflecting badly on their overall product.

Table 5-3: Example information, intelligence and insight services for private consumers

Provider	Service description	Use of transport data	Service types
JustPark / Parkopedia (BMW)	Online marketplace for commercial and privately owned parking spaces (formerly known as parkatmyhouse.com)	<p>Uses Google Map and Places APIs to overlay parking locations and availability in real-time.</p> <p>Aggregates parking space locations, price and availability data on real-time from public and private car parks, as well as individuals.</p> <p>Creates usage data and insight into parking bay usage patterns and value.</p>	Supply, Aggregate, Develop, Enrich
MyCityWay (BMW)	Free e-travel guide app. Provides customised wayfinding and lifestyle information. Integrated into BMW iDrive service.	<p>Uses live city-level data feeds on traffic conditions, transport services, Google Places API, TripAdvisor API, weather, social network trends, fuel price, local 'what's on'.</p> <p>Creates usage data and insight into urban dwellers movement patterns and sentiments.</p>	Aggregate, Develop, Enrich
ChargeNow (BMW)	Free app and integrated service for BMW i-vehicles that identifies electronic vehicle charging points and their availability.	<p>Uses Google Maps and Places APIs to overlay electric vehicle charging point locations collated from city authorities and national datasets.</p> <p>Aggregates real-time data on EV charging point location usage.</p> <p>Creates usage and payment data for insight into electric vehicle parking and charging patterns.</p>	Supply, Aggregate, Develop, Enrich
Embark (BMW)	Free app providing transport planning journey planning for 10 US cities.	<p>Uses Google Maps and Places APIs, and GTFS scheduled/real-time transit data feeds.</p> <p>Creates user data on origin-destination of public transport trips.</p>	Develop, Enrich

Provider	Service description	Use of transport data	Service types
Life 360 (BMW)	Free app for group location tracking, check-in & sharing service.	Uses Google Maps and Places APIs to overlay personal and group location data. Aggregates multiple users GPS location feeds and check-ins.	Develop, Enrich
Citymapper Transit App OTP Hopstop	Free apps for personal journey planning. Provide free journey mapping integration for other apps.	Uses Apple Maps, OpenStreetMap, Foursquare, TransportAPI's transport service feeds derived from the traveline national dataset, TfL live data feeds, ATOC's national rail data and Cyclestreets cycle routing API. Creates user location dataset, origin-destination intention and completed journey path data from GPS traces.	Develop, Enrich
FatAttitude	Freemium public transport smartphone apps for tube and bus users in the UK.	Uses TfL real-time tube and bus feeds, TfL bike hire station and real-time availability, Traveline National Dataset, Google Maps and Places APIs. Creates BusChecker , tubechecker smartphone apps.	Develop, Enrich
Strava Strava Metro	Freemium running and cycling activity tracking, metro data service	Collects and aggregates user-generated data on running and cycling activities from smartphones and other GPS devices. Provides freemium activity tracking tools to encourage personal data recording. Creates anonymised geo-located activity datasets for roads and paths for purchase and third-party re-use/analysis.	Supply, Aggregate, Develop

- 5.17 The extent to which these services are actually free varies according to the value users, and data aggregators/exploiters, place on the digital exhaust data⁶ they create. In most cases, such free mobility information services actually involve a user trading their personal usage/location data in return.
- 5.18 As such many emerging personal mobility information, intelligence, and insight products and services are strategically building new user-based datasets that have not previously existed at such a granular level. For example – the possibility of interrogating train or bus ticket sales in silos has always existed, but knowing an individual’s end-to-end journey patterns in the way that an organisation like Google Maps and Citymapper now can (by retaining and examining user’s intended and completed journey GPS traces) presents a potential step change in understanding travel behaviour..
- 5.19 The current trend, driven by the success of social networking tools such as Facebook and Twitter, is for personal mobility information and insight services to be deeply integrated into tangible mobility products and services. The auto industry’s Connected Vehicle programmes; which are focused on integrating smart navigation and parking, crowd-sourced sentiment, related contextual ‘lifestyle’ information, and media platforms into the vehicle environment; provides an example of this. The data journey illustrated in Figure 5-3 depicts how crowd-sourced and aggregated parking space location, price and availability data through the Parkopedia/Just Park platform are being used in a range of contexts – including integration into existing forms of mobility.

For transport industry

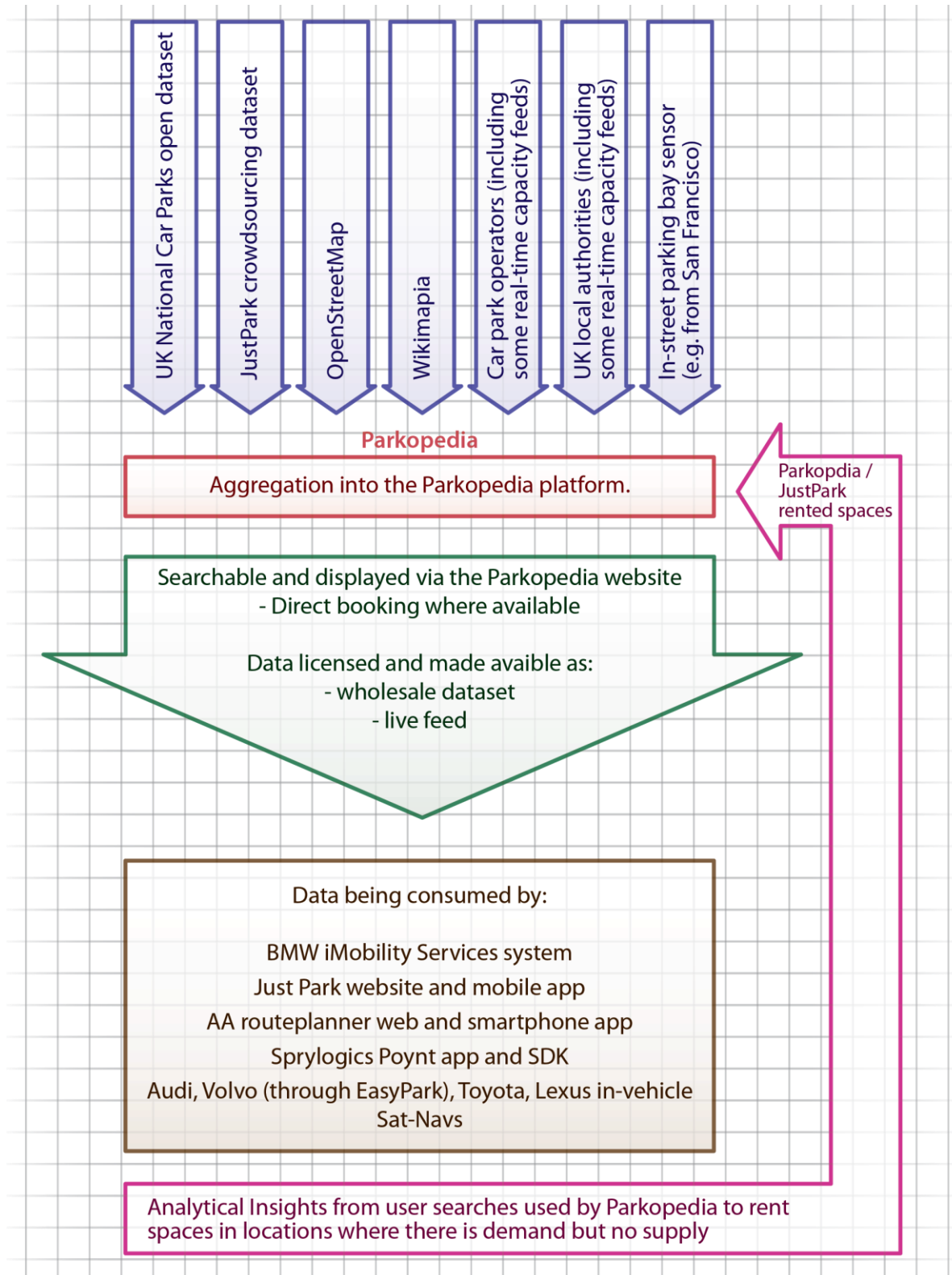
5.20 Data-driven information, intelligence and insight services for the transport industry increasingly involve making sense of the vast range of data sources that are created and stored; and aggregating this data in a way that permits strategic analysis to derive actionable insights. Historically, these insights have been confined to individual modes of travel, or simple analyses of multi-modal transport interactions, with individual/small numbers of databases being interrogated and researched. The technological trends driving the pace, volume, and accuracy of data collection and storage identified earlier in this report; and computational analysis techniques for handling and enquiring multiple datasets; are changing the scope of multi-modal analyses that are possible on behalf of the transport industry.

for Transport Industry

- Operational insight
- Strategic planning
- Network coordination
- Infrastructure management

- 5.21 The data-driven information, insight and intelligence services being delivered for the transport industry (see Table 5-4 for examples), cover the following service areas:
 - Data-as-a-Service, whereby raw datasets are aggregated and processed into formats or feeds that the transport industry needs in order to function. Examples include [Cloudmade](#) and [INRIX](#) providing Sat-Nav and connected vehicle data services to the automobile industry.
 - Converting raw transport system/operational data into both public information and operational outcomes, such as [Cloud Amber’s](#) maintenance, and aggregation, of UTMC and traffic count data from multiple sources to power public information tools and traffic control systems.

Figure 5-3: Data journey through Parkopedia's platform



5.21 Collecting and maintaining data, such as the roadworks dataset maintained by [Elgin](#) on behalf of English local authorities; or [Traveline's](#) work to maintain the national public transport dataset for regions outside of London.

- **Analytics and insights as consultancy services to the transport industry**, such as the Transport Data Management Platform that [ITO! World](#) has created and is using to provide connectivity and environmental analyses to UK transport operators and local authorities. This service area also includes the wide range of existing transport planning, modelling and research services undertaken by central government, local authorities, transport operators, and private sector/not for profit consultancy providers. The Pattern Based Strategy³⁵ predictive traffic model³⁶ developed by ATOS and deployed on a proof of concept basis is a similar example of analytics and insights in action. The work of Mike Barry and Brian Card to demonstrate the value of using big data analytics to visualise and analyse operations on the Boston MBTA subway system is also a good example of these types of analytical service³⁷.

5.22 Unlike private consumers, most transport industry organisations are used to paying for data in order to build models, perform strategic analyses, or provide information services to their customers. As such the services listed above, and the examples cited in Table 5-4, are almost always delivered on a paid-for basis; with the exception of revenue-sharing and joint commercial exploitation models that exist in some cases. Innovations in this field are primarily occurring in respect of how existing databases held and stored by the transport industry can be made publicly available, aggregated alongside newer crowd-sourced datasets, and exposed to computational analyses techniques to deliver new insights that result in greater operational efficiency and strategic expansion of transport system capacity and services.

5.23 As depicted in Figure 5-4, [Elgin's](#) approach is less conventional and has involved spending 10 years brokering and crowd-sourcing roadworks data from over 175 different streetworks systems used by English local authorities to build a de-facto industry standard set of tools and dataset. Local authorities now pay a low subscription fee to licence use of the Elgin's software tools, and then receive free access to the aggregated national roadworks dataset thereafter (which is also publicly available on Elgin's website). Elgin is exploiting this data by syndicating it on to a number of sources, with the expectation of future market growth as connected vehicle environments drive up demand for a consolidated data feed on roadworks and road network disruptions such as that it has created in partnership with English local authorities. Further value is being derived through the development of associated Traffic Management Order coordination and communications tools which function along similar principles.

³⁵ Gartner (2009) *Introducing Pattern Based Strategy*. Available online at: <https://www.gartner.com/doc/1117912>, last accessed on 17/07/14

³⁶ ATOS (2013) *Ascent White Paper Real Time Traffic Forecast*. Available online at: <http://atos.net/content/dam/global/ascent-whitepapers/ascent-whitepaper-real-time-traffic-forecast.pdf>. Last accessed, 20/07/14.

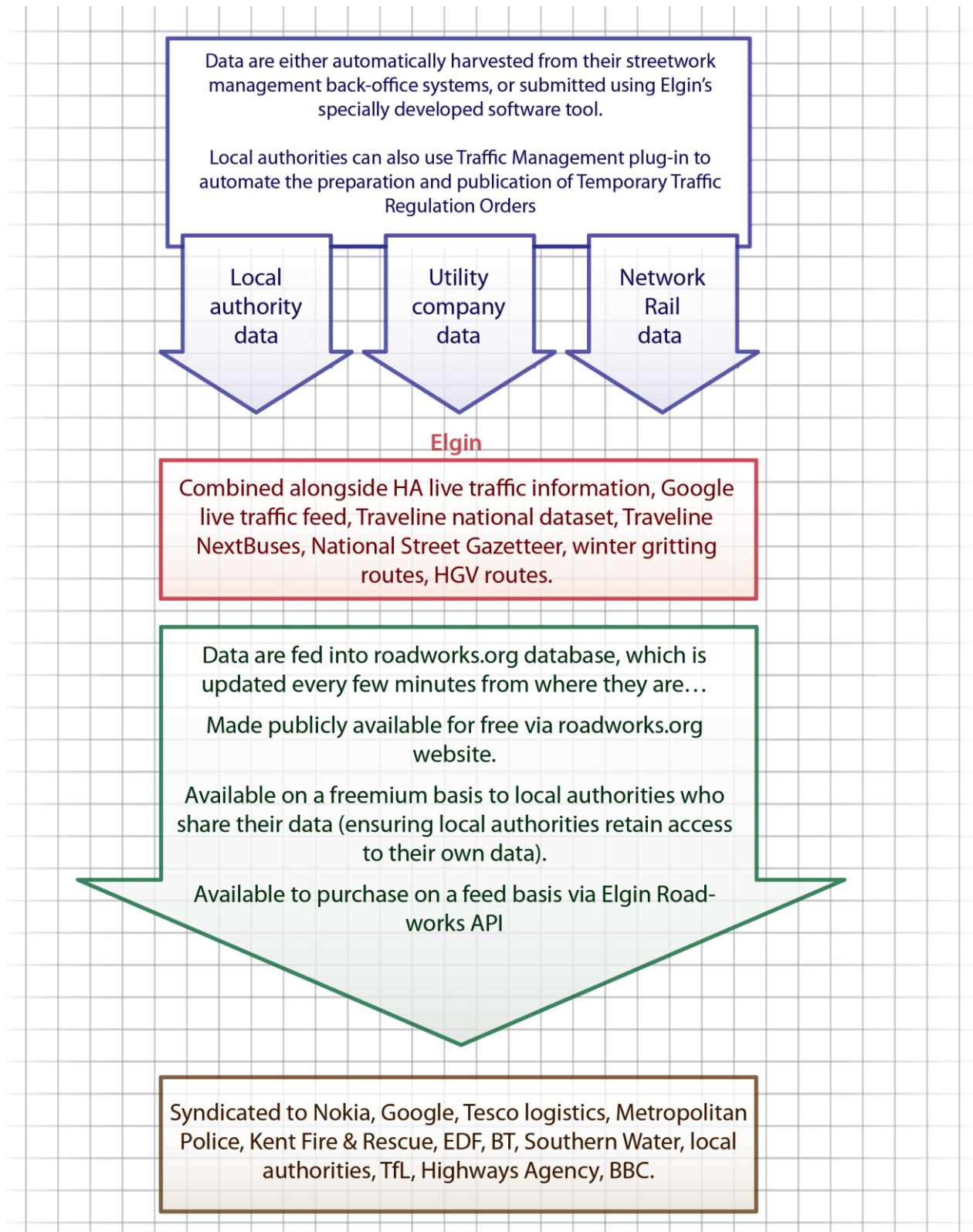
³⁷ Mike Barry and Brian Card (2014) *Visualizing MBTA Data*. Available online at: <http://mbtaviz.github.io/>, last accessed on 24/06/14.

Table 5-4: Example information, intelligence and insight services for the transport industry

Provider	Service description	Use of transport data	Service types
Cloud Amber Envia Mott MacDonald Siemens Traak systems	Urban traffic management control data storage and system analytics.	<p>Uses live feeds of data from SCOOT counting loops, ANPR cameras, Bluetooth networks, traffic signals, real-time bus vehicle location, and other sensors using common UTMC standards.</p> <p>Creates insight into the real-time, and predicted, performance of local road networks. Real-time public transport information systems and public-facing multi-modal travel information services delivered on clients' behalf.</p>	Aggregate, Develop, Enrich
Cloudmade	Content, map data, and user interface software for connected vehicles provided as OEM services to the automobile industry.	<p>Uses pre-cached & fetched data feeds from over 1,000 global sources: points of interest (Foursquare, Yelp, Google Places), road networks (HERE, TomTom, OpenStreetMap), traffic data (INRIX, HERE), sentiment (Facebook, Twitter), fuel prices (OPIS, Gas Buddy), weather (The Weather Channel, Weather Decision Technologies) , and connected vehicles.</p> <p>Creates Single-line/voice searchable database for connected vehicle navigation, destination discovery, mapping. Connected vehicle exploitation through car & driver analytics, and traffic data crowd-sourcing.</p>	Aggregate, Develop, Enrich
Elgin	Roadworks and traffic disruption data collection, management and communications services.	<p>Aggregates official roadworks data from 175 English and Welsh local authority streetworks systems and publishes them on an open web map.</p> <p>Uses Google Maps and Traffic data, HA National Traffic Information Service, National Street Gazetteer, winter gritting routes, HGV routes, TransportAPI data services.</p> <p>Creates roadwork and Temporary Traffic Regulation Order management and communications software tools for local authorities, JourneyMapper public journey planning tool as a service for local authorities.</p>	Aggregate, Enrich, Enable

Provider	Service description	Use of transport data	Service types
<u>INRIX</u>	Collecting GPS and mobile phone data for multi modal travel analysis and Sat-Nav services	<p>Aggregates traffic data from Sat-Nav users, app users, connected vehicle environments, fleet customers, traffic cameras, HA National Traffic Information Service, fuel price, weather, EV charge networks, GSM data.</p> <p>Creates real-time and predictive traffic data services and routing data for automobile industry, road network managers, members of the public, Sat-Nav manufacturers, fleet operators.</p>	Supply, Aggregate, Develop, Enrich
<u>ITO! World</u>	Collect, aggregate, manage, analyse complex public transport and spatial data, including real time data feeds	<p>Aggregates traveline national dataset, HA National Traffic Information Service, all available real-time UK public transport data feeds.</p> <p>Creates transport maps, accessibility planning tools, Transport Data Mgt. Platform, sustainable transport marketing information, data visualisation.</p> <p>Supplies cleaned & re-formatted UK public transport data feeds to Google, transport modellers, local transport authorities, navigation apps.</p>	Aggregate, Develop, Enable

Figure 5-4: Data journey through Elgin’s platform



For other industrial sectors

5.24 Other industries, that are users of transport systems and services, also exploit transport data to varying degrees in order to develop and enrich their products and services. The wide range of uses for this data beyond the transport sector makes it hard to accurately categorise them into service areas. Based on the small number of examples derived from this review, and set out in Table 5-5, it is possible to identify the following groupings:

for Other Industries

- Peripatetic services
- Retail
- Event organisation
- Emergency response

- **Collection and exploitation of data through activities that are not transport-related at their core.** [Trafficmaster](#), [Microlise](#) and [The Floow](#) all depict organisations that specialise in collecting and exploiting data that relate to transport services, but are ultimately being used by industries that are not necessarily specialised in transport (e.g. insurance / retail). The Floow’s use of telematics data to improve the quality of driver insurance and driver education services; and Microlise’s use of similar data to optimise retailer operations and peripatetic service delivery – for whom transport is a means to an end – are great examples of this. Their clients are concerned about how more detailed information on the movements of their staff, vehicle fleets, and goods, affect the sound operation and profitability of their businesses.
- **Use of open and available transport data to add value to non-transport services.** Whether it be aggregating and geo-locating the sentiments of local people to proposed planning and regeneration schemes ([CommonPlace](#)), or using open transport data to define travel times to/from specific places or opportunities at different times of day ([TravelTime](#) and [Mapumental](#)), these services all use transport data placed in a geo-spatial context to deliver greater meaning and value for their clients – who include planning authorities, property developers and emergency services. Services like [Screach.tv](#), which provide information feeds to redundant TV screens (commonly public transport information in pubs when there is no live sport to show) also fall into this category.

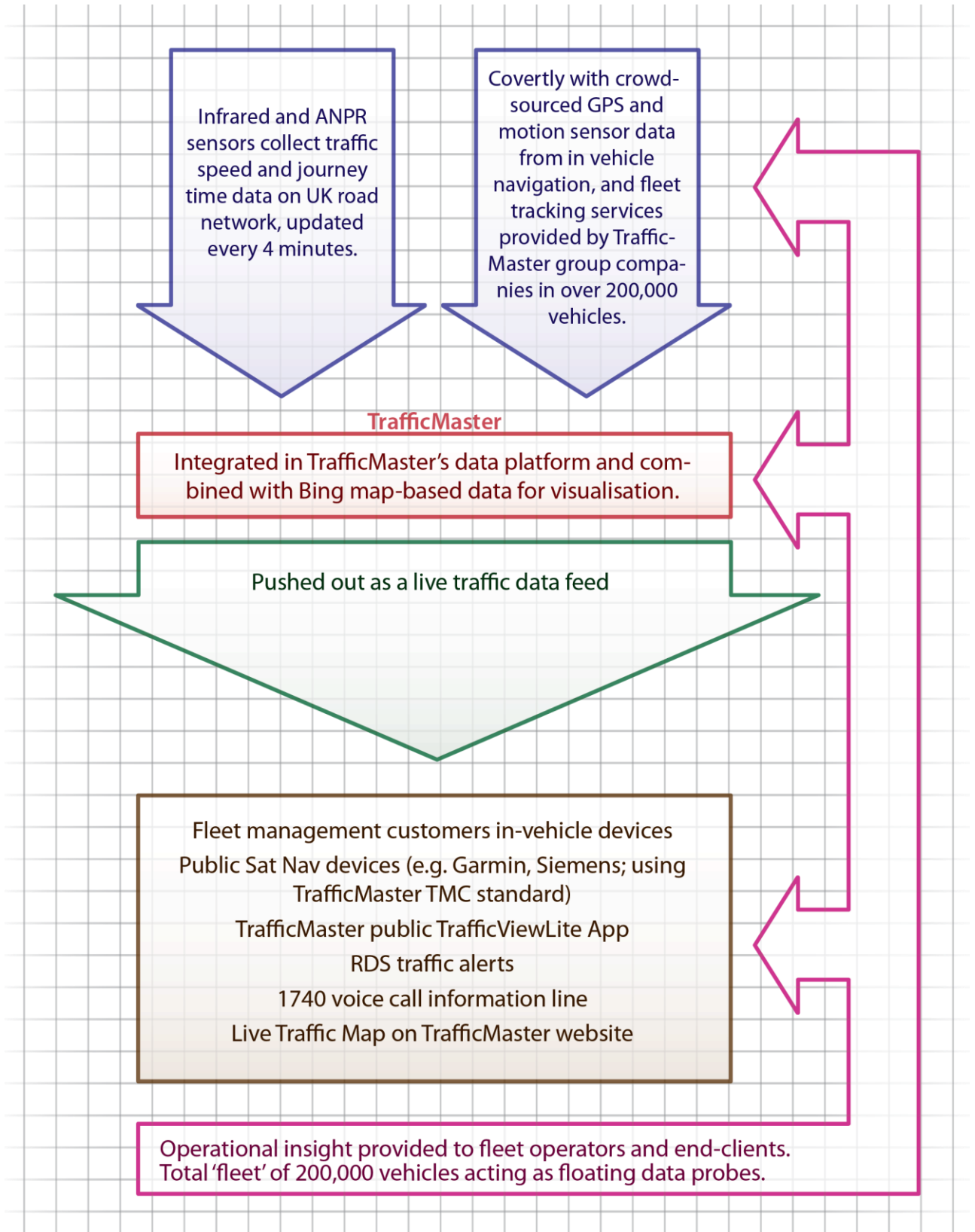
5.25 The example of [Trafficmaster](#), and its use of road traffic data depicted in Figure 5-5 is particularly interesting. It demonstrates how a UK firm has evolved from a conventional fixed-location sensor-derived data company (leveraging its network of radar hot-spot sensors and ANPR cameras), to one which aggregates data from a legacy sensor network (which is declining in size) alongside the floating vehicle ‘probes’ made up of its anonymised fleet users’ 200,000 vehicles. The result is a larger sample of traffic data and richer insights for its clients in respect of traffic congestion, route guidance, and operational efficiencies. Similar services are also provided to non-transport industries, such as real-estate and media companies by [INRIX](#); which additionally infers high-level traffic movement, and traveller origins and destinations, using data purchased from two of the major mobile communications service providers.

Table 5-5: Example information, intelligence and insight services for other industries

Provider	Service description	Use of transport data	Service types
<u>CommonPlace</u>	Aggregation and exploitation of people’s sentiment data	<p>Uses map data from OpenStreetMap styled using the Leaflet open source tool.</p> <p>Collects and aggregates sentiment data from bespoke web and smartphone interfaces and via common social media channels.</p> <p>Creates Analytics of aggregated sentiment to the planning of anything in the built environment (land-use, transport, shop locations etc).</p>	Aggregate, Develop, Enrich
<u>Microlise</u>	Vehicle tracking, telematics and proof of delivery services	<p>Uses map and point of interest data, client-defined locations</p> <p>Aggregates GPS and vehicle sensor data from bespoke sensors and customised communications devices (phone/tablet) installed in participating fleets.</p> <p>Provides remote vehicle tracking, driver alerts, 2-way messaging, tachometer integration, driver performance analytics, fleet/driver speeding and cornering reports, incident recording, safety camera feeds, trailer security, scheduling, route management, Sat Nav, KPI reports, delivery tracking, business intelligence.</p> <p>Delivers 5-10% fuel efficiency for clients, reduced speeding and safety incidents.</p>	Aggregate, Develop, Enrich
<u>The Flow</u>	App-based telematics services leveraging Android and iOS devices	<p>Uses map and point of interest data, client-defined locations.</p> <p>Aggregates GPS and telematics device data from vehicles with app deployed.</p> <p>Provides actuarial insight into individual and aggregated driver behaviours and risk factors to enable fairer insurance premium pricing. Telematics-driven insurance services for high risk (17-25 year old) drivers. Automated crash notification system. Tailored driver education programmes. Contextual accident causation analyses</p>	Aggregate, Develop, Enrich

Provider	Service description	Use of transport data	Service types
<p><u>Trafficmaster</u> (Danaher)</p>	<p>Fleet management , GPS tracking and Sat-Nav services across all industry sectors. Stolen vehicle tracking.</p>	<p>Uses map and point of interest data, client-defined locations, TrafficMaster IR and ANPR sensors on UK motorway and trunk road network</p> <p>Aggregates GPS and vehicle sensor data from Sat Nav and fleet users</p> <p>Provides in-vehicle navigation, fleet management and monitoring tools, freemium public information traffic and navigation apps, traffic data for SatNav device and vehicle manufacturers, RDS alerts.</p> <p>Creates large anonymised vehicle location/speed/journey time/behaviour dataset across all clients for use as real-time feeds and historic journey data.</p>	<p>Aggregate, Develop, Enrich, Supply</p>
<p><u>TravelTime</u> (iGeolise)</p> <p><u>Mapumental</u> (mySociety)</p>	<p>Creation of time-base travel maps to present journey information in time rather than distance for Great Britain.</p>	<p>Uses Traveline national dataset, Network Rail open data, HA National Traffic Information Service</p> <p>Creates travel time-based maps, search, and data insights for any organisation’s internal company systems (e.g. sales force management), website or operational management.</p>	<p>Aggregate, Develop, Enrich</p>

Figure 5-5: Data journey through TrafficMaster’s platform



Tangible mobility service delivery

5.26 As noted in Chapter 2 of this report, a key market is for transport operators to increasingly be aware of the scope to exploit usage and operational datasets they have always collected and stored. At the same time they are increasingly moving into the tangible provision of mobility services, looking to exploit the crowd-sourced and user-based datasets they are building. The examples in this section demonstrate these trends in respect of both the tangible services that exist to move people and things.

For People

5.27 The examples of data-driven mobility services for people presented in Table 5-6 are understood to represent only a modest sample of the full range that exist. We have deliberately not included reference to conventional modes of transport (as listed in the graphic to the right), but note there is scope for these services to be operated more intelligently and adaptively through smarter use of operational and passenger-transaction data (discussed in Chapter 6). The services listed in Table 5-6 can be grouped as follows:

for People

- Bike hire
- Private car / car club
- Taxi / ride-on-demand
- Bus / light rail / metro
- Rail / coach
- Air / ferry / intl. rail

- **Ancillary services that make it easier and cheaper to use existing transport systems.** These include smart payment systems such as those provided by [ACT](#), [corethree](#) and [Droplet](#) (not included in Table 5-6, but active in this space), [Assist-Mi](#) which is aimed at travellers who experience limited mobility, and [Ingenie's](#) pay-as-you-drive insurance products which refine young driver insurance costs based on telematics data gathered to determine their behaviours. These services go beyond information and insight, by enhancing or facilitating a consumer's use of existing forms of mobility.
- **On-demand mobility using socially shared/community assets.** These include well-established virtual platforms for car-sharing such as [Liftshare](#) and [Faxi](#); as well as by-the-minute vehicle hire services such as BMW's [DriveNow](#), [CityCarClub](#), and [ZipCar](#).
- **On-demand door-to-door mobility services using operator-provided assets.** In their data-driven context, these services represent an emerging group that are rapidly, and disruptively, changing the way that urban mobility is provided and procured. Examples include existing personal mobility services such as [Lyft](#) and [Uber](#), which provide a totally integrated mobility experience (e.g. mobility request and purchase) for consumers, and those which are in the early stages of coming to market to provide similar demand-responsive collective transport systems (e.g. [Simply Connect](#) and [Bridj](#)).

5.28 Almost all of these user-focused services; and in particular the on-demand mobility services illustrated in Figure 5-6; only rely on map-based data that relates to the locations of public transport and road networks, points of interest, and places (e.g. addresses). The main reason is that these services have been designed in a closed way to resolve discrete mobility challenges, and are currently concentrating on scaling-up to win market share. As such, integration with other mobility services (e.g. Lyft/Uber booked as part of a seamless journey involving coach/rail/air service interchange) is not yet happening, but could conceivably do so in the future.

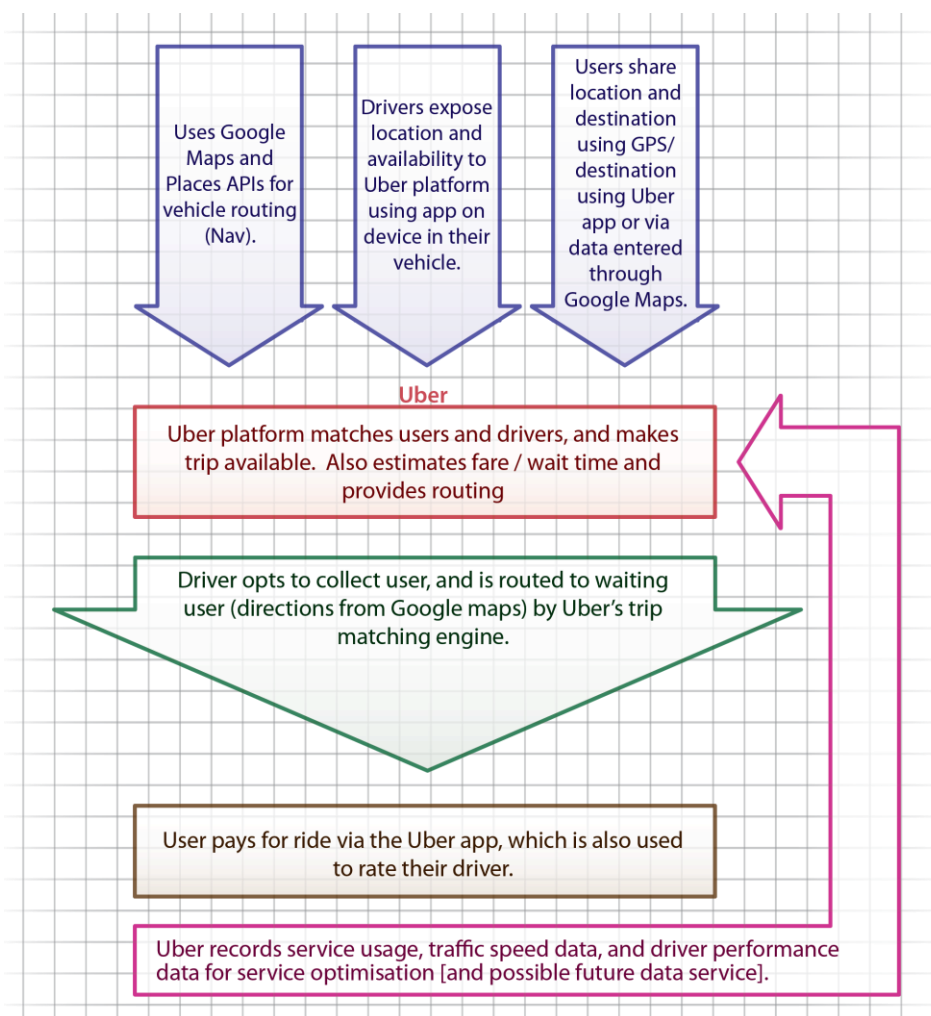
Table 5-6: Examples of tangible mobility services for people

Provider	Service description	Use of transport data	Data Service types
ACT (Applied Card Technologies)	Manage and provide cashless ticketing and payment systems adhering to UK (ITSO) smartcard standards.	Creates volumes of transactional data relating to customer travel (boarding and alighting) and ticket purchasing behaviours, as well as related use of tourist services.	Aggregate, Develop, Enrich
corethree	Mobile ticketing, real-time information and multi-modal services direct to passengers smartphones	Uses live public transport service data feeds and schedule data. Creates volumes of transactional data relating to customer travel (boarding and alighting), ticket purchasing behavioural analytics to inform operator’s promotional ticketing strategies, advertising opportunity.	Aggregate, Develop, Enrich
DriveNow CityCarClub ZipCar	By the minute vehicle hire delivered via smartphone / smartcard services.	Uses Google’s Maps and Places APIs to power map-based search for car-share locations and capacities. Creates data on customer rental origin-destinations and GPS data from vehicle trips crowd-sourced into traffic speed dataset.	Aggregate, Develop, Enrich
Simply Connect and Bridj (Boston)	Demand responsive public transport via smartphone booking and trip-demand aggregation tools.	Uses map data and points of interest related to service area. Collects and Aggregates user origins and stated destinations using GPS. Delivers demand responsive public transport in limited locations / along key corridors (both organisations are proof of concept start-ups).	Aggregate, Develop
Faxi , Liftshare	Social journey sharing tools	Uses Google Maps and Places APIs. Aggregates user-contributed journey origins and destination locations to identify opportunities for regular journey-sharing. Delivers journey sharing service reducing travel costs for users.	Aggregate, Develop

Provider	Service description	Use of transport data	Data Service types
<u>Assist-Mi</u>	Personal assistance for disabled people ordered via smartphone	<p>Combines map and point of interest data, and transport stop and station locations, with user’s personal location data shared via smartphone.</p> <p>Creates service requests for personal assistance from disabled users, means of communication between service providers and disabled clients.</p>	Aggregate, Develop, Enrich
<u>Ingenie</u>	Pay-as-you-drive car insurance product for younger drivers.	<p>Uses GPS and motion sensor data from ‘black box’ device fitted to vehicle.</p> <p>Creates dataset on driver behaviour and location patterns of young drivers</p> <p>Delivers Driving score via smartphone app every 10 days to influence driver behaviour. Insurance price reviewed every 3 months.</p>	Aggregate, Develop, Enrich
<u>Lyft</u> <u>Uber</u>	Crowd sourced ride-sharing via smartphone	<p>Uses map and point of interest datasets, user/driver GPS location data.</p> <p>Delivers Virtual car service and integrated payment via smartphone.</p>	Develop, Enrich

- 5.29 These data-driven ‘ride on demand’ mobility service operations also create their own datasets based on personal location data harvested through customer interactions with their vehicle fleets. Although yet to have been exploited widely, there is scope for Lyft/Uber’s underlying operational datasets to come to market as a data service for third parties. Such a move would allow other service providers (e.g. restaurants, hotels, rail operators, etc) to make calls for ‘ride on demand’ services on behalf of their customers. It would also expose a new dataset for analysis and strategic transport planning / socio-economic analyses.
- 5.30 Uber has been a particularly high profile disruptor in the UK, eliciting protests from taxi operators in London who contest the service should be regulated under the same legislation that grants the taxi trade the right to ply for trade.³⁸

Figure 5-6: Data journey through Uber’s platform



³⁸ Dan Gillmor in the Guardian (2014) *The Uber revolution is just beginning, and no cabbie protest can stop it*. Available online at: <http://www.theguardian.com/commentisfree/2014/jun/13/uber-cabbie-protest-sharing-economy>, last accessed on 08/07/14.

For Things

5.31 The impact of data-driven decision making and analytical technologies on the provision of mobility services for ‘Things’ (i.e. bulk freight, consumer goods) is the area least explored by this study. Our understanding is that scope exists for a considerable amount of data exploitation to happen in this field, and it is being driven by the main freight operators. [DHL’s Big Data in Logistics](#)³⁹ publication confirms this, but suggests that while Big Data is predicted to become a disruptive trend in the logistics industry the application of Big Data analytics is not immediately obvious in the sector. We believe this is due to fierce commercial pressures that exist, and competing organisation’s desires to exploit (rather than share) the intellectual property they develop.

- for Things**

 - Cycle couriers
 - Local delivery vans
 - HGV
 - Rail freight
 - Air freight
 - Sea container

- 5.32** The services defined in Table 5-7 focus on:
- The optimisation of retail deliveries ([Activ8 Virtual Parking Solutions](#))
 - Last mile delivery and collection of parcels ([Collect Plus](#), [MYWAYS](#), [SmartTruck](#))
 - Long-distance rail-freight service optimisation ([FreightArranger](#)).

Table 5-7: Tangible mobility services for things

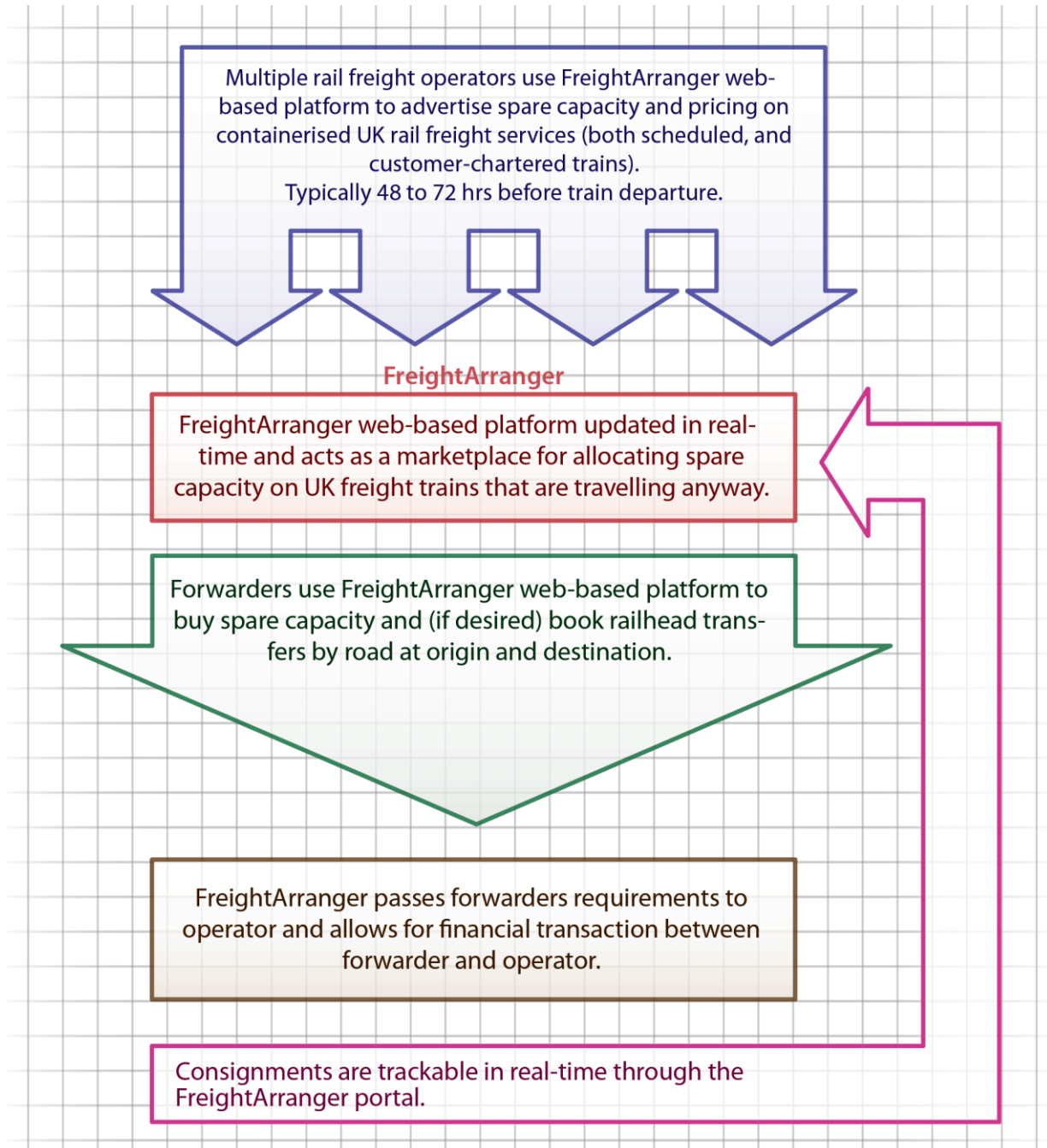
Provider	Service description	Use of transport data	Service types
Activ8 Virtual Parking Solutions	Virtual kerb-space booking for loading and unloading at a particular time and place.	<p>Uses map data of road alignments and markings.</p> <p>Aggregates Geo-fenced virtual parking bay locations and user-requests to book.</p> <p>Provides managed parking bay services</p>	Develop, Enricher
Collect Plus	Personalised parcel sending/goods collection services.	<p>Combines map data, locations of collect/drop locations, purchase data from users, smart sensor data from collection cabinets or retail outlets handling goods.</p> <p>Provides personalised goods collection and delivery services.</p>	Develop, Enricher

³⁹ DHL Customer Solutions and Innovation (2013) *Big Data in Logistics – A DHL perspective on how to move beyond the hype*. Available online at: http://www.dhl.com/content/dam/Local/Images/go/aboutus/SpecialInterest/innovation/SI_Studie_BIG_DATA_FINALE-ONLINE.pdf, last accessed on 23/07/14.

Provider	Service description	Use of transport data	Service types
FreightArranger	Containerised freight transport as a service	<p>Matches data on user requirements with real-time rail and connecting road freight availability capacity data, optimised for booking 48-72 hrs before train departure.</p> <p>Provides online brokerage platform for buying, selling and tracking intermodal rail freight capacity and connecting services.</p>	Develop, Enricher
MYWAYS (DHL)	Crowd-based parcel delivery service	<p>Uses data from user app interaction regarding parcels to be delivered directly to them. Map-based data on local area and address/location for delivery.</p> <p>Provides Crowd-based delivery for B2C parcels. Flexible delivery in time and location. Using existing movement of city residents.</p>	Develop, Enrich
SmartTruck (DHL / Quintiq)	Dynamically adapting route guidance and delivery sequencing based on local traffic conditions	<p>Uses live data feed on current (crowd sourced from Berlin taxi fleets) traffic disruption and congestion, map-based data, precise delivery locations.</p> <p>Provides dynamic re-routing and delivery order sequencing to optimise driver’s time and reduce delivery route mileage.</p>	Develop, Enrich

5.33 Like the example of FreightArranger depicted in Figure 5-7, several of these services commonly seek to improve the efficiency of freight services by acting as a brokerage for buyers and sellers of services. Where FreightArranger simplifies the process of connecting the forwarders and consigners of containerised railfreight (and supporting road-to/from-rail services); Activ8 connects people wishing to use and rent kerb space to perform deliveries, and MYWAYS connects people who can collect parcels with those who need them collecting and bringing to them. Only FreightArranger and CollectPlus have moved beyond a pilot or startup stage at the time of preparing this report.

Figure 5-7: Data journey through the FreightArranger platform



6 Where is the future value in transport data exploitation?

6.1 Our review explored transport industry/related data-service providers’ views on how transport data can be catalysed and exploited to provide foundations for intelligent mobility. This chapter collates the key outputs from workshop discussion sessions to identify the motivations for exploiting transport data to enable intelligent mobility; and emerging business models for, and barriers to, exploiting complex and multi-source transport data.

Intelligent mobility motivations for exploiting transport data

6.2 Table 6-1 sets out a range of possible motivations for seeking to exploit complex and multi-source datasets related to transport. These emerged through workshop discussions with transport and data specialists, and have been mapped against the 5 intelligent mobility themes referenced in Chapter 1 of this report. The drivers and opportunities set out below are not mutually exclusive to each intelligent mobility theme, but have been presented against the topics against which they most closely align.

Table 6-1: Motivations for exploiting transport data

Intelligent mobility theme	Drivers and opportunities for transport data exploitation
Autonomous systems	<ul style="list-style-type: none"> Optimised performance of existing transport and logistics services Reduced operating costs and staffing requirements Better-use of existing transport network capacity Reductions in fuel costs and transport emissions
End-to-end journeys	<ul style="list-style-type: none"> Greater convenience for transport users Inclusive access for people with limited mobility Mobility as a service procured on-demand Reduce time and costs associated with moving people and things
Information exploitation and customer experience	<ul style="list-style-type: none"> Easier to use transport networks through enhanced information Tailored, contextual assistance for travellers with particular needs Reduce time and costs associated with moving people and things Add value by leading consumers to relevant goods and services
Resilience	<ul style="list-style-type: none"> Faster response to emergencies and incidents Better-informed strategic plans for winter readiness Dynamic switching between transport networks Reduced costs associated with service delays and cancellations

Intelligent mobility theme	Drivers and opportunities for transport data exploitation
Smart infrastructure	Adaptive capacity to accommodate primary movement flows New sources of digital exhaust data from connected infrastructure Reduce/delay need for additional road/rail/air infrastructure Optimise maintenance and repair activities based on sensor data

6.3 The trend towards the greater provision and use of open data will force businesses to conceive new commercial models, derive higher quality insights, develop new ways of engaging with stakeholders and ultimately deliver economic and social benefits⁴⁰. We believe that most of the drivers and opportunities identified in Table 6-1 come with value attached, either in the form of:

- Time and money savings for transport operators, service users, and the economy.
- Passenger/customer revenues arising from the use of new forms of mobility, and increased sales of related services.
- Payment for strategic planning and operational insights, and advice, derived using transport-related data.
- Payment for data itself, which will continue to flow between mobility service providers, organisations, and individuals as they interact.
- Revenues from partnerships/sales of tools and applications to major technology companies that are growing by acquisition.
- There is every reason to believe that the same will apply to the public sector, with local authorities moving from a supplier of services to an enabler of services. There have been a number of estimates from across Europe covering the impact of open data on Gross Value Added (GVA). These are summarised below, and have been normalised to provide GVA uplift per capita:

Table 6-2: Estimates of GVA per capita from Open Data

Source	GVA per Capita
Gartner, Innovative use of public data, report to the National IT and Telecom Agency, and the Danish Agency for Science, Technology and Innovation (Danish), 2009	£54
Analysis by CS Transform Limited, based on modelling by Pollock, R. (2011), “Welfare gains from opening up Public Sector Information in the UK” ⁴¹	£122

⁴⁰ Deloitte (2012) *Open Data: driving growth, ingenuity and innovation*. Available online at: http://www.deloitte.com/view/en_GB/uk/market-insights/deloitte-analytics/deod882044ad7310VgnVCM3000001c56f00aRCRD.htm, last accessed on 24/06/2014.

⁴¹ Pollock (2012) *Welfare gains from opening up public sector information in the UK*. Available online at: http://rufuspollock.org/papers/psi_openness_gains.pdf, last accessed on 25/07/14.

Source	GVA per Capita
Open data measures in the Autumn Statement 2011, Cabinet Office as quoted in <i>Open Data - driving growth, ingenuity and innovation</i> , Deloitte, 2012.	£255
Data equity - Unlocking the value of big data, Centre for Economics and Business Research Ltd, April 2012 ⁴²	£650

6.5 This gives a very wide range; from £54 uplift per capita to £650, which suggests Open Data could be worth anywhere between £3.24bn and £39bn to the UK economy in Gross Value Added terms. Open transport data would represent a subset of this value.

Costing these values in relation to each of the drivers and opportunities was outside the scope of this study, but could be useful in order to identify short, medium, and long-term priorities for transport-related data exploitation.

A key challenge associated with valuing these opportunities are the flows of venture capital that currently skew the marketplace for transport data services and innovative tangible mobility services. In several cases significant venture capital ‘bets’ are being placed on future data markets, forecast values for user-derived datasets, and whole mobility services (e.g. ride-on-demand) at an early stage of their anticipated development and emergence.

Business models for exploiting transport data

6.6 During the second project workshop we focused on the question ‘Where does the value lie?’ for transport data service providers in future intelligent mobility contexts. Ten different business models for exploiting transport-related data emerged through this dialogue with transport operators, practitioners, and data specialists. These were subsequently validated through desk-based research and analyses^{43,44} and are documented in Table 6-3 below. They are anticipated to provide the high level data building blocks, and key enabling services, on which future intelligent mobility services will exist.

⁴² CEBR (2012) *Data equity - Unlocking the value of big data*. Available online at: <http://www.sas.com/offices/europe/uk/downloads/data-equity-cebr.pdf>, last accessed on 28/07/14.

⁴³ The motivations set out in Table 6-2, and the business models presented in Table 6-3 are largely an extension of the five techniques for leveraging big data to create value identified in:

McKinsey Global Institute (2011) *Big data: The next frontier for innovation, competition, and productivity*. Available online at: http://www.mckinsey.com/~/media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx, last accessed on 24/07/14.

⁴⁴ The business models defined in Table 6-3 were also tested alongside an empirically-derived taxonomy of data-driven business models used by start-up firms, identified in:

Hartmann, Zaki, Feldmann, and Neely - University of Cambridge Service Alliance (2014) *Big Data for Big Business? A Taxonomy of Data-driven Business Models used by Start-up Firms*. Available online at: http://www.cambridgeservicealliance.org/uploads/downloadfiles/2014_March_Data%20Driven%20Business%20Models.pdf, last accessed on 28/07/14.

Table 6-3: Business models for exploiting transport data

Service type	Business model descriptions	Key markets
Core transport data services	Create and maintain data market infrastructure where transport-related data from the public and private sector can be easily discovered bought, sold, and exchanged as wholesale datasets / as a commodity.	Gov to Biz Biz to Gov Biz to Biz
	Aggregate, clean, collate, and standardise multiple feeds of both open and available data; and re-sell them using common online data catalogues/marketplaces.	
	Create, manage, and curate EU/globally standardised data formats and conventions.	
Information, insight and intelligence products & services	Create standalone applications and tools that enable people and transport operators to make optimised, and more integrated use of existing transport networks.	Biz to Con Biz to Gov Biz to Biz
	Create white-label applications and tools for existing transport service providers to give-away to customers/re-sell/integrate into existing products.	Biz to Biz Biz to Gov
	Create and share/sell new sets of digital exhaust data derived from free/freemium personal mobility applications and tools.	Gov to Con Biz to Con Biz to Biz
	Create applications and tools, and provide analytical support/consultancy services, that enable non-transport businesses (e.g. estate agents, retailers, restaurants, builders) to reduce costs, attract customers, and optimise flows of goods and people to their services	Biz to Biz
	Create applications and tools and provide analytical support/consultancy services that enable transport businesses (e.g. service operators, network managers, planners), to reduce costs, attract customers, optimise capacity relative to demand, and strategically plan operations predictively and in real-time.	Biz to Biz Biz to Gov
Tangible mobility service provision	Develop and provide entirely new/re-imagined data-driven forms of mobility for people and things (e.g. ride-on-demand services, crowd-based parcel delivery, autonomous vehicles) to make better use of existing transport networks, and improve consumer choice and experience.	Biz to Con Biz to Biz
	Develop tools, applications and data-driven technologies that can be bundled with different forms of mobility (e.g. car/bus/train/plane), search engines, social media platforms, wearable technology, and personal comms devices to enhance integration across mobility services in real-time.	Biz to Biz

6.7 Our meta-analysis suggests that, for UK businesses, the most exploitable future value for intelligent mobility services in transport data terms will most likely reside in the components of value chains that major international tech companies are currently overlooking or failing to exploit. These are anticipated to be:

- Opening-up, promoting and releasing unexploited datasets currently held by the public sector and its contracted service providers (e.g. UTMC traffic flow data, bus fares) as open data.
- Autonomous sensor-based monitoring and data capture for transport networks and services via the growing Internet of Things.
- Aggregating and collating multiple transport data feeds so they can be readily combined and analysed for patterns, deployed in models, and up-scaled from representative population samples to inform strategic transport planning and analyses.
- Statistical and computational analyses/modelling of transport datasets, alongside wider consumer and socio-economic datasets, to identify and deliver market intelligence - actionable insights for and strategies for mobility service improvements and product innovation/development.
- Predictive and real-time analytics, and automated control systems, that optimise capacity in relation to demand across all transport networks.
- Products and information services that are compelling enough to be ‘baked-in’ to the next generation of intelligent mobility services (e.g. connected/autonomous vehicles).
- Developing ‘intermediate technology’ transport applications, tools and open data standards relevant to emerging economies’ transport systems.
- Establishing ethically-focused tools that fairly and transparently trade user’s digital exhaust data in return for contextually relevant information, insight, and/or incentives.
- Coordination and exploitation of the OpenStreetMap, including geo-coding improvements.

Identified challenges and obstacles to exploiting transport data

6.8 Building on the discussion at the end of Chapter 3, where we identified the critical data gaps that exist for intelligent mobility, our workshop discussions also identified a related group of barriers that may prevent the emergence of intelligent mobility. These are documented in

Table 6-4, and are framed around five themes:

- | | |
|--------------------------------|--------------------------|
| • Data availability and demand | • Data relevance |
| • Usability of data | • Structural impediments |
| • Market forces | |

Table 6-4: Possible transport data exploitation barriers for intelligent mobility

Theme	Identified transport data barriers
<p>Data availability and demand</p>	<p>Cost of data collection/purchase.</p> <p>Ethical/privacy issues relating to collection, storage and use of sensitive personal (location and sentiment) data.</p> <p>Incomplete or partial datasets - require knowledge of what's missing.</p> <p>Data not being open / lack of willingness to share (e.g. data islands such as Google/Bing, connected vehicle data platforms for each car manufacturing group).</p> <p>Data not available, simply because no one has asked for it yet.</p>
<p>Data usability</p>	<p>Incompatible data, resulting from lack of adherence to/existence of standardised formats and conventions, and proprietary formats.</p> <p>Data not machine-readable / lacking in metadata, making it hard to discover and access.</p> <p>Poorly structured and curated datasets deter time investment.</p>
<p>Data relevance</p>	<p>Data not real-time, or sensibly archived from real-time sources, to allow for detailed computational pattern and statistical analyses.</p> <p>Scale and feasibility of data samples outside of major conurbations (e.g. will X work in Coventry, or Bath?).</p> <p>Lack of precision over where and when data were initially collected (e.g. time intervals, and geospatial granularity).</p>
<p>Structural impediments</p>	<p>Lack of technical skills to handle and process big-data sets (Skills shortage).</p> <p>Relationships between organisations (e.g. public/private sector) may prevent data sharing.</p> <p>Internet bandwidth and mobile data coverage outside major cities.</p> <p>Government willingness/action on opening-up datasets.</p> <p>Personal/societal attitudes to data privacy and trust + nervousness of data owners to exploit anonymised personal data they hold.</p> <p>Data storage / archiving / recording – who is responsible?</p>
<p>Market forces</p>	<p>Competing/vested commercial interests disrupt open data sharing and standardisation - e.g. multiple geospatial data ‘stacks’ for map-based data.</p> <p>Rapid pace of technological change – may deter investment due to high costs involved.</p> <p>Lack of incentives to ‘add value to data’ – expectation that developers will respond to available data could be misguided.</p>

High 'expectation of free' for data services and consumer information based on major technology company's ability to cross-subsidise these using other revenue streams.

Most of these challenges are considered solvable; and are likely to be dealt with organically as the Internet of Things develops, and service provider's demands for existing public datasets grows. But those which relate to market forces are less easy to address, and instead may come to represent the unavoidable features of the transport data marketplace around which intelligent mobility services will need to work.

- 6.9 On the topic of 'expectation of free', and making raw data feeds available as high quality public services, a finding from the [Stride](#) project was that data-as-service models are valuable to software developers and innovators in the field of transport. As such they are likely to become key enablers of tomorrow's intelligent mobility services. While Stride, and [i-Move](#), demonstrated that personal mobility and transport network management applications *can* be developed using freely available and easily discoverable data, there are a number of easier ways for app developers to generate revenue. This suggests there needs to be a reason beyond simply app development for local authorities to go to the trouble of sourcing and opening-up datasets such as traffic flow from their UTMC systems.
- 6.10 While there remains an expectation that free, open data will continue to be released by the public sector (and in some cases provided back to the public sector free of charge by private sector operators - e.g. Elgin's roadworks, or ATOC's DARWIN datasets), some developers are demonstrating a willingness to pay for data feeds that are professionally aggregated, cleaned, and re-formatted for easier machine-readability. A good example of this is the vehicle manufacturing industry and its long-term investments in hyper-local traffic, roadwork/disruption, point of interest, parking and other dynamic routing datasets/feeds. Here the market for buying private or covertly assembled (e.g. from fleet vehicles using Sat Nav and job-scheduling/routing tools) datasets has been established for around 10-15 years. This is longer than that for public transport datasets – where there is a greater reliance on the data remaining free to re-use and openly available, because that is the basis on which they were originally released.

In chapter 9 we reflect on the opportunities and challenges identified above in order to make recommendations where intervention could help decisively move the market for transport data and, in-turn, accelerate the emergence of intelligent mobility.

7 What data-related capabilities are needed?

7.1 Much of the dialogue between transport and data specialists during the workshop discussions held through the course of this project related to individual and organisational capabilities and skills. This chapter sets out our findings on these topics.

Key data-related capabilities for intelligent mobility

7.2 Techniques and technologies developed and adapted to aggregate, manipulate, analyse, and visualize complex and multi-source transport data draw on several fields including; statistics, computer science, applied mathematics, and economics. McKinsey’s 2011 report on Big Data⁴⁷ describes 26 techniques and 24 technologies which, while not exhaustive, are commonly applied to big data sets. These are re-presented, and slightly adapted, in Table 7-1, below. Full definitions can be found on pages 27-31 of that report.

Table 7-1: Big data techniques and technologies relevant to intelligent mobility

Techniques	Technologies
A/B testing	Big Table distributed databases
Association rule learning	Business Intelligent application software
Classification	Cassandra open source database management
Cluster analysis	Cloud computing
Crowdsourcing	Data mart used to provide data to users
Data fusion and data integration	Data warehouse optimised for reporting
Data mining	Distributed systems for parallel computing
Ensemble learning	Data extract, transform and load (ETL) software
Genetic algorithms	Hadoop framework for very large data processing
Machine learning	HBase non-relational distributed database
NLP: Natural Language Processing	MapReduce data processing tool
Neural Networks	Mashups combine data/functions from >1 source
Optimisation of service from insight	Metadata describe content & context of data files

Pattern recognition and analysis	Non-relational databases (data not in tables)
Predictive modelling	R. Open source statistical computing tools
Regression analysis	Relational databases (data stored in tables)
Sentiment analysis	Semi-structured data containing tags & markers
Signal processing	SQL relational database management language
Spatial analysis	Stream processing of large real-time event datasets
Statistical analyses	Structured data that reside in fixed fields
Techniques	Technologies
Supervised learning	Unstructured data not stored in fixed fields (video)
Simulation and modelling	Visualisation tools for data analysis outputs imagery
Time series analysis	
Unsupervised learning	

- 7.3 These techniques and technologies will not be homogenously applicable to all large and multi-source datasets that relate to intelligent mobility. Instead they will likely be used to varying degrees in the specific context of the types and sizes of datasets being combined, the proportion of sample population(s) they represent, and the rate at which the datasets in question were originally collected.
- 7.4 We believe the techniques and technologies documented in Table 7-1 can be simplified into a broader set of capabilities, and example activities relevant to intelligent mobility. These are presented in Table 7-2.

Table 7-2: Data-related capabilities and example activities for intelligent mobility

Capability grouping	Common activities
Raw data creation, collection and curation	Exposing feeds of data currently collected/stored in silos so they become linked open data (e.g. UTMC data). Collaborative definition and maintenance of national / regional / global standards for data. Cataloguing transport-related datasets to improve discoverability. Sensor design and deployment to create new datasets. Archiving and storing data collected in real-time for future analyses.
Dataset handling and manipulation	Data mining, cleaning, and aggregation. API creation and consumption. Design and maintain standardised/bespoke data services. Cloud-hosting and stream processing.
Computational and statistical analyses	Algorithm design and application to multi-source datasets. Geospatial analyses using relational and non-relational databases. Predictive modelling, sampling and inferential analyses (statistical probabilities) using real-time and historic datasets.
Capability grouping	Common activities
Human intelligence and use of data insights	Transport service scheduling based on deeper understanding of partially connected networks and services (road/rail/air/sea). Data-driven optimisation of transport services Digital economy business model identification and exploitation. Management-level decision-making utilising ‘big data’ insights.
Software and technology development	Applying outputs from human intelligence + data analysis insight to design new mobility services which directly meet user needs. Innovation arising from new datasets and intelligence they create. Designing automated systems that exploit computational analyses and human intelligence from data insights. Creating new ICT technologies to keep-up with the rate of data sample from transport systems and personal movement. Sales and marketing to raise awareness of new intelligent mobility products and services.

Requirements for building capability and capacity for intelligent mobility

- 7.5 Three key sets of capability development and capacity building requirements were identified through the course of our meta-analysis, and workshop discussion sessions. They were:
- Skilled technical talent capable of handling and analysing very large datasets compiled from multiple sources.
 - Organisational capability that ensures key business leaders and decision-makers understand these analytical processes in outline. They need to be able to use actionable insights derived from new data analysis processes to inform strategic company decision-making, and exploit associated creative business models.
 - Technological investment to ensure access to requisite data storage capacity and computational processing power, for example through the use of cloud-hosted servers.

7.6 The remainder of this section explores these three requirements in more detail.

Finding people with the right combinations of skills

7.7 All the workshop discussions held during this review identified a lack of appropriately skilled people as a major barrier to the adoption of data-driven mobility services over the next 10 years. The view of data specialists we engaged with was that the skills, as identified in Table 7-1 and Table 7-2, do exist in the marketplace, but are in short supply; and that few (if any) people will demonstrate all of the skills defined in these tables. As such the challenge for transport service providers and systems coordinators will be to recruit and manage teams of people that comprise these skills. The widespread expectation is that demand for these skills will outstrip the available talent pool in the next 5-10 years.

7.8 These assertions are backed-up by market reviews conducted by McKinsey Global Institute, and the CEBR, who have published separate reports on the marketplace for 'Big Data' and 'Open Data' analytical jobs:

- The CEBR's assessment of Open Data jobs created concludes that the Professional Services, Manufacturing and Retail sectors are expected to gain the most from open data driven business creation. These sectors have higher levels of SME concentration when compared to Banking, Energy and Telecoms sectors. The CEBR forecast 58,000 new UK jobs to be created between 2012 and 2017 as a result of reduced barriers to entry and increased demand for data analytics skills.⁴²
- McKinsey's report stated there will be a shortage of talent necessary for organisations to take advantage of big data. By 2018, the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions.⁴⁷

7.9 The transport sector will compete with other major industry and technology sectors for people with these skills, and specifically seek to develop individuals who are able/motivated to bring them to bear on the optimisation of transport networks and mobility services. We predict that:

- Based on the latest ONS industrial sector employment figures⁴⁵, up to 5% of the ‘Data Scientist’ shortfall numbers quoted above are related to transport and mobility services.
- By 2017, as many as 3,000 data specialists will therefore be needed by the UK transport industry to achieve data-driven efficiencies and optimisation.
- Additional skilled analytical jobs of this nature are also expected to be created through the emergence of new forms of intelligent mobility over the next 10 years.

7.10 One of the largest challenges to filling these skill-gaps is that formally taught University courses that will accredit the next wave of Data Scientists are yet to emerge. [Nigel Shadbolt](#), has been quoted as stating that while “*Bits of it do exist in various departments around the country, and also in businesses, but as an integrated discipline it is only just starting to emerge*”⁴⁶. This highlights a gap in the training market for both data science, and its application to transport systems and personal mobility challenges. In the meantime it is widely expected that data scientists will have to be largely self-taught:

“The ability to ask questions about the data is the key, not mathematical prowess. You have to be confident at the math, but one of our top people used to be an architect.”

Nick Halstead, CEO of [DataSift](#)⁴⁶

Organisational and leadership capabilities

- 7.11 Alongside skilled data analyst roles, organisational capability needs to develop in order that people leadership roles, and decision-making groups within transport service providers and coordinating authorities are able to make strategic company decisions based on intelligence being derived from the manipulation and aggregation of multiple datasets.
- 7.12 McKinsey projected a need for 1.5 million additional managers and analysts in the US who will be able to ask the right questions and effectively interpret Big Data analytics:

“... having a core set of deep analytical talent is not enough to transform an organization, especially if the key business leaders and analysts do not know how to take advantage of this big data capability. All of the business leaders in an organization will have to develop a baseline understanding of analytical techniques in order to become effective users of these types of analyses.

Organizations can modify their recruiting criteria to take this requirement into account, but more importantly, they will need to develop training programs to increase the capabilities of their current management and analyst ranks. A basic statistics program or a series of classes in data analysis at a

⁴⁵ ONS (2012) *EMP13: Employment by industry (last updated February 2012)*. Available online from: <http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=Employment+by+Industry+Sector#tab-data-tables>, last accessed on 28/07/14. To derive this estimate we applied the percentage of jobs that make up Transport & storage (4.86%), to the CEBR’s assessment of skills shortages for open data exploitation.

⁴⁶ Wall Street Journal (2012) *Big Data’s Big Problem: Little Talent*. Available online at: <http://online.wsj.com/news/articles/SB10001424052702304723304577365700368073674>, last accessed on 28/07/14.

local college or university, for instance, could create a team of highly motivated managers and analysts that could begin this transformation.”

McKinsey Global Institute (2011)⁴⁷

- 7.13** In the transport context, this is particularly true of local authorities outside of London, and smaller companies responsible for the delivery of transport services. Not only are these organisations less-able to hire data-analyst talent needed to optimise their service delivery, but they are also generally less likely to appoint higher-calibre management talent that fully understands and appreciates these possibilities.
- 7.14** A further need for this understanding at senior/managerial level is also highlighted by recent reports that ‘freemium’ and ‘micro-payment’ business models commonly-used by app developers are becoming unprofitable. The Developer Economics report for Q3 2014 revealed for the first time that most mobile app businesses are actually unsustainable. Most of the App Store revenue was found to be generated by just 1.6% of the 2.9 million app developers worldwide (see Figure 7-1).
- 7.15** This finding was based on a large-scale online survey and one-to-one interviews with over 10,000 mobile app developers from 137 countries worldwide over 5 weeks in April and May. The survey revealed that half (50%) of iOS developers and even more (64%) Android developers are operating below an “app poverty line” of \$500 per app per month.⁴⁸

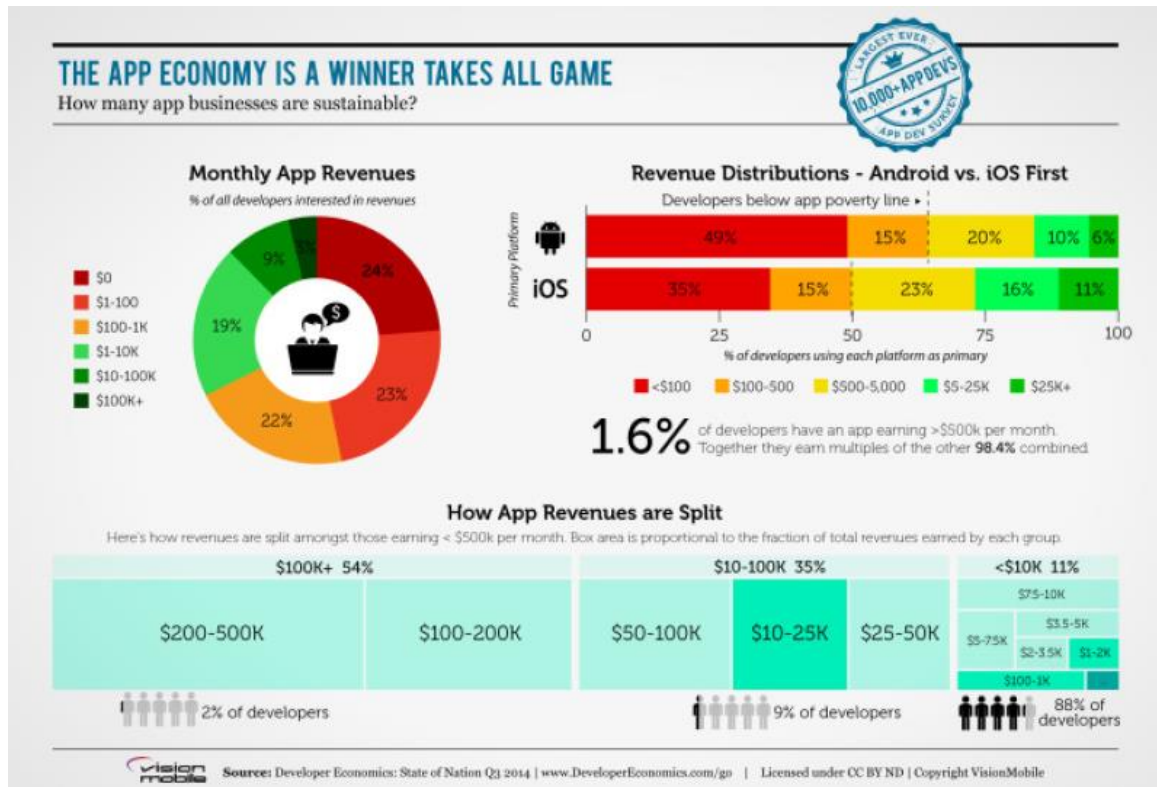
These findings emphasise the need for organisational capability within the transport industry to understand and determine the marketplace for data-driven services, applications and optimisation. Being able to build, manage and make best-use of teams of skilled people that can answer strategic optimisation questions will be an equally valuable skill to the analysis itself. The typology of transport-related data business models in Table 7-3 will help senior managers and decision-makers at established transport companies seeking to adopt new data-driven business models alongside their existing service offering, as well as prospective new ventures. It also suggests the third party app developer market may be limited in scope - explaining the importance of free/low-cost open data to this community.

⁴⁷ McKinsey Global Institute (2011) *Big data: The next frontier for innovation, competition, and productivity*, page 94. Available online at:

http://www.mckinsey.com/~media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx, last accessed on 24/07/14.

⁴⁸ Vision Mobile (2014) *Developer Economics Q3 2014: State of the Developer Nation*. Available online at: <http://www.developereconomics.com/reports/developer-economics-q3-2014/>, last accessed on 28/07/14.

Figure 7-1: The (un)sustainability of global app-based businesses⁴⁹



Investment in technology and resources

- 7.16 Ability to invest in the technologies being deployed to power data-driven mobility service is also expected to disadvantage less well-resourced organisations (such as local authorities outside of London, and smaller transport service providers) when compared with major cities and large transport operators/coordinating agencies.
- 7.17 This is already happening at [TfL](#), the [Highways Agency](#), [Network Rail](#) and [ATOC](#); who have been able to invest in real-time data collection equipment (sensor networks) and the internal/outsourced capacity to maintain live data feeds for public consumption as Open Data. This is demonstrably less common in smaller UK cities, where there have been lower levels of investment in real-time data collection systems and the organisations also have less in-house capacity to open-up and maintain the datasets they do hold. A cursory review of real-time transport datasets available on UK data catalogues and platforms reveals these are currently concentrated on strategic national road and rail networks, and in London, while almost none exist for smaller cities in the UK.

⁴⁹ Reproduced from Sarah Perez in Techcrunch (2014) *The Majority of Today's App Businesses Are Not Sustainable*. Available online at: <http://techcrunch.com/2014/07/21/the-majority-of-todays-app-businesses-are-not-sustainable/>, last accessed on 28/07/14. Copyright VisionMobile.

- 7.18 This disparity is widening in an age of austerity and public sector budget cuts – particularly those affecting the revenue needed to run data services and subscriptions. Of 2,000 surveyed Regional and National Capital cities with 200,000 to 1 million inhabitants, less than 40% plan to increase spending on traffic management in the next five years. “Lack of funding” was also cited by 40% of cities as the principal reason they were not more effective at managing traffic. Too few experts and HR resource were commonly cited as preventing effective traffic management, which demonstrates the need to be proactive about the forecast lack of data scientist and managerial talent highlighted previously⁵⁰.
- 7.19 Another field in which few smaller organisations have awareness is the procurement of cloud computing and data processing as a service. Such services can be highly cost-effective and flexible to an organisation’s data storage and processing needs. The main barriers (in particular for the public sector) to established transport service providers unlocking these technologies are:
- Awareness of its benefits, capabilities and availability.
 - Previous ‘sunk’ investment in bespoke servers that are not web-connected, but not yet obsolete (so hard to make a case for replacement/investment).
 - The ability to make regular payments for such services (revenue expenditure) versus one-off payments for ICT equipment (capital expenditure).

These findings are transferrable to transport operators/coordinators and mobility service providers in general. These organisations have scope to understand the importance of technology investment to support data-driven analytics; and tailor these to suit their needs; but are only likely to do so by: developing in-house knowledge, learning from best practices, and using these as the basis for being bold, knowledgeable and accountable⁵⁰.

⁵⁰ Siemens (2014) *Traffic Management Transformed: Moving cities forward*. Available online at: http://www.credo-group.com/downloads/WP_TrafficManag+CaseStudies_e_interaktiv.pdf, last accessed on 28/07/2014.

8 What do roadmaps for intelligent mobility look like?

- 8.1 This chapter contains five ‘roadmaps’ that link the intelligent mobility services identified in Chapter 2, with datasets discussed in Chapter 3, and the five Intelligent Mobility themes. These roadmaps are based on outputs from the final project workshop run, which we used the findings from our desk-based meta-analyses to reflect upon and validate.
- 8.2 Each roadmap is based on a ten year time period running from 2014 to 2024 and is linear in nature. The intelligent mobility services, and datasets on which they rely, are considered symbiotic - they both relate to, and enable, one another. Table 8-1 provides commentary on these roadmaps, while more details on the specific data requirements of each intelligent mobility service are documented in Appendix A to this report.

Table 8-1: Commentary on intelligent mobility roadmaps (by intelligent mobility theme)

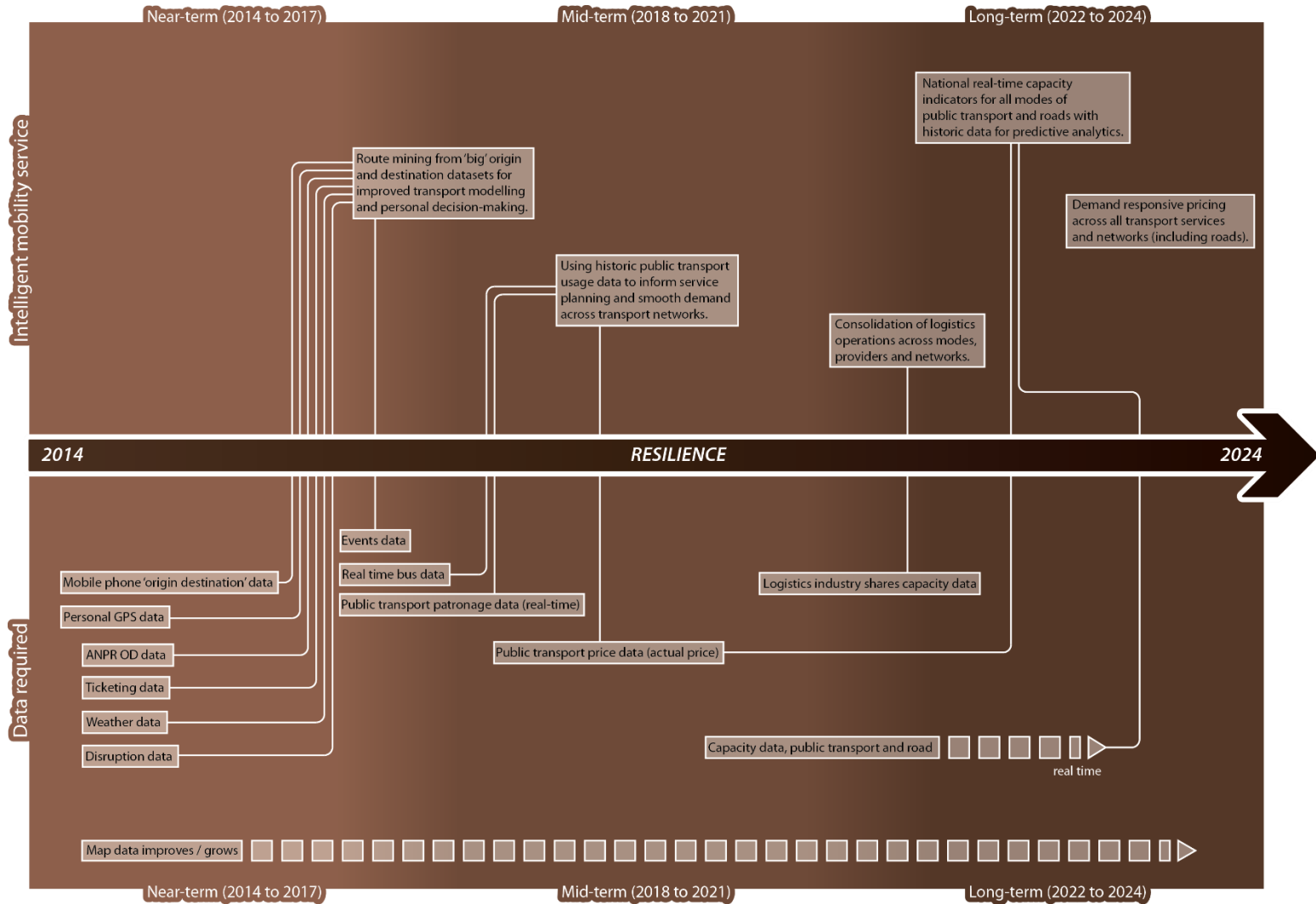
Resilience	Data created/consumed by autonomous systems will enhance resilience. Archived real-time datasets key to underlying predictive/pattern analytics. Statistical sampling and inference techniques required for composite data. Political/market forces, rather than data, are expected to delay demand responsive pricing for roads, and optimisation across UK logistics networks.
Autonomous Systems	Real-time data from network sensors, personal and vehicle locations is key. Maintained geospatial parking, road speed-limit, and ‘places’ datasets critical for driverless road-based vehicle services and systems. Person/‘thing’-level origin-destination demand data needed to power real-time trip aggregation and vehicle route definition. Legislative and public acceptability challenges considered greater than data availability and technological possibilities for driverless vehicles and drones.
Smart Infrastructure	Integrated cashless payment systems embedded in transport & retail key to adaptive public transport and road-based mobility services. Individual’s purchasing ‘exhaust’ data can be aggregated to predict demand for infrastructure, but current £20 ⁵¹ limit of e-purse legislation is a barrier. Expansion of smart sensor networks (on-street parking bay monitoring) and vehicle-based systems (e.g. monitoring ‘parked’ status) also important.

⁵¹ From September 2015, the current £20 e-purse limit will rise to £30, partly alleviating this issue: http://www.theukcardsassociation.org.uk/contactless_merchant/contactless_limit.asp

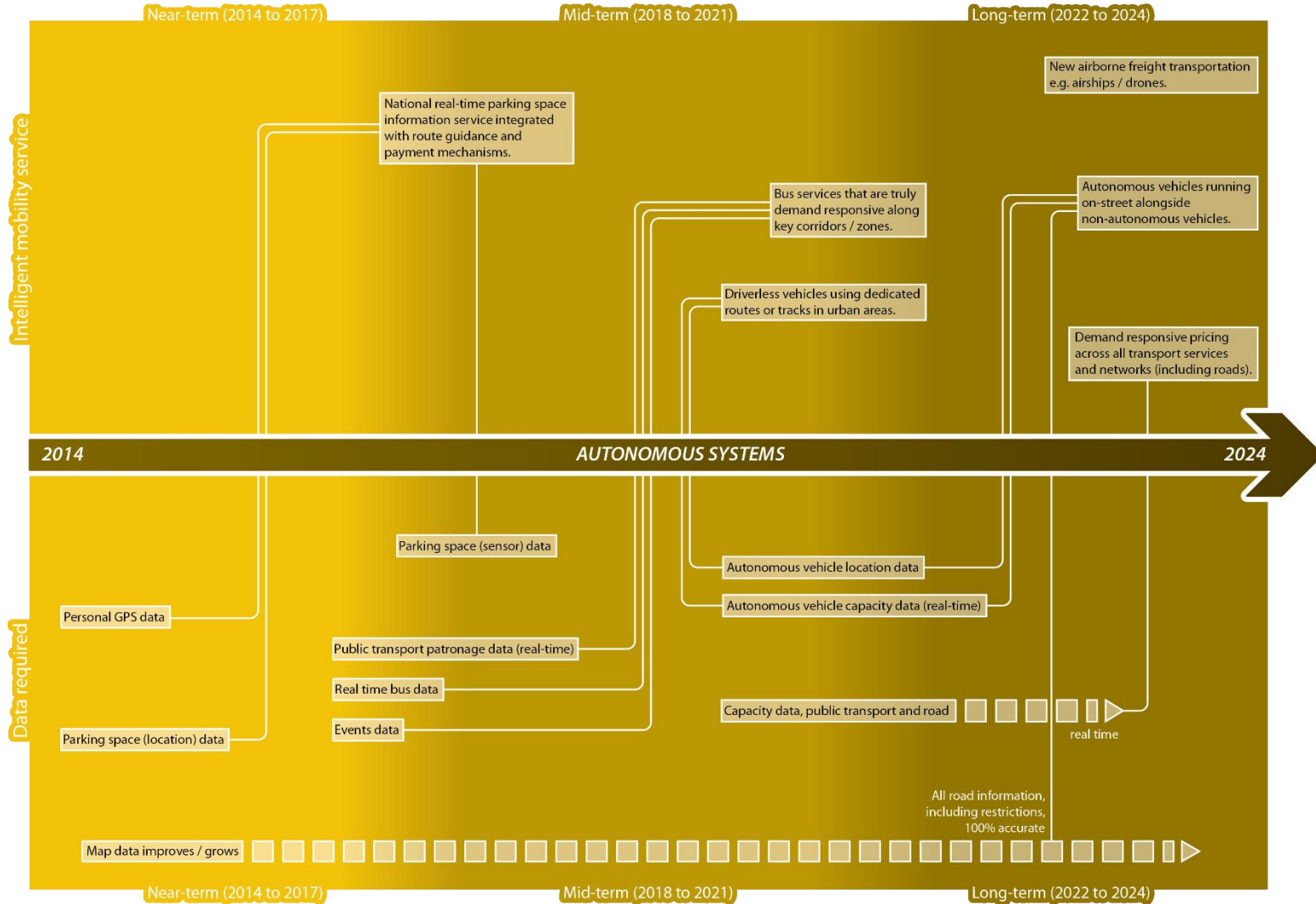
<p>End-to-End journeys</p>	<p>Greatest number of new IM services expected to emerge in this field. Critically rely on location/schedule data shared by people & operators.</p> <p>Pricing and patronage data feeds key to demand-responsive services that adjust quickly to changes in patronage/capacity demand.</p> <p>Hyperlocal, and national, event and weather datasets also required.</p>
<p>Information Exploitation</p>	<p>Collecting and sharing personal location/retail/consumption data is key for personalised and context-specific consumer information services.</p> <p><u>Right to be Forgotten online</u> and other <u>privacy issues</u> overcome by account-based services (Uber, Google Now), but they are closed and prevent data sharing across different transport systems and services.</p> <p>Some data owners deterred from sharing client data by Data Protection Act.</p> <p>Public awareness of value of their ‘digital exhaust’ may drastically change availability of covertly crowdsourced datasets relating to mobility.</p> <p>Indoor mapping quality vital for personalised routing and improving accessibility of public places (interchanges) for disabled people.</p>

8.3 The remainder of this chapter documents each of these roadmaps in-turn. We recommend these roadmaps are maintained and updated as progress is made (in terms of data, technology, legislation, capabilities, and public attitudes) against each of its five intelligent mobility themes.

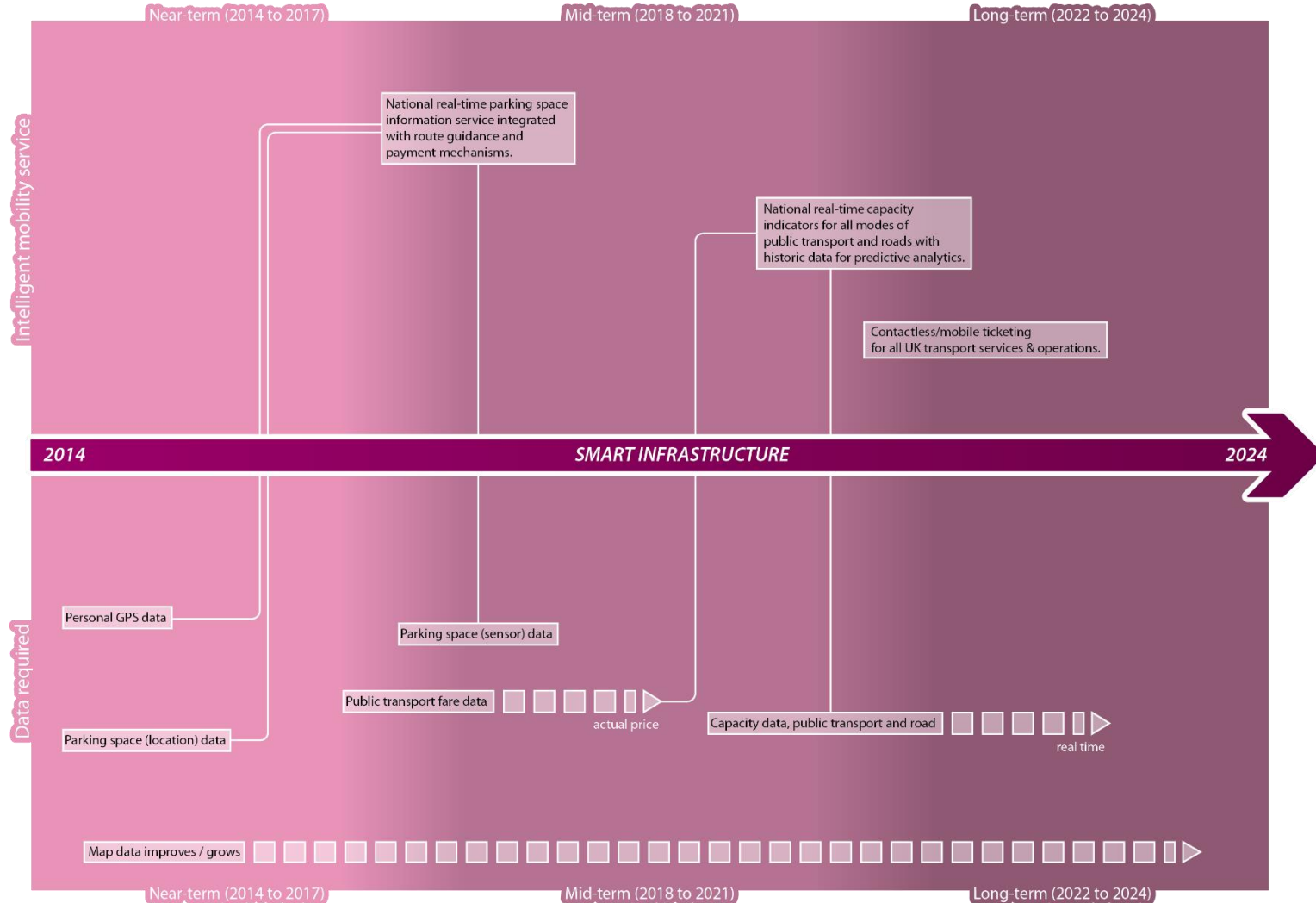
Resilience



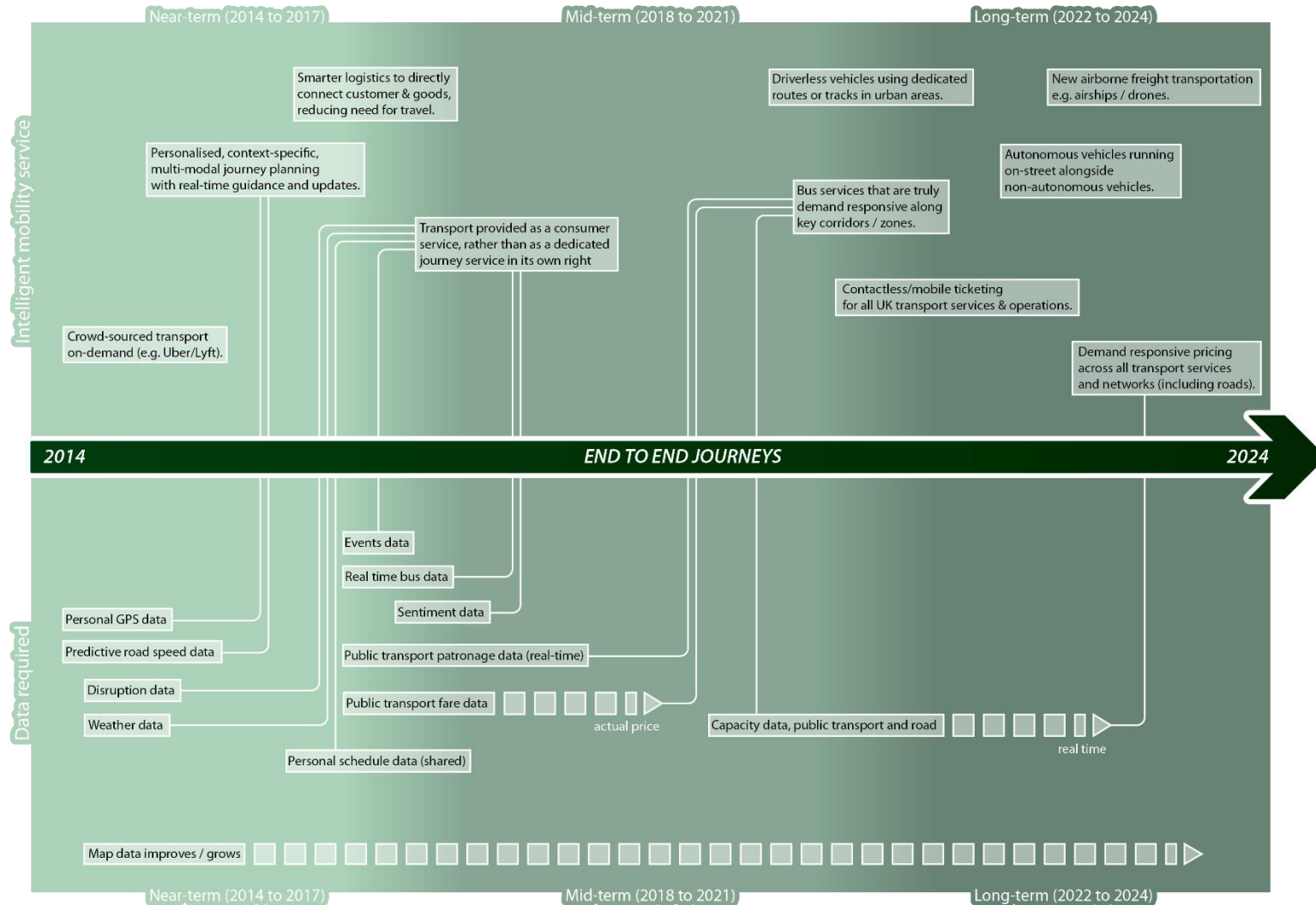
Autonomous Systems



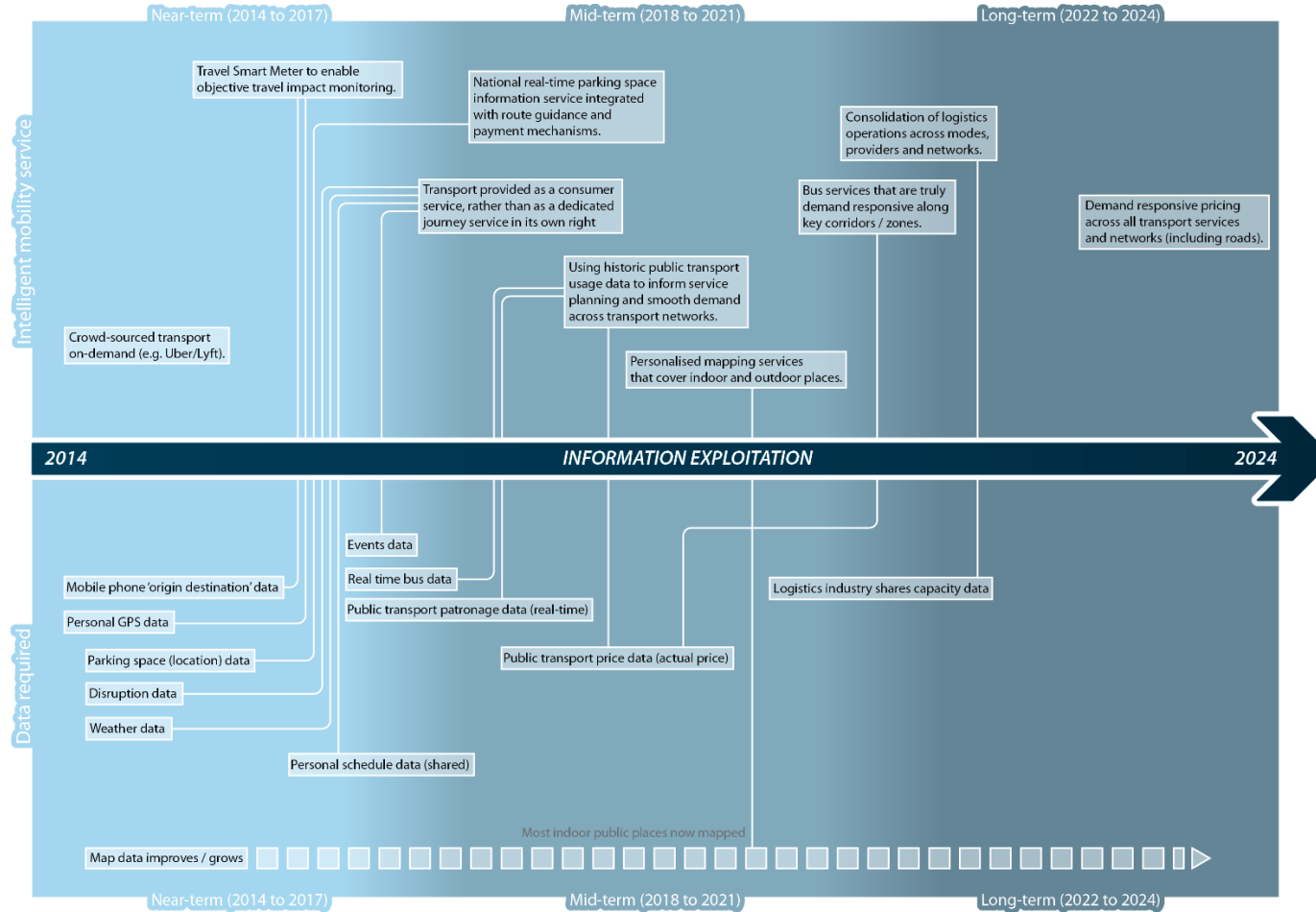
Smart infrastructure



End to End journeys



**Information
Exploitation**



9 What can the UK do to catalyse transport data exploitation?

- 9.1 This report has comprehensively documented our team’s findings from the 4 month review of transport-related data, and data-related services that will power the intelligent mobility services predicted to emerge over the next 10 years. In this final Chapter we reflect on our findings to identify 20 possible actions required to accelerate the adoption of intelligent mobility services, deepen integration between existing transport networks, and establish data-driven analytics that support more intelligent and optimised mobility services for people and things.
- 9.2 We have separated our recommendations into two sections. The first sets out the identified actions that could potentially be delivered independently by government, industry, or the Transport Systems Catapult. The second defines actions that required coordinated action across industry, including the public and private sector. Both are divided into quick wins, and strategic long-term activities.

Independent activities

Quick Wins	
1	Establish an advice service, with accompanying data collection and publishing tools, to help UK public sector open-up, maintain, and exploit the core transport datasets identified by this review. Promote linked open data in line with ODI good practice.
2	Incubate, publicise, and maintain an online transport data ‘catalogue of catalogues’ that becomes the ‘Argos’ or ‘Amazon’ for UK transport data. Requirements include: <ul style="list-style-type: none"> • Focus on core datasets for intelligent mobility identified in Table 3-3 of this report. • Include both open (free) and available (licensed/paid-for) datasets. • Showcase transport datasets so they are easy to discover and learn about. • Achieve national coverage for each dataset by grouping locally maintained feeds. • Use the IoT Hypercat standard, so data are easy to find and exploit by both humans and machines; and readily related to datasets ‘siloes’ in other data hubs.
3	Raise awareness of this catalogue, and the benefits associated with data-driven mobility services, analytics, and associated data-services; through regular presentations, workshops and best-practice/learning events aimed at UK city/local authorities and transport operators/coordinating agencies.
4	Compile regular (quarterly/annual) market monitoring and intelligence reports to build an evidence base valuing and tracking ‘user needs, drivers and opportunities’ for exploiting transport data, as identified in paragraph 6.3 of this report. Publish and publicise this intelligence to send commercial signals to UK transport industry and data marketplace.

Strategic longer-term activities	
5	Develop and maintain a publicly accessible cloud-hosted archive of core transport datasets (both ‘static’ and ‘real-time’), so their availability is guaranteed for historic data pattern analyses / predictive analytics in the future. This is considered important for migration to intelligent mobility based on predictive analytics (2014-2021) as a precursor to real-time service automation and optimisation (2021+).
6	Establish a centre of excellence for data-driven business model development to provide guidance and advice for innovative companies seeking to exploit transport-related data. This would draw on market intelligence reports and provide much needed user-needs focus.
7	Incubate, launch and publicise an online marketplace for commoditised and wholesale transport datasets catalogued in #2 above. The relevant party will need to act as honest broker to connect buyers and sellers of transport data together online (including re-sale through existing platforms like transportAPI). Improved transparency and reduced cost of transport data-as-a-service are expected to accelerate the development and commercialisation of user-focused intelligent mobility services.

Activities requiring co-ordinated action

Quick Wins	
8	UK transport operators, agencies, and local authorities working together to broker open releases of as many core transport datasets as possible, specifically focusing on data gaps identified in Table 3-2 of this report. Work with public sector transport agencies to access the Release of Data Fund, and Breakthrough Fund , and catalogue new open data sets in-line with recommendation #2.
9	An organisation engaging with major private sector organisations that hold data (bus operators, retailers, payment system providers) to share their data within ‘sandpit’ environments for innovative feasibility projects. Outcomes will need to showcase the intelligence, efficiencies, optimisations and mobility services that data-driven intelligent mobility providers can deliver if private sector datasets are made available for exploitation. Such projects need to provide direct routes-to-market for innovations delivering tangible value in meeting private sector intelligence/mobility needs.
10	Engage with emerging service providers (e.g. Uber) and existing transport network operators to foster data-driven integration between ‘old’ and ‘new’ mobility services.
11	Collaborate with UKTI , international development agencies, and developing/transitional city authorities to identify overseas data-driven mobility user-needs, and opportunities for UK companies to exploit intelligent mobility expertise.

12	Develop Transport Data Science courses and modules in UK Universities. Their focus should be on increasing the number of graduates with the skills that meet the core capabilities and skills identified in Chapter 7 of this review.
Strategic longer-term activities	
13	Strategic work between Innovate UK, Transport KTN, and relevant Catapults (Transport Systems, Satellite Applications, Connected Digital Economy, Future Cities) to co-ordinate innovation projects and harness the datasets they create so they are included in the catalogue described in #2.
14	Establish strategic placement opportunities in the UK transport industry to ensure these graduates are retained by UK intelligent mobility service providers.
15	Professional bodies; such as the Transport Planning Society (TPS), Chartered Institution of Highways & Transportation (CIHT), and Chartered Institute of Logistics and Transport (CILT); to build organisational & management capabilities through transport data analytics training schemes for professionals in the sector.
16	Support the growth of IoT sensor networks through demonstration/collaborative technology projects with transport industry partners. These could be established through demo projects (e.g. Finnish Traffic Lab) to quantify benefits and impacts of integrating IoT sensor data from road, parking, public transport, walk/cycle sources; and accelerate uptake of demand-responsive mobility networks and services.
17	Transport sector to participate in debate around data privacy issues, and ethical personal data/'digital exhaust data' re-use. Mobility organisations keen to share user's data, and providers looking to consume it, would benefit from clear guidance on legitimate uses and appropriate practices for anonymising and aggregating user-derived data. Consumer guidance on protecting personal (location) data and their rights
18	Transport sector to play an active role in supporting, defining, and maintaining increasingly global common data formats (DATEX II and GTFS) and defining new standards (e.g. for the exchange of personal location data) that will be needed to drive intelligent mobility service delivery over the next 10 years.
19	Engage in/influence government discussion around cashless payment and e-purse legislation. The current £20 threshold will become a barrier to integrated cashless payment systems for future intelligent mobility.
20	Engage with all UK major freight handlers and forwarders, e.g. through industry representative groups, to identify non-competitive/collaborative opportunities for data-driven sharing of logistics networks/operational data so as to maximise optimisation and efficiency across the sector.

Appendix A – Intelligent Mobility services identified through this review

Emerging intelligent mobility services: Short-term (2014-2017)

The six intelligent mobility services that could feasibly emerge within the next 3 years primarily reflect information services providing the kind of analytics and agglomerated information feeds that enable smarter, system-wide operational enhancements and individual journey planning in the short-term. In the longer-term future these information-focused services are also likely to power more tangible mobility services that seek to build upon the real-time analytics and network usage information they create.

Intelligent Mobility Service	Related innovation challenges	Complex / multi-source Big Data challenge?	Key datasets needed
Route mining from ‘big’ origin and destination datasets for improved transport modelling and personal decision-making.	<ul style="list-style-type: none"> Improving the traveller experience at interchanges Improve traveller experience using relevant information Minimising disruption using adjacent networks. Developing insights to improve network performance. 	<ul style="list-style-type: none"> Multiple location-based data sources (GPS, cell mast, WiFi, Bluetooth). Noisy data needs cleaning Vast real-time quantities 	User’s location data, transport operator’s location data, comms provider’s network usage data.
National real-time parking space information service integrated with route guidance and payment mechanisms.	<ul style="list-style-type: none"> Improving the traveller experience at interchanges. Improve traveller experience using relevant information. Developing insights to improve network performance. 	<ul style="list-style-type: none"> Whole network scale. Multiple raw data feeds. Integration with routing & transactions. 	Parking space locations, prices, ticket sales/barrier entries, speed limits, live road traffic conditions.
Crowd-sourced transport on-demand (e.g. Uber/Lyft).	<ul style="list-style-type: none"> Offering end to end mobility as a service. Improve traveller experience using relevant information. 	<ul style="list-style-type: none"> Managing & prioritising service requests Store historical O-D data. 	Road alignments, points of interest, public transport stops/stations.

Travel Smart Meter to enable objective travel impact monitoring	<ul style="list-style-type: none"> • Improve traveller experience using relevant information. 	<ul style="list-style-type: none"> • Collecting, storing and visualising personal data for easy self-analysis. 	Road & public transport routes, service schedules, user's location data
Personalised, context-specific, multi-modal journey planning with real-time guidance and updates.	<ul style="list-style-type: none"> • Improve traveller experience using relevant information. • Improving the traveller experience at interchanges. • Minimising disruption using adjacent networks. 	<ul style="list-style-type: none"> • Multiple real-time data feeds across all modes. • Smart analysis of options & availability in real-time. 	User's location data and preferences, all transport route/schedule & current operating status
Smarter logistics to directly connect customer & goods, reducing need for travel.	<ul style="list-style-type: none"> • Delivering seamless freight. 	<ul style="list-style-type: none"> • Matching locations of customers with goods to be delivered in real-time. 	Deliverer/recipient real-time location, live road traffic conditions.

Emerging intelligent mobility services: Medium-term (2018-2021)

The five intelligent mobility services anticipated to emerge 4-7 years from now are generally more tangible, direct 'mobility' offerings that will increase the range of ways people and goods can move around. In some cases they represent a significant re-imagining of existing transport services (e.g. truly demand responsive bus services), whilst others are new modes of travel (e.g. automated vehicles running on tracks in urban environments).

Intelligent Mobility Service	Related innovation challenges	Complex / multi-source Big Data challenge?	Key datasets needed
Transport provided as a consumer service, rather than as a dedicated journey service in its own right	<ul style="list-style-type: none"> • Offering end to end mobility as a service • Enabling whole journey accessibility 	<ul style="list-style-type: none"> • Deeper understanding of why people travel • Integration between retailer and transport provider datasets 	Personal schedule data, real-time & historic travel demand, real-time & historic consumer service demand.

<p>Personalised mapping services that cover indoor and outdoor places.</p>	<ul style="list-style-type: none"> • Improving the traveller experience at interchanges • Improve traveller experience using relevant information. 	<ul style="list-style-type: none"> • Whole world scale and micro-level of detail. • Multiple raw data feeds. 	<p>All place & space data, user’s sentiment data, user location data.</p>
<p>Using historic public transport usage data to inform service planning and smooth demand across transport networks</p>	<ul style="list-style-type: none"> • Developing insights to improve network performance • Improving the traveller experience at interchanges • Quality of life & city economy factored into transport. • Systems approach to transport investment and policy. 	<ul style="list-style-type: none"> • Probabilistic analysis of historic capacity datasets • Real-time analysis and automated intervention 	<p>Real-time and historic demand for travel and capacity data, all transport routes/schedules & real-time operating status</p>
<p>Bus services that are truly demand responsive along key corridors / zones.</p>	<ul style="list-style-type: none"> • Offering end to end mobility as a service. • Improving the traveller experience at interchanges. • Enabling whole journey accessibility 	<ul style="list-style-type: none"> • Matching bus & passenger locations. • Navigating to specific locations / avoiding traffic. 	<p>Real-time location data, place & space data, service schedules, real-time bus-loads & traffic.</p>
<p>Driverless vehicles using dedicated routes or tracks in urban areas</p>	<ul style="list-style-type: none"> • Offering end to end mobility as a service • Enabling whole journey accessibility 	<ul style="list-style-type: none"> • Probabilistic analysis of historic capacity and demand datasets 	<p>Real-time & historic travel demand, related transport schedules and delays.</p>

Emerging intelligent mobility services: Long-term (2022-2024)

The final group of future mobility services are chiefly considered innovative on account of the fact they require both technological and transport data innovation before they are likely to come to market; and are therefore considered harder to deliver:

- Services such as drone-based freight transportation and autonomous vehicles that respond and react to human-controlled vehicles operating on-street in traffic are unproven technologies.

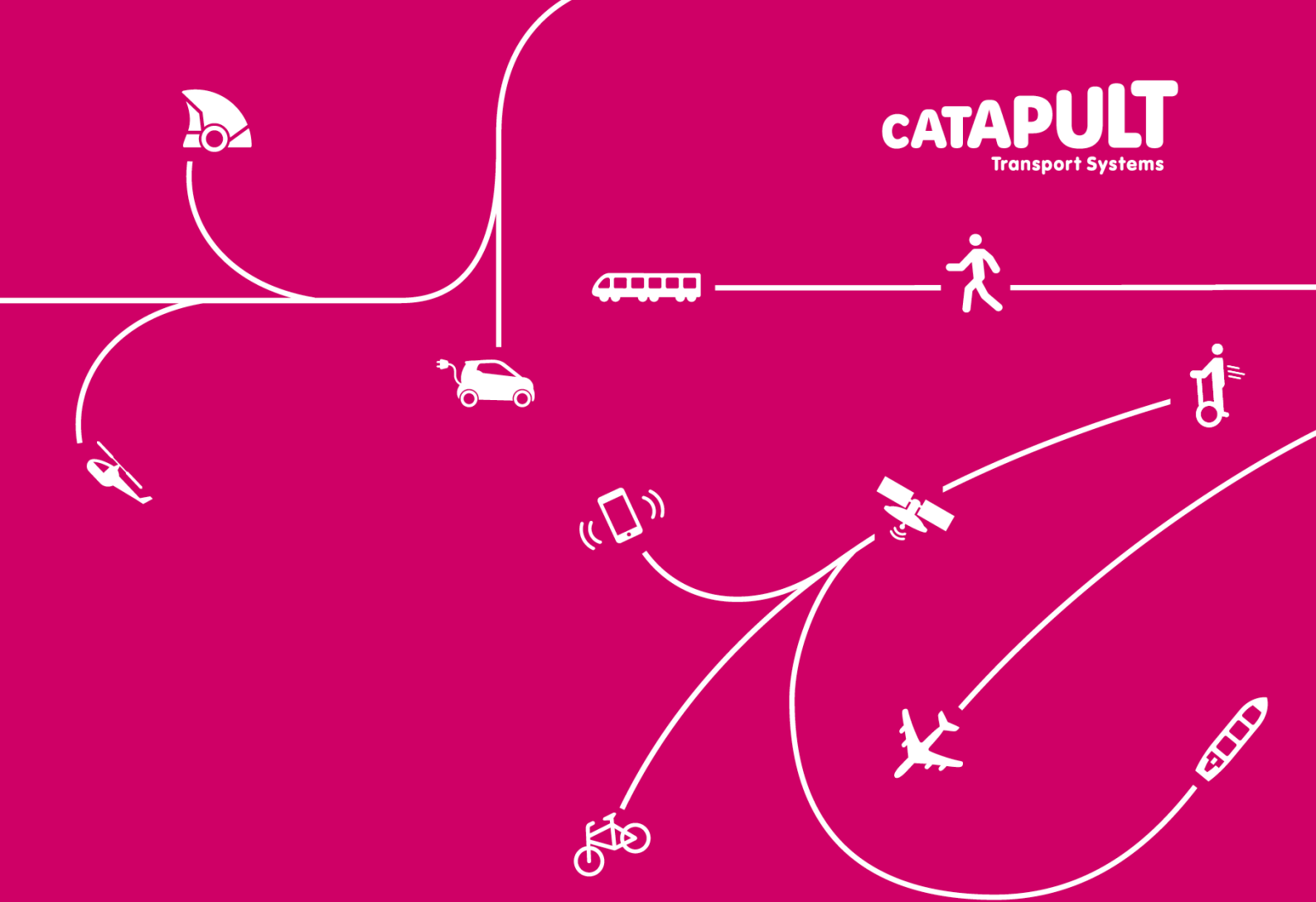
- Consolidating logistics networks to the extent good can be shipped by different service providers using shared systems is considered complex because of the vertically integrated nature of existing supply chains and existing competing interests.
- Ubiquitous contactless/mobile ticketing and network-wide demand responsive ticketing require both technological advances, changes to legislation and information-sharing agreements between service providers that currently have competing interests.

Intelligent Mobility Service	Related innovation challenges	Complex / multi-source Big Data challenge?	Key datasets needed
National real-time capacity indicators for all modes of public transport and roads with historic data available for predictive analytics.	<ul style="list-style-type: none"> • Minimising disruption using adjacent networks. • Developing insights to improve network performance. • Quality of life & city economy factored into transport. • Systems approach to transport investment and policy. 	<ul style="list-style-type: none"> • Whole network scale. • Multiple raw data feeds. • Constant data collection • Volume of stored data growing over time. 	Road & public transport routes, service schedules, ticket sales, boarding data, vehicle counts, traffic speeds.
New airborne freight transportation e.g. airships / drones	<ul style="list-style-type: none"> • Seamless freight • Quality of life & city economy factored into transport 	<ul style="list-style-type: none"> • Real-time locations of people receiving deliveries • Guidance and avoidance of other airborne vessels 	Real-time location data of service users and delivery vessels, air traffic control data, place & space data
Autonomous vehicles running on-street alongside non-autonomous vehicles.	<ul style="list-style-type: none"> • Offering end to end mobility as a service. • Enabling whole journey accessibility of transport systems. 	<ul style="list-style-type: none"> • Real-time locations of people requiring a vehicle • Guidance and avoidance of other road users 	Real-time location data of vehicles and users, precise road geometries and positioning
Contactless/mobile ticketing for all UK transport services & operations.	<ul style="list-style-type: none"> • Improving the traveller experience at interchanges • Incentivising seamless journeys through interchanges 	<ul style="list-style-type: none"> • Match boarding & alighting points across all transport networks 	Transaction data and fares for all UK transport services

<p>Demand responsive pricing across all transport services and networks (including roads)</p>	<ul style="list-style-type: none"> • Improve traveller experience using relevant information • Minimising disruption using adjacent networks 	<ul style="list-style-type: none"> • Probabilistic analysis of historic pricing datasets • Real-time analysis and automated intervention 	<p>Transaction data and fares for all UK transport services.</p>
<p>Consolidation of logistics operations across modes, providers and networks</p>	<ul style="list-style-type: none"> • Delivering seamless freight • Developing insights to improve network performance 	<ul style="list-style-type: none"> • Supply chain capacity and availability across different firms / transport modes 	<p>Real-time and historical freight capacity and movement data</p>

CATAPULT

Transport Systems



Transport Systems Catapult
The Pinnacle, 170 Midsummer Boulevard
Milton Keynes MK9 1BP
Tel: 01908 359 999

www.ts.catapult.org.uk
[linkedin.com/company/transport-systems-catapult](https://www.linkedin.com/company/transport-systems-catapult)
Twitter: @TSCatapult