



LEVERAGING BIG DATA FOR MANAGING TRANSPORT OPERATIONS

Deliverable 1.2

Big Data Policies

Jasper Tanis¹, Tharsis Teoh¹, Anna Pauliina Aavik¹, Arnaud Burgess¹, Jasmien César², Remy Russotto³

Panteia B.V.¹, Bird & Bird², Confederation of Organizations in Road Transport Enforcement³

May 2018

Work Package 1

Project Coordinator

Prof. Dr. Rajendra Akerkar (Western Norway Research Institute)

Horizon 2020 Research and Innovation Programme

MG-8-2-2017 - Big data in Transport: Research opportunities, challenges and limitations



Distribution level	Public (P)
Due date	31/05/2018 (M7)
Sent to coordinator	28/05/2018
No. of document	D1.2
Title	<i>Big Data Policies</i>
Status & Version	<i>Final</i>
Work Package	<i>1: Setting the stage on big data in transport</i>
Related Deliverables	<i>D1.1, D1.3, D2.1, D2.2, D2.3, D2.4, D3.1, D3.2, D4.1, D4.2, D4.4</i>
Leading Partner	<i>Panteia B.V.</i>
Leading Authors	<i>Jasper Tanis , Panteia, (Chapter 1, 2, 4) Tharsis Teoh, Panteia, (Chapter 1, 3, 4) Anna Pauliina Aavik, Panteia (Chapter 2) Arnaud Burgess, Panteia Jasmien César, Bird&Bird (Chapter 2) Remy Russotto (Chapter 2)</i>
Contributors	<i>Krijn Wijnands, Panteia, (Chapter 2, 3) Gerben Zwart, Panteia, (Chapter 2, 3) Thijs Driessen, Panteia (Chapter 2) Julien Debussche, Bird&Bird (Chapter 2) Sophie Roizard, CORTE (Chapter 2) Kim Hee, GUF (Chapter 2) Rajendra Akerkar (Chapter 2)</i>
Reviewers	<i>Otto Andersen, WNRI Minsung Hong, WNRI</i>
Keywords	<i>Big Data, Transportation, Public Policies, Private Policies</i>

Disclaimer:

This report is part of the LeMO project which has received funding by the European Union's Horizon 2020 research and innovation programme under grant agreement number 770038.

The content of this report reflects only the authors' view. The European Commission and Innovation and Networks Executive Agency (INEA) are not responsible for any use that may be made of the information it contains.

Executive summary

Big data applications in the transport sector have achieved national and EU-level interest as a driver for future economic growth and at the same time a source of concern, in terms of negative socio-economic impacts. This report reviews current policies implemented in the EU, its Member States and internationally, which support or restrict the (re-) use, linking of and sharing of data, in the context of big data techniques and in the transport sector. Also, the report illustrates in selected examples of transport-related private companies, the types of private sector policies that have been adopted or promoted.

While there are not any distinctly big data policies, each political entity has implemented some policies aimed at protecting the privacy of its citizens, encouraging data sharing among private and public sector entities, and develop policies that support the digitalization of the transport sector. Some of the key areas of policy in the transport sector are for instance the implementation of Intelligent Transport System, the increased Open Data policies, Automated Driving, and Smart Mobility.

Preceding and in light of these developments, the private sector has also moved ahead to incorporate the use of big data techniques into their own business models as process or product innovations. The potential applications in the transport sector are diverse, as digitalization is a major trend of the transport sector. The report covers six distinct transport sub-sectors, where the application of big data is or potentially could be used. The aim is to highlight the challenges and enablers of data sharing in the different cases. The cases are: Railway Operators, Open Data in the Airport Operator Context, Real-time Road Traffic Management, Big Data in Supply Chain Management, Managing Port operations, and Connected and Automated Vehicles.

The development of a policy roadmap to foster the growth of big data in transport will require an understanding of how existing policies affect the economic, political, social and legal environment for government and private actors in the transport sector. This work will be carried out Work Package 2.

Table of contents

Executive summary	II
List of Figures	V
List of Tables	VI
Abbreviations.....	VII
1 Introduction	1
1.1 Abstract.....	1
1.2 Purpose of the document	2
1.2.1 <i>Scope of policies considered</i>	2
1.2.2 <i>Review approach</i>	3
1.3 Target audience	3
2 Understanding policies on big data in the public sector	5
2.1 Overview of policies on big data at the European level.....	5
2.2 National policies that are restricting or that hinder access to, linking of, and (re-) use of big data	17
2.2.1 <i>General Data Protection Regulation (GDPR)</i>	17
2.2.2 <i>Data Privacy Policies in Third Countries</i>	18
2.3 National policies that are facilitating the access to, linking of, and (re-) use of big data	19
2.3.1 <i>Intelligent Transport Systems Directive (“ITS Directive”)</i>	20
2.3.2 <i>Open Data</i>	23
2.3.3 <i>Automated driving</i>	28
2.3.4 <i>Smart mobility</i>	29
2.3.5 <i>General policies</i>	30
3 Understanding big data policies in the private sector	31
3.1 Railway operators	31
3.2 Open data in the airport operator context	35
3.3 Real-time road traffic management	40
3.4 Big data in supply chain management	44
3.5 Managing port operations.....	48



3.6	Connected and automated vehicles	52
4	Conclusion	59
	References	60

List of Figures

Figure 1 Data required for track maintenance planning systems (Source: (Ghofrani, He, Goverde, & Liu, 2018)) .	32
Figure 2 Description of the digital railway programme of the Rail Delivery Group	34
Figure 3 Illustration of the aviation, travel and tourism ecosystem (Source: (World Economic Forum, 2017))	35
Figure 4 Value of digital transformation in aviation industry (World Economic Forum, 2017)	37
Figure 5 Enabling technologies for becoming a digital airport (ACI, 2017)	38
Figure 6 Overview of function of RITIS.....	43
Figure 7 SCM data according to volume, velocity and variety (Rozados & Tjahjono, 2014)	45
Figure 8 Illustration of data in the supply chains of the future (KPMG, 2017)	47
Figure 9 Traditional communication in a port versus communication with a port community system (Posti, Yliopiston, Koulutus-Ja, & Julkaisuja, 2012)	50
Figure 10 Multi-layer structure of PCS functions and modules (Carlan, Sys, & Vanelslander, 2016)	50
Figure 11 Description of different levels of automation in vehicles	54
Figure 12 Traffic management 2.0 stakeholder overview and interactions (Vlemmings et al., n.d.)	56

List of Tables

Table 1: General European Union Policies related to big data impacting the transport sector	6
Table 2: European Union road transport policies impacting big data	10
Table 3: European Union air transport policies impacting big data	11
Table 4: European Union maritime and inland waterway transport policies impacting big data	13
Table 5: European Union rail transport policies impacting big data	14
Table 6: European Union general transport policies impacting big data	15
Table 7: Overview of the status of national data access points as a mechanism for accessing, exchanging and reusing transport related data under Delegated Acts of ITS Directive 2010/40/EU	22
Table 8 Big data sources for RTS operations (Adapted from (Ghofrani et al., 2018))	33
Table 9 Types of Open APIs offered by Schiphol Airport	39
Table 10 ITS data sources and data collection technology (Khan, Rahman, Apon, & Chowdhury, 2017)	41
Table 11 Description of RITIS components	44
Table 12 Applications of big data in supply chain operations (KPMG, 2017)	45
Table 13 Benefits and cost for the PCS (Carlan et al., 2016)	51
Table 14 Functions of connected vehicle systems (summarized from Shladover, 2017)	53

Abbreviations

Abbreviation	Expression
ACEA	European Automobile Manufacturers Association
ACI	Airports Council International
ACRIS	Airport Community Recommended Information Services
API	Application Programming Interface
ATO	Automatic Train Operation
AV	Automated Vehicle
B2B	Business-To-Business
B2C	Business-To-Consumer
BMVI	Bundesministerium für Verkehr und digitale Infrastruktur (Germany's federal ministry of transport and digital infrastructure)
CCTV	Closed-circuit television
CDAS	Connected Driver Advisory Systems
CV`	Connected Vehicle
DaaS	Data-as-a-Service
DSRC	Dedicated Short-Range Communications
ERP	Enterprise Resource Planning
ETCS	European Train Control System
FPS	Federal Public Services
GDPR	General Data Protection Regulation
GPS	Global Positioning System

I2V	Infrastructure-To-Vehicle
IATA	International Air Transport Association
ITS	Intelligent Transport Systems
IWT	Inland Waterway Transport
JSON	JavaScript Object Notation
LeMO	Leveraging Big Data for Managing Transport Operations
LTA	Land Transport Authority
MDM	Mobility Data Marketplace
OEM	Original Equipment Manufacturer
PCS	Port Community System
POS	Point-of-Sales
PSI	Public Sector Information
RFID	Radio-Frequency Identification
RIS	River Information Services
RITIS	Regional Integrated Transportation Information System
RPSI	Re-use of Public Sector Information Regulations 2015
RTS	Railway Track Systems
SCM	Supply Chain Management
SKU	Stock-Keeping Unit
TM	Traffic Management
TM2.0	Traffic Management 2.0
V2I	Vehicle-To-Infrastructure

V2P	Vehicle-To-Pedestrian
V2V	Vehicle-To-Vehicle
V2X	Vehicle-To-Anything

1 Introduction

Recent developments in the quantity, complexity and availability of data collected from and about transport, combined with advances in information and communication technology, are presenting new opportunities to create more efficient, smarter, safer and more secure transport systems for people and freight, and to improve their performance for the operators which enables them to increase their service level. These new opportunities can be supported by public and private policies that facilitate and support access to, linking of and (re-) use of big data. These policies will improve transparency, leverage innovation and encourage economic growth.

In this deliverable the outcomes are reported of Task 1.3 'Big data policies' of the Leveraging Big Data to Manage Transport Operations (LeMO) project. In this task existing public and private sector policies on big data are reviewed and the application of these policies is identified through good practices and examples from the LeMO case study selection.

It will examine policies:

- at the European level;
- in selected Member States;
- in selected third countries (e.g. the US, Japan, etc.);
- and the policies of different corporate actors and institutions (for example, IT companies, transport companies, etc.).

This approach will allow for an identification of policies at different levels, and an analysis of how particular (set of) policies at specific levels have an impact on the policies at other levels. In addition, this approach also allows for a cross comparison between different actors at the same level.

We will examine policies that:

- **facilitate and support access to**, linking of and (re-) use of big data and;
- **may hinder access to**, linking of and (re-) use of big data.

Most information that is collected is of secondary nature and has been collected through desk research. In some instances, information is validated with local stakeholders that have a good knowledge about existing policies.

1.1 Abstract

Big data applications in the transport sector have achieved national and EU-level interest as a driver for future economic growth and at the same time a source of concern. This report reviews current public policies implemented in the EU, its Member States and internationally, which support or restrict the (re-) use, linking of and sharing of data, in the context of big data techniques and in the transport sector. Also, the report illustrates in selected examples of transport-related private companies, the types of private sector policies that have been adopted or promoted.

1.2 Purpose of the document

The LeMO project will contribute to developing a strategy that defines the research efforts necessary for the realisation of the big data economy through a consideration of the opportunities, limitations and challenges associated with big data in the transport sector. It will thus aid European stakeholders in improving adoption of technology and support actions that amplify constructive opportunities (e.g., new products and services, efficiencies, economic competitiveness, etc.) associated with big data, while diminishing limitations (e.g., privacy infringements, legal barriers, etc.). As such, the LeMO project has three main objectives:

1. To produce a **research and policy roadmap towards data openness, collection, exploitation and data sharing** to support European transport stakeholders in capturing and addressing issues, that range from technical to institutional, including legitimacy, data privacy and security.
2. To **involve European transport sector actors** in order to identify and analyse concrete opportunities, barriers and limitations of the transportation systems to exploit big data opportunities.
3. To **disseminate the LeMO findings, recommendations and the contribution of the LeMO** project to evidence-based decision making by improving knowledge on methodological and exploitation issues taking also into account economic, legal, social, institutional and technical aspects.

Task 1.1 and task 1.2 are crucial first steps to explore big data opportunities and to identify barriers and limitations to exploit big data opportunities. Task 1.3 is building on this by identifying which policies facilitate and which policies hinder the linking of and (re-) use of big data. Hereby the identified barriers and limitations to exploit big data need to be taken into account, as policies should be designed in such a way that they facilitate the linking of and use of big data, by removing the barriers to do so.

The purpose of task 1.3 is to examine policies that facilitate and support access to, linking of and (re-) use of big data, as well as those that may hinder access to, linking of and (re-) use of big data. The mapping of already existing public and private policies, and identifying if they facilitate or hinder the use of big data is a crucial first step in the process towards a research and policy road towards data openness, collection, exploitation and data sharing.

The purpose of this document is to summarise public and private policies on big data and to identify which of those policies facilitate and support access to, linking of and (re-) use of big data, and which of those policies may hinder that. The findings of this analysis will be applied to the LeMO case studies, which, among others, will provide insights into what policies are supporting access to, linking of, and (re-) use of big data, and which policies are not. The research approaches allow for an analysis between different (governmental) levels and for a cross comparison between actors at the same (governmental) level.

1.2.1 Scope of policies considered

In this document, we have restricted our study to the public and private policies according to the following criteria:

- General public policies related to (big) data impacting the transport sector, focusing more on economic or social developments.
- Public policies that related to the creation, collection, use, and sharing of data specifically within the transport sector
- Private sector policies, best practices or actions in the transport sector with relation to big data

For public policies, we have focused on EU and individual Member States' policies. However, in some notable instances, we have also included policies from countries external to the EU, such as the US, Singapore, and Japan.

For private policies, we have focused on several big data business models related to the preliminary cases study descriptions in the LeMO Grant Agreement.

1.2.2 Review approach

The primary methodological approach is desk research. For public policies, the contributors to the document have searched for policies in the respective legislation portals and in general search engines, according to the following questions:

1. Are there any national / local government or political initiatives about the use of (big) data? (e.g. White Paper, government policy, etc.)
2. Are any of the initiatives listed under Question 1 specific to the transport sector?
3. Is there any specific legislation about (big) data?
4. Is there any specific legislation about (big) data in the transport sector?
5. Is there any legislative proposal to regulate (big) data?
6. Is there any guidance / opinion / recommendation from a national or local authority relating to (big) data?
7. Are any of the documents listed under Question 5 specific to the transport sector?

These set of questions aim to capture to the extent possible, all the policies, legislation and legislative proposals, political initiatives, public guidance/opinion/recommendation, which are related to big data in the transport sector. And as such, we take primarily policies related to data, since policies of big data are not yet available.

For the private sector policy review, the approach focuses on the types of actions and conditions necessary for big data business models in the transport sector to develop and to operate. We start by defining key big data business models then perform a literature search on academic or industry descriptions of the business models. Some of the document types are best practices, white papers, literature reviews, webpages, and academic articles on the various topics. Although a search was performed for "policies" specific to a company, these were usually not obtainable for confidentiality reasons. This will be in scope of the work to be done in the course of LeMO case studies.

1.3 Target audience

This document will be made publicly available. The results of the research are especially interesting for people working for organisations in the public sector (e.g. politicians, policy

makers, policy consultants etc.) and in the private sector (e.g. managers, directors, and consultants). The analysis of public and private policies at different levels (EU level, national level) and the cross comparison of public and private sector policies in different countries, transport sectors, and company services provide benchmarks for those who are working on developing policies and strategies that maximize the opportunities for the exploitation of big data. This includes both public stakeholders that try to maximize opportunities for the gain of the society and private stakeholders that try to maximize opportunities for the gain of their corporate interests.

2 Understanding policies on big data in the public sector

In this chapter big data policies by the public sector are summarised. A key aspect in this is the open access of big data. Open access to the widest variety of information helps to achieve the highest impact, especially relating to its social impact, when using new methods or exploiting the opportunities offered by the data deluge. Open access is the key to the linking, sharing and re-use of data, which also implies high utility. Thus, the development of big data research and policy goals should be aligned and combined with open data practices to serve decision-making and meet scalability and performance challenges.

Big data contains massive amounts of personal data. Therefore, another key aspect is privacy and the balance between privacy and open access of big data. As the LeMO project is progressing, a big debate is going on worldwide about the balance is between privacy and openness, whereby some recurring questions are:

- To what degree must the privacy of citizens be protected?
- What is the role of public authorities in this?
- Who (in the private and public sector) is allowed to make use of what particular data?
- What data needs to be anonymised?
- Under what conditions can personal data be used and how will the General Data Protection Regulation (GDPR) affect the use of big data in transport in Europe?

The following sections summarize a comprehensive, but not exhaustive list of policies at the EU and national level that are related directly or indirectly to big data.

2.1 Overview of policies on big data at the European level

The mobilisation of big data in Europe is promoted by the Europe 2020 Strategy (European Commission, 2010a), established in March 2010, to provide smart and sustainable economic development of the continent. European policy on big data is entwined with policy on open data, which is also expected to foster significant innovations, public-private partnerships and increase European competitive advantage.

The 2020 Strategy is interlinked with the Horizon 2020 EU Framework Programme, which targets the promotion of research and innovation for major European concerns in economy and society. One key aspect of the Horizon 2020 programme is the provision of open access to scientific data, through which Europe and Member States should provide open access to scientific data generated by publicly funded research, particularly European Commission-funded research (European Commission, 2010c).

Table 1 - Table 6 provide an overview of policies (including legislation, legislative proposals and governance) at EU level that are related to the collection, (re-) use, sharing and linking of big data and that may impact the transport sector.

General policies related to (big) data impacting the transport sector

Table 1: General European Union Policies related to big data impacting the transport sector

Policy	Personal/non-personal data	Policy topic	Policy description
Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)	Personal data	Privacy	The GDPR has been adopted to harmonize data protection laws across Europe and to give greater protection and rights to data subjects as well as impose more stringent requirements on organizations that process personal data. It also aims to encourage the free movement of personal data within the EU.
Directive 96/9/EC on the legal protection of databases	Personal and non-personal data	Legal protection of databases (intellectual property)	This Directive aims to create a legal framework for the protection of databases and to ensure the legitimate interests of users to access the information in the databases.
<ul style="list-style-type: none"> Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector 	Personal data	Privacy	The Directive (also called the ePrivacy Directive) regulates the privacy concerns related to the electronic communications services. It aims to harmonise the legal, regulatory and technical provisions concerning the protection of personal data, privacy and the legitimate interest of legal persons in the electronic communications sector in order to avoid obstacles for electronic



<ul style="list-style-type: none"> • Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (COM(2017)10) 			<p>communications. In January 2017 the EU Commission published a draft ePrivacy Regulation (COM(2017)10), which is to replace the current ePrivacy Directive.</p>
<ul style="list-style-type: none"> • Directive 2003/98/EC on the re-use of public sector information amending Directive 2003/98/EC on the re-use of public sector information), known as the Public Sector Information (PSI) Directive • Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information (recast) (COM(2018) 234) 	<p>Personal and non-personal data</p>	<p>Re-use of public sector information</p>	<p>The Directive aims to encourage Member States to make as much public sector information available for re-use as possible. It provides a general framework for conditions governing the re-use of public sector documents.</p> <p>The European Commission has performed a public consultation end-2017 to evaluate the implementation of the PSI directive and to get feedback on how to improve accessibility and re-use of public and publicly funded data as well as on access to privately held data of public interest.</p> <p>On 25 April 2018, a Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information was published. Such proposal explicitly includes the transport sector into the scope of the rules on the re-use of public sector information.</p>



<p>Proposal for a Regulation of the European Parliament and of the Council on a framework for the free flow of non-personal data in the European Union (COM(2017)495)</p>	<p>Non-personal data</p>	<p>Free flow of data</p>	<p>The proposal for the Regulation intends to remove the obstacles related to the free flow of non-personal data which currently exist due to differing national requirements. Its objective is to achieve a more competitive and integrated EU market for data storage and/or processing services and activities.</p>
<ul style="list-style-type: none"> • Communication from the Commission of 2 July 2014 – "Towards a thriving data-driven economy" (COM(2014) 442 final) • Commission Staff Working Document of 2 July 2014 – "Report on the Implementation of the Communication 'Unleashing the Potential of Cloud Computing in Europe'" (SWD(2014) 214 final) 	<p>Personal and non-personal data</p>	<p>EU data-driven economy (focus on cloud computing and big data)</p>	<p>In this Communication, the Commission sketches the features of the data-driven economy of the future and sets out some operational conclusions to support and accelerate the transition towards it. It also sets out current and future activities in the field of cloud computing.</p>
<ul style="list-style-type: none"> • Communication from the Commission of 6 May 2015 – "Digital Single Market Strategy for Europe" (COM(2015) 192 final) 	<p>Personal and non-personal data</p>	<p>EU Digital Single Market</p>	<p>The Digital Single Market denotes the strategy of the European Commission to ensure access to online activities for individuals and businesses under conditions of fair competition, consumer and data protection, removing geo-blocking and copyright issues. The Digital Single</p>



<ul style="list-style-type: none"> Commission Staff Working Document of 6 May 2015 – "A Digital Single Market Strategy for Europe - Analysis and Evidence" (SWD(2015) 100 final) 			<p>Market Strategy is built on three pillars: (i) better access for consumers and businesses to online goods and services across Europe; (ii) creating the right conditions for digital networks and services to flourish; and (iii) maximising the growth potential of the European Digital Economy.</p>
<ul style="list-style-type: none"> Communication from the Commission of 10 January 2017 - "Building a European data economy" (COM(2017) 9 final) Commission Staff Working Document of 10 January 2017 on the free flow of data and emerging issues of the European data economy (SWD(2017) 2 final) 	<p>Personal and non-personal data</p>	<p>EU data economy</p>	<p>This Communication explores the following issues: free flow of data; access and transfer in relation to machine-generated data; liability and safety in the context of emerging technologies; and portability of non-personal data, interoperability and standards. This Communication also sets out suggestions for experimenting with common regulatory solutions in a real-life environment.</p>
<ul style="list-style-type: none"> Communication from the Commission of 25 April 2018 - "Towards a common European data space" (COM(2018) 232 final) Commission Staff Working Document of 25 April 2018 – 	<p>Personal and non-personal data</p>	<p>B2B and B2G data sharing</p>	<p>In this Communication the Commission presents measures that will make it easier for businesses and the public sector to access and re-use data coming from different sources, sectors (including the transport sector) and disciplines in the EU.</p> <p>The Staff Working Document aims to provide a toolbox for companies that are data holders, data users, or both. For</p>

<p>"Guidance on sharing private sector data in the European data economy" (SWD(2018) 125 final)</p>			<p>this purpose, it contains a "How to" guide on legal, business and technical aspects of data sharing that can be used in practice when considering and preparing data transfers between companies coming from the same or different sectors.</p>
---	--	--	--

Transport-specific policies per transport mode

Table 2: European Union road transport policies impacting big data

Road transport			
Policy	Personal/non-personal data	Policy topic	Policy description
<p>Directive 2010/40/EU on the framework for the deployment of intelligent transport systems in the field of road transport and for interfaces with other modes of transport</p>	<p>Personal and non-personal data</p>	<p>Efficient transport system</p>	<p>The Directive provides a legislative framework for the coordination of and coherent roll-out of Intelligent Transport Systems (ITS) across the EU. It aims at making more use of ITS solutions to achieve a more efficient management of the transport network for passengers and businesses.</p>
<p>Regulation (EU) No 165/2014 on tachographs in road transport, repealing Council Regulation (EEC) No 3821/85 on recording equipment in road transport and amending Regulation (EC) No</p>	<p>Personal and non-personal data</p>	<p>Road safety</p>	<p>The Regulation sets out the requirements on the construction, installation, use, testing and control of tachographs used in road transport in the European Union (EU). Digital tachographs must record data. The access to tachograph data may be granted at all times to the control authority and the transport operator concerned. Digital</p>

<p>561/2006 of the European Parliament and of the Council on the harmonisation of certain social legislation relating to road transport</p>			<p>tachographs warn drivers about exceeding the maximum allowed continuous driving time, in order to help them comply with legislation. Processing of personal data must be solely to verify compliance with this and other relevant regulations.</p>
<p>Regulation (EC) No 715/2007 on type approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information</p>	<p>Non-personal data</p>	<p>Access to vehicle repair and maintenance information</p>	<p>Manufacturers must share on websites unrestricted and standardized access to vehicle repair and maintenance information for independent operators. The manufacture is entitled to charge a reasonable fee for this service.</p>

Table 3: European Union air transport policies impacting big data

Air transport			
Policy	Personal/non-personal data	Policy topic	Policy description
<p>Regulation (EC) No 80/2009 on a Code of Conduct for</p>	<p>Personal data and non-personal data</p>	<p>Non-discriminatory</p>	<p>The Regulation sets out a harmonised Code of Conduct on the use of computerised reservation systems to ensure</p>

<p>computerised reservation systems and repealing Council Regulation (EEC) No 2299/89</p>		<p>sharing of marketing, booking and sales data</p>	<p>fair competition and protection of consumers' rights. System vendors must, <i>inter alia</i>, load and process data provided by airlines with equal care and timeliness as well as process personal data only for the purposes for which they were given. Furthermore, system vendors may make available marketing, booking and sales data, provided these are offered on a non-discriminatory basis to all participating carriers and do not identify the customer.</p>
<p>Council Directive 2004/82/EC on the obligation of carriers to communicate passenger data</p>	<p>Personal data and non-personal data</p>	<p>Communication of passenger data</p>	<p>According to this Directive, air carriers are required to communicate information concerning their passengers travelling to a European Union (EU) border crossing. This information is supplied, at the request of the authorities responsible for carrying out checks on persons at the external borders of the EU, to improve border control and to combat illegal immigration more effectively.</p>
<p>Regulation (EC) No 552/2004 on the interoperability of the European Air Traffic Management network</p>	<p>Non-personal data</p>	<p>Interoperability of air traffic systems</p>	<p>The aim of this Regulation is to define common requirements to guarantee interoperability between the various air traffic management systems used. It establishes a harmonised system of certification for components and systems. Part B of Annex II of the Regulation sets out the requirements that are specific to each one of the systems, among which flight data processing systems and surveillance data processing systems.</p>

Table 4: European Union maritime and inland waterway transport policies impacting big data

Maritime and inland waterways transport			
Policy	Personal/non-personal data	Policy topic	Policy description
Communication from the Commission of 21 January 2009 – Communication and action plan with a view to establishing a European maritime transport space without barriers	Non-personal data	e-maritime systems	In order to simplify the administrative formalities applicable to intra-EU maritime transport, the Commission proposes in this Communication to make electronic data transmission more efficient by creating “e-maritime” systems, as announced by the Commission in its White Paper on Transport Policy.
Directive 2005/44/EC on harmonized river information services (RIS) on inland waterways in the Community	Non-personal data	River information services	The Directive establishes rules on the use of harmonised RIS. These rules are designed to ensure the safety, efficiency and environmental friendliness of inland waterways in the EU. They apply to canals, rivers, lakes and ports able to take vessels between 1,000 and 1,500 tonnes. EU countries must: (i) supply all RIS users with the relevant data on navigational and voyage planning; and (ii) designate authorities to oversee RIS application and the exchange of international data.
Directive 2002/59/EC establishing a Community vessel	Non-personal data	Vessel traffic monitoring and	This Directive sets up a vessel traffic monitoring and information exchange system. This seeks to enhance

traffic monitoring and information system and repealing Council Directive 93/75/EEC		information exchange	maritime safety, port and maritime security, environmental protection and pollution preparedness. It also permits the exchange and sharing of additional information facilitating efficient maritime traffic and transport.
---	--	----------------------	---

Table 5: European Union rail transport policies impacting big data

Rail transport			
Policy	Personal/non-personal data	Policy topic	Policy description
Directive 2012/34/EU establishing a single European railway area	Non-personal data	Information exchange between regulatory bodies	The Directive establishes a single European railway area. It also adds important substantive changes to tackle the lack of competition, poor regulation and low investment observed in the rail market in the last decade. The regulatory bodies must exchange information about their work and decision-making principles and practice and, in particular, exchange information on the main issues of their procedures and on the problems of interpreting transposed Union railway law.
Regulation (EC) No 91/2003 on rail transport statistics	Non-personal data	Rail transport statistics	The Commission requires statistics on freight transport and rail transport of passengers as well as rail accidents in order to ensure the monitoring and development of the



			common transport policy and the preparation of measures in the field of transport safety. The scope of this Regulation covers all railway undertakings in the EU. Each EU country must report statistics which relate to rail transport on its national territory.
<ul style="list-style-type: none"> • Directive 2008/57/EC on the interoperability of the rail system within the Community (Recast) • Directive (EU) 2016/797 on the interoperability of the rail system within the European Union 	Non-personal data	Interoperability of rail systems	<p>Directive 2008/57 aims to improve the interoperability of and access to the national rail networks, including the technical installations for data processing and telecommunications intended for services on the network and efficient traffic management.</p> <p>Directive 2008/57 will be repealed by Directive 2016/797 as from 15 June 2020.</p>

Table 6: European Union general transport policies impacting big data

General (all transport modes)			
Policy	Personal/non-personal data	Policy topic	Policy description
Regulation 1315/2013 on Union guidelines for the development of the trans-European transport	Non-personal data	Efficiency of transport system	The Regulation establishes guidelines for the development of a trans-European transport network in order to improve the management of the infrastructure of the network. It provides priorities for the development of the network



network and repealing Decision No 661/2010			and measures for the implementation of the transport network.
White paper – Roadmap to a single European Transport Area – towards a competitive and resource efficient transport system (2011)	Personal and non-personal data	Efficiency of transport system	This White paper describes a policy agenda to enhance the internal market for transport and to build a competitive transport system that will increase mobility, remove major barriers in key areas and fuel growth and employment. It lays down 10 goals to guide the policy actions and to measure progress which are underpinned by 40 concrete initiatives to be developed over the next decade.

2.2 National policies that are restricting or that hinder access to, linking of, and (re-) use of big data

The large majority of policies that are restricting or hinder access to, linking of, and (re-) use of big data are policies aiming to protect the privacy of individuals. Therefore, the regulation of data in most cases encompasses the regulation of personal data too. The European Union is taking more control over the protection of personal data by enacting the GDPR, in force from 25th May 2018. It replaces the Directive 95/46/EC and provides uniform rules for every Member State. In doing so, the Regulation aims to encourage the free flow of personal data throughout the Union while setting the conditions under which personal data can be used. Since these are uniform rules applying for every organization and every individual in the European Union, access to, linking of, and (re-) use of big data in the transportation sector is also restricted by the GDPR.

2.2.1 General Data Protection Regulation (GDPR)

The GDPR is significantly more stringent in respect of the processing of personal data than the Directive 95/46/EC that it is replaced. Compared to the previous directive it expands the rights of data subjects to control how their personal data is collected and processed. In addition, it also places a range of new obligations on organisations that process personal data.

Two key concept of the GDPR are highlighted below.

A first key concept of the GDPR is ‘personal data’. It refers to any data related to an identified or a directly or indirectly identifiable person. This person is referred to as the “data subject”. The GDPR has clarified that any data that can be linked to a particular individual constitutes personal data, such as a name, an address, localisation data, health information, income, cultural profile qualify as personal data. The GDPR now also formally includes online identifiers such as IP addresses in its definition of personal data.

The second key concept of the GDPR is “processing”, which is understood as an operation or set of operations, whether or not by automated means, performed on personal data or on sets of personal data. These operations involve a wide range of activities including the collection, recording, organisation, storage, adaptation, alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment, combination, blocking, erasure, destruction, etc. Any activity which involves personal data, including the (re-)use, sharing of and linking of personal data is therefore subject to the principles of the GDPR.

The Regulation introduces new responsibilities and duties for companies processing personal data. For instance, they are required to “implement appropriate technical and organisational measures” taking into account “the state of the art and the costs of implementation” and “the nature, scope, context, and purposes of the processing as well as the risk of varying likelihood and severity for the rights and freedoms of individuals.”

As the GDPR is a regulation, it is directly applicable in Member States without the need for implementing legislation. However, the GDPR has left the Member States some margin of appreciation to adopt derogations or impose additional requirements in some areas such as freedom of expression and information, public access to official documents or national identification numbers. In those areas, Member States may take divergent approaches. Member States are currently in the process of drafting and adopting such national laws implementing the GDPR principles.

To the extent that big data analytics also involve the use of personal data, the GDPR may be a limiting factor for the uptake of big data in general and more specifically in the transport sector.

2.2.2 Data Privacy Policies in Third Countries

No uniform, comprehensive federal U.S. law regulating the collection, use, and sharing of personal information is in application in the USA. Instead, a collection of laws and regulations are imposed at the federal and state levels, and they do sometimes contradict one another. Compliance to the laws may be an extremely complex issue as in some cases state laws preempt over federal laws regulating the same activities and in other case it is the contrary. In addition, governmental agencies and industry groups (such as the payment card, mobile marketing and online advertising industries) develop self-regulatory guidelines that are not legally enforceable but considered as best practices and increasingly used by regulators. Most states have enacted privacy legislation. California leads the way¹. New laws and amendments are proliferating as technological threats change and progress towards a uniform federal legislation is too slow.

US privacy laws do not generally provide individuals a right to access their data (except the HIPAA and some California laws) and neither to object to the processing of their data. The Children's Online Privacy Protection Act allows a parent to view the personal information collected by a website about a child, and to delete and correct that information. The California Shine the Light Law (Civil Code §§1798.83 to 1798.84) allows consumers to learn how their personal information is shared by companies for marketing purposes and encourages businesses to let their customers opt out of this.

Big data is a major economic topic for the **Japanese government**. The 2013 White paper on information and communication states that developing big data in Japan could create large benefits. Japan has also developed strong industries in the Web 2.0 sector.

However, the debate around big data revolves as well around privacy protection. As private actors want to analyse big data, Japan plans on setting a means to certify big data users in order to leverage privacy concerns.

In **Singapore** the Personal Data Protection Act (2012) regulates personal data by prescribing rules for the collection, use and disclosure of personal data by organisations. It further provides rules on the transfer of personal data outside Singapore. Under that Act linking data is allowed

¹ <http://oag.ca.gov/privacy/privacy-laws>, Accessed 18.08.2018

even without consent of the person when it is used for a research purpose, including historical or statistical research.

Personal data and the sharing of such data between public bodies are regulated by Singapore's Public Sector (Governance) Act 2018. All agencies that request data, not just those that own it, are responsible for protecting that data. The act sets up centralised agencies that must ensure that raw data is properly anonymised before released to relevant agencies. The act lists seven purposes under which data can be shared among public sector agencies – those purposes are, for example, to improve the efficiency or effectiveness of policy planning and service delivery, and to support a whole-of-government approach in public sector work.

2.3 National policies that are facilitating the access to, linking of, and (re-) use of big data

This section provides an overview of a couple of most important policies in the European Union that are facilitating the access to, linking of, and (re-) use of big data. This refers in particular to making data publicly available so that everyone (with no or limited effort) can access it and use it. All regulations in this section refer to non-personal data, since the access to, linking of, and (re-) use of personal data is restricted by the GDPR.

Every jurisdiction has in place regulations prescribing rules for transportation in general and for mode specific transport – road, rail, maritime, freight/passenger, etc. Most of these legislations are of technical nature and do not prescribe rules about the use of (big) data in transport. The most commonly found reference to data in those legislations is found in connection to specific registers being created. For example, such registers may be in the field of national public transport, railway traffic, maritime traffic, air traffic, traffic victims (SP), licences, etc. Such registers fall under the scope of the general data protection rules of the country insofar personal data is being used. Reference to data is also often being made in connection with passenger transport where passenger personal data is being used for various reasons. The use of passenger personal data falls under the general data protection rules available in each country.

The transport policies and initiatives or strategies published by the studied countries in the field of transport share some common elements and indicate a direction where the countries are focusing their efforts in the field of transport. Those areas are:

- Intelligent Transport Systems;
- Open Data;
- Automated driving; and
- Smart mobility.

Some policies that are relevant to big data, but do not deal with the transport sector per se, are summarized in the last section.

2.3.1 Intelligent Transport Systems Directive (“ITS Directive”)

The increase of road transport in the European Union has given rise to an increasing congestion of road infrastructure, growing energy consumption, as well as environmental and social problems. In light of these challenges, on the 7th of July 2010, the European Parliament and Council adopted the ITS Directive (Directive 2010/40/EU). This was considered necessary to ensure sustainable, efficient and competitive mobility in Europe, given the projected increase in road transport, and the limitations of new physical infrastructure development.

In EU countries Directive 2010/40/EU on the framework for the deployment of ITS in the field of road transport and for interfaces with other modes of transport has been implemented into national legislation. The Directive seeks to encourage the development of innovative transport technologies to create ITSs, in particular in cross-border situations. It is applicable to road transport sector, both commercial and passenger transport. The Directive provides for the development of legally binding specifications and standards for interoperability and continuity through delegated acts. The specifications are measures, which lay down provisions (functional, technical, organizational, service) containing requirements, procedures or any relevant rules when deploying certain systems and services.

The main aim of the ITS directive is to establish interoperable and efficient ITS services while leaving EU countries to individually decide which systems to invest in. In doing so, ITS services should contribute to a reduction of the negative effects of road transport, e.g. reducing air pollution, CO₂ emissions, congestion and improving road safety.

The deployment and use of ITS includes also the processing of personal data. For that the Directive has provided rules on privacy, security and re-use of such information. The rules provided in the Directive are also compatible and add to the general data protection rules available in the EU, which all must be followed when using ITS.

The directive provides for *priority areas* and *priority actions* to be implemented at EU level through delegated acts, which are binding for the deployment of the related ITS applications and services. The delegated acts provide for the necessary specifications in order to fulfil the priority action in hand. Such specifications refer to binding measures “Laying down provisions containing requirements, procedures or any other relevant rules” (Art. 4(17)).

The *priority areas* are listed in Article 2:

1. Optimal use of road, traffic and travel data;
2. Continuity of traffic and freight management ITS services;
3. ITS road safety and security applications; and
4. Linking the vehicle with the transport infrastructure.

The *priority actions* are listed in Article 3:

1. Provision of EU-wide multimodal travel information services;
2. Provision of EU-wide real-time traffic information services;
3. Data and procedures for the provision, where possible, of road safety related minimum universal traffic information free of charge to users;

4. Harmonised provision for an interoperable EU-wide eCall²;
5. Provision of information services for safe and secure parking places for trucks and commercial vehicles; and
6. Provision of reservation services for safe and secure parking places for trucks and commercial vehicles.

For the first five of these priority actions the Commission has adopted delegated regulations. The last action is still to be developed. Any specification issued under any of these priority actions shall comply with the following principles (Annex II of Regulation 2010/40/EU):

- Be effective;
- Be cost-efficient;
- Be proportionate;
- Support continuity of services;
- Deliver interoperability;
- Support backward compatibility;
- Respect existing national infrastructure and network characteristics;
- Promote equality of access;
- Support maturity;
- Deliver quality of timing and positioning;
- Facilitate inter-modality; and
- Respect coherence.

The two priority actions which most closely concern non-personal transport data are the priority actions one and two.

Priority action 1 has been adopted by the Commission ((EU) 2017/1926) in 2017 as part of the Commission's Mobility Package "Europe on the move". It aims to encourage member states to look and develop cost-effective ways to digitise existing static and dynamic data of different transport modes. The specifications are applicable to all transport modes in the EU (schedule based, transport on demand and personal based). The rules affect transport authorities, transport operators, transport on demand service providers and infrastructure managers. A national access point must be set up by each member state that should be a single point of access for users of static travel and traffic data and historic traffic data. The inclusion of dynamic data is not mandatory, but suggested. The regulation provides rules for the linking of travel information services where they shall provide one another with routing results based on static, and where possible, dynamic information, if so requested. Such requests, and the terms and conditions should be defined in a contractual agreement between the information service providers.

The Commission Delegated Regulation (EU) 2015/962 adopted for the *second priority action* seeks to improve the accessibility, exchange, re-use and update of the road and traffic data

² The eCall is Europe's "single emergency number 112 in the event of a serious road accident and communicates the vehicle's location to the emergency services" (https://ec.europa.eu/transport/themes/its/road/action_plan/ecall_en, Accessed 28.05.2018)

required for the provision of high quality and continuous real-time traffic information services across the EU. The rules apply to both road authorities, road operators and service providers. The regulation foresees that each member state sets up a national access point for the exchange of data. The rules require that road status (both static and dynamic) and traffic data are made accessible via those access points in a standardised format. That access point can be the same as established under any other delegated regulation adopted under Directive 2010/40/EU. Such data must be accessible for exchange and re-use on a non-discriminatory basis, and within a time-frame that ensures the timely provision of the real-time traffic information services. The regulation also establishes rules on data updates including the timeliness of such updates.

The delegated regulations do not provide specific rules for the use of such transportation data. Establishing those is left up to every member states. The regulations only harmonize the data that shall be made available and the (suggested) format of such data. The national data access points can take various forms, such as databases, data warehouses, data marketplaces, repository, register, web portal or similar. The form depends on the type of data concerned. Most member states have already started to develop such national access points as required under each of the priority actions. Most member states have chosen to develop one common access point for all actions. However, some countries (e.g. Finland, Spain) have different access points for some actions, or the access points overlap. For numerous countries the work to develop these access points is still in process, and some have not yet started the development. An overview of the current status of national access points is given in Table 7.

Table 7: Overview of the status of national data access points as a mechanism for accessing, exchanging and reusing transport related date under Delegated Acts of ITS Directive 2010/40/EU

	Countries	
Status	Data action point for all delegated Acts	No data action point yet for all delegated acts
Countries that develop different data action points for some actions, or that develop overlapping action points	Netherlands, Spain, United Kingdom	Finland
Countries that develop one common data action point for all actions	Austria, Germany,	Belgium, Cyprus, Czech Republic, Denmark, France, Ireland, Italy, Lithuania, Luxembourg, Malta, Norway, Poland, Slovenia, Sweden
Countries that have not developed a data action point,		Bulgaria, Croatia, Estonia, Greece, Hungary, Latvia,

<p>or countries for which no information if they have developed a data action point is available</p>		<p>Portugal, Romania, Slovakia, Switzerland</p>
--	--	---

In 2013 the **Dutch government** started the programme “Beter Benutten³”. In this program the national government co-operates with local governments and the private sector to improve the accessibility of different regions. With this programme the government is stimulating to use existing roads and infrastructure more intelligently. An important part of this programme is the development of ITS. For example the project “Blauwe golf” is about all the facilities that provide better traffic flow and travel time gains for traffic over the water and on the road where water and roads intersect. The exchange of data plays an important role in this project.

An initiative by **Belgium** specific in the field of transportation is known as CITRUS. With the CITRUS (Cooperative Intelligent Transport Systems for Trucks) project, the Flemish Government wants to improve road safety on the Belgian roads and motor ways, reduce CO2 emissions, and promote mobility using C-ITS systems. Data is collected from vehicles and roadside systems and will be used to inform road users proactively, in real time.

In **Estonia**, data exchange between government agencies and other entities is governed by the Data Exchange Layer of Information System Act. This Act stems from the Public Information Act. It regulates the use of the so-called X-road which is a technical infrastructure and environment between X-road members that enables them to exchange information online through safe means. X-road is a software designed for state institutions and organisations to use in the course of exercising their duties. X-road is used for storing all the data related to E-Estonia (see 2.3.5) and is rooted in a blockchain called K.S.I., which is developed by Guardtime. The members of the X-road must use a specific safe-server approved by the X-road, where they use e-stamps that have to be created in the server according to the EU Regulation No 910/2014 on electronic identification and trust services for electronic transactions in the internal market (“eIDAS Regulation⁴”).

2.3.2 Open Data

Open data portals can be found in every country studied, both in EU and outside EU. The purpose of those portals is to allow public access and free (re-) use of public sector information. Such information includes areas like agriculture, fisheries, culture, sports, economy, finance, education, energy, environment, health, justice, population, regional, science, technology, and transport. The themes and categorization of data depends on a particular country, but they cover in one way or another every aspects the country has an interaction with.

The goal of opening up data is:

³ <https://www.rijksoverheid.nl/onderwerpen/wegen/wegen-beter-benutten>

⁴ Infosüsteemide andmevahetuskiht 30.09.2016

- To increase economy;
- To increase transparency;
- To provide access to information;
- To decrease the number of information requests and, consequently, also reduce the workload of public sector workers;
- To facilitate creation and management of open services for private and community sectors; and
- To encourage migration to future technologies such as linked data, big data and internet of things⁵.

The open data portal, ideally, should be a one-stop portal of publicly available datasets. That would mean there is only one portal where all data is being collected. That would result in benefits to the public who wishes to locate a specific set of data and only has to go through one portal in order to find it. Having only one portal in place also provides for uniform rules regarding privacy, sharing, re-use and access to the data, and may promote the use of linked data.

In the EU there is a legislation on the re-use of public sector information. *The Directive 2013/37/EU* provides a common legal framework for all member states to make use of government-held data. The two key aspects of that directive are transparency and fair competition. The directive has put more focus on the economic aspects related to re-use of information, and does not deal with rules on access of citizens to such information.

Under that directive, member states are encouraged to make public sector information available. It regulates the information held by public sector bodies at all levels in the country – national, regional and local levels, including ministries, state agencies, municipalities, as well as organisations funded by or under the control of public authorities. The information should ideally be provided to citizens free of charge, however for certain types of information setting a user fee is allowed on a condition that the amount set is non-discriminatory and the calculation of it is transparent.

The open data portals available in the EU member states are all influenced by that directive and thus share many common features and their set-up.

In the field of transport, all identified open data portals include datasets related to transport. In some countries however, specific transport data is also provided in different data portals focused on a specific field of transport. That however may result in duplication of the information and different rules applicable to the use of that data.

Such data portals specific in the field of transport are focused on providing data in the following listed fields:

- public transport services;
- public roads and ports;
- statistics road data;

⁵ Opendata.riik.ee

- mobility and related topics,
- environment detectors,
- parking spaces, carpark availability;
- road works;
- hazards and incidents;
- traffic management measures; and
- taxi availability.

The open data portals of governments must set clear boundaries to how they are managed, how the information is being used and gathered etc.

Most portals provide free use of data for both commercial and non-commercial purposes. The data is allowed to be freely re-used and redistributed. In Singapore, the transport related open data portal, the so-called Land Transport Authority (LTA) DataMall, that provides information on road and passenger transport modes has been used since 2015 by the public and third-party developers in individual business ventures and government organised hackathons. Having open and free access to transport data has helped developers to put such data into good use in creating people-centric transport solutions. In the hackathons, countless innovative ideas and applications are being presented, and many of them find use in real life in order to provide better mobility in Singapore.

Such open and free access however requires certain standards of safety and security. The data provided should comply with the rules and regulations concerning the private data. In Estonia, for example, the Information System Authority has published the following guidelines for users of the open data portals: (1) Guide for the creation and publication of open data; (2) Data exchange with Estonian open data portal, (3) Portal guidelines for holders of information, and (4) Open data guidelines for holders of information.

These documents provide a comprehensive overview of the rights and obligations of data holders. It provides information how to submit data, what type of data and how it should be presented (both mandatory requirements and suggested requirements to facilitate future developments), it covers anonymization and making synonymous of the data, quality control etc. However, there are no guidelines as to how such data shall be used when downloaded from the portal, and what the rules and obligations are for the people or entities to make use of such free data.

Germany has the *Federal Open Data Act* of 18 May 2017. It obliges the federal authorities to make their data publicly accessible. The act also establishes a central support agency for open data. German Government has their open data publicly available⁶.

In **Germany**, there are two open data portals specific in the area of transport. The first is *mCLOUD*, which is part of the Geo-information strategy where many geoformations from the business area can already be researched and downloaded. The mCLOUD is a research platform

⁶ Govdata.de, Accessed 18.05.2018

for open data in the field of mobility and related topics. With the mCLOUD the BMVI⁷ provides a central access point to all open data of its business area and also opens the portal for private providers from the mobility sector to offer their data there. Second is the *Mobility Data Marketplace (MDM)*, which is a B2B platform for data exchange. The MDM provides data such as traffic and environment detectors, parking spaces, road works, hazards and incidents, traffic management measures. It is only a platform for exchange of data, it does not change the data nor store it.

The Netherlands has also adopted several acts in relation to making data open, accessible and re-usable. With regard to re-use of (big) data, The Netherlands has transposed the Directive 2013/37/EU into the Act hergebruik van overheidsinformatie. This law provides options for citizens and companies to submit requests for the provision of government information. Information and data from the government should be actively made available in a machine-readable format to be reused for other purposes, even commercially. The main goal of this law is to create economic added value.

Furthermore **the Dutch** government is actively providing open data to the public and has 'open unless' as a starting point and is focused on providing as many datasets as possible. Currently more than 7,000 datasets on different subjects are freely available on the internet among which traffic data. For example: Current traffic data gives an up-to-date picture of the situation on the road. Every minute, data from more than 24,000 measurement locations in the Netherlands are processed in a database and distributed to customers. It concerns data about intensities, travel times, point speeds and vehicle categories.

Belgium has transposed the Directive 2013/37/UE into the federal legal system⁸. As far as the terms used and the definitions given to them are concerned, Belgian law departs from the European text in two respects, but only terminologically. As regards the scope of application, that of Belgian law seems to be slightly narrower than that envisaged by the Directive. Indeed, the federal text contains an exception that does not appear in the European text in that it places unfinished or incomplete administrative documents outside its scope of application. The most remarkable feature resides in the creation of a "Public Sector Information" sectoral committee whose task is to give prior authorisation for the communication of documents containing personal data to be used for re-use, and will also be able to deliver opinions on the "open data strategy" and anonymization methods used by public authorities. The same directive is also being transposed into regional law by every region separately.

One of the key development that Belgium is looking towards, is opening up data. The federal strategy of 'Open Data' was approved in July 2015. Public data belonging to the federal government must by definition be accessible, with a few exceptions based on privacy and security. In concrete terms, the Open Data strategy includes 14 guidelines, among others:

⁷ Bundesministerium für Verkehr und digitale Infrastruktur (Germany's federal ministry of transport and digital infrastructure)

⁸ La loi du 4 mai 2016 relatif à la réutilisation des informations du secteur public

- I. The goal is free reuse of data without reference to the source. This makes it easy to combine data sets for the development of innovative applications;
- II. Data are provided in technical formats that facilitate reuse;
- III. By 2020, the federal government will proactively make its data available, not just on demand;
- IV. A single federal portal will be set up with all publicly available and usable data;
- V. Maximum continuity will be ensured: "re-users" must be able to rely on the availability of data in the future as well; and
- VI. Each public service develops an open data strategy and appoints a responsible person. The Open Data approach covers among others aspects of the environment, mobility and economy. Belgium has already opened up some of its public data through its government open data portal data.gov.be, where more than 473 datasets related to transport are being provided.

As part of the Open Data Strategy, the Federal Public Services (FPS) Mobility and Transport has chosen to make available to the public a series of data from all the services comprising it. These data can be reused free of charge, in the spirit of the European Directive on the re-use of public sector information.

Spain has implemented the PSI Directive through *Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público*. This law is applicable to all Spanish public administrations: Central State Administration, autonomous communities, councils, state owned companies and public universities. However, some documents are not affected by this law, such as those that are potentially affecting the national security or those containing intellectual property rights.

Four types of reutilization are developed by this law:

- i. Documents which are not subject to any condition;
- ii. Documents subject to a license;
- iii. Documents that need to be requested to certain public administration; and
- iv. Exclusive agreements of disclosure of documents in order to provide a public service. This act is further developed through Royal Decree 1495/2011, which established new obligations for public administrations such as informing through their e-offices about the re-usable documents that they are keeping. The Royal Decree details the provisions for the state sector. It promotes and facilitates the availability of public sector information as much as possible. Out of those regulations Spain has developed its government open data portal⁹, which is available to the public free of charge.

Estonia has implemented Directive 2013/37/EU on the re-use of public sector information through the Public Information act, which entered into force in 2001 and was last modified in January 2018. Public information refers to information that is recorded and documented in any manner and on any medium, and which is obtained or created upon performance of public duties provided by law or legislation issued on the basis thereof. This Act ensures that the public

⁹ www.datos.gob.es , Accessed 18.05.2018

has the opportunity to access information intended for public use and it creates opportunities for the public to monitor the performance of public duties. It provides for the conditions of, procedure for and methods of accessing and re-using public information and the bases for refusal to grant access. It is important to note that the exchange of information between holders of information for the performance of their public duties does not constitute re-use of information. The most important feature of this act is that it regulates databases and it created an open data portal for Estonia¹⁰. It provides a single point of access for the general public to unrestricted public sector data with the permission to re-use and redistribute such data for both commercial and non-commercial purposes¹¹.

In countries outside the EU initiative to make data publicly available are also being taken.

In **Singapore**, the available legal framework specific for the field of transport is of technical nature and does not deal with the use, processing, sharing, etc. of data. The only reference to data that can be found is in the Road Traffic Act, where Article 6C-E provides for the rules for trials and use of autonomous motor vehicles. Such vehicles can be required to be equipped to capture and store sensor data and such records may be required to be kept. In other modes of transport no such rules are provided.

Although the legal framework surrounding data in Singapore is still to be developed, the country is working towards becoming a highly developed country through its Smart Nation initiative. Singapore plans to harness digital technologies, which includes several national projects, like National Digital Identity, E-Payments, Smart Nation Sensor Platform, Smart Urban Mobility and Moments Of Life. The Smart Nation initiative also provides for data to be open for public through online portals. For that purpose, the Government has launched the one-stop government portal¹² that provides access to its publicly available datasets, including 86 datasets in the field of transport. Besides the general governmental open data portal, Singapore also has an open data portal specifically for the transport sector, called LTA DataMall. It provides data on the road and passenger transport sector. LTA DataMall is part of a series of information, eServices and tools that are available to all land transport users in Singapore. This includes real-time data on bus arrival timings, taxi availability, traffic conditions and carpark availability.

2.3.3 Automated driving

Countries like Singapore, France, Italy, Estonia and Germany have all included the promotion and development of connected and automated vehicles into their national strategies. The goals of those strategies is simple – to develop a framework that will allow for the use of fully autonomous road vehicles. That requires fundamental changes in the current legislation. So far only Germany has taken the first step by introducing the Automated Driving Act (2017)¹³ which prescribes rules for the driver behaviour of self-driving cars by allowing them to give control

¹⁰ opendata.riik.ee/en, Accessed 18.05.2018

¹¹ Public information act 01.01.2018

¹² data.gov.sg, Accessed 15.05.2018

¹³ <https://www.bundesregierung.de/Content/DE/Artikel/2017/01/2017-01-25-automatisiertes-fahren.html>, Updated 12.05.2017, Accessed 24.05.2017

over to the car. It prescribes rules for the allocation of fault and liability, for which purposes each such car must be equipped with a black box. The law prescribes the type of data and the duration of keeping that data collected by the black box, it does not specify rules concerning the recording and deleting of such data nor rules about the protection of such data.

Germany is the first country to introduce such a legislation. Other countries remain at the strategy level, however it is recognized by all, that fundamental changes in legislation are necessary in order to fully allow automated driving into real traffic situations. The main concern in this field is the interaction between vehicles and infrastructure which presupposes the collection, processing and intelligent linkage of data. Automated driving requires accurate information (e.g. traffic situation, spatial conditions, the state of the infrastructure and the weather situation). The vast amount of information that is needed to collect for automated driving are creating new requirements to be met by cyber security for vehicles and infrastructure and the protection of data. Countries must develop clear and internationally applicable cyber security standards and data protection requirements. These data protection requirements are mostly considered to be connected with the personal data collected.

2.3.4 Smart mobility

Smart urban mobility is one of the main strategy points for most studied countries. Smart mobility requires leveraging both data and digital technologies. In most countries smart mobility is related and sought to enhance public transport. Other modes of transportation are being less considered under the smart mobility initiatives.

The development of smart mobility foresees the adoption of a new transport data collection technologies. That involved several consideration like (1) expansion of the information delivery channels to quickly and effectively reach out to the public; (2) quality of information; and (3) presentation and visualisation of information. That is a need to establish more open data formats, interfaces and communication protocols. Such vast amount of data necessary needs to be harmonised and standardized.

Most smart mobility policies are taken at city level and not at national level. Therefore, not many smart mobility policies are identified analysing the policies at country level. Some smart mobility policies have been identified for a non EU-country, **Singapore**. This is most likely because of the small size of the country.

- 1) Singapore has published its Smart Mobility 2013 strategy, which comprises of four key focal areas : informative, interactive, assistive and green mobility. Under informative, it foresees the adoption of new transport data collection technologies. That involves several considerations: expansion of the information delivery channels to quickly and effectively reach out to the public;
- 2) Quality of information;
- 3) Presentation and visualization of information. There is a need to establish more open data formats, interfaces and communication protocols. Data needs to be harmonized and standardized (ITS standardization strategy).

2.3.5 General policies

Some countries undertake initiatives to promote the use of big data in general. A couple of these are summarized below.

In the field of re-use of data the **United Kingdom** has enacted the *Re-use of Public Sector Information Regulations 2015 (RPSI)*. RPSI does not apply to information that would be exempt from disclosure under information access legislation, i.e. among others the Data Protection Act and the Freedom of Information Act. In the same field the UK has had in place since 2017 the *Digital Economy Act* that increases government access to public data. It provides the government free access to citizens' data, with the power to share said data among departments, local authorities and even researchers.

Estonian legislation in the field of transportation is divided into different modes – road, rail, passenger, freight etc. These acts however do not deal with the use of big data. The acts only foresees the creation of different databases related to either public transport, traffic, motor vehicles, etc. Such databases store both personal and non-personal data. All personal data fall under the scope of the Personal Data Protection Act, and the non-personal data is subject to the rules of the regulations governing such specific databases.

Estonia is in the forefront of electronic governance through the initiative E-Estonia, which is a movement by the government of Estonia facilitate citizen interactions with the state through the use of electronic solutions. It foresees (and has already produced) several e-services, which include i-voting, e-tax board, e-business, e-ticket, e-banking, e-school, University via internet, e-governance academy etc. The new Digital Agenda 2020 focuses on creating an environment that facilitates the use of ICT and the development of smart solutions. Some specific initiatives under that agenda are:

- 1) the creation of a Nordic Digital Infrastructure Institute - an international development centre for the joint development of X-Road, e-identity, digital signature and other components of the basic service infrastructure;
- 2) Greater control over the use of personal data; and
- 3) The development and increase of the public sector's capacity to apply data analytics solutions. The Digital Agenda has also produced a study on, Linked Estonia which is to provide input into the implementation of the Digital Agenda plans and to prepare changes in the Estonian legal system and policy measures necessary for the efficient implementation of activities related to the use of linked data.

In the field of transportation, Estonia has developed the Transportation Development Plan 2014-2020, which stems from the Digital Agenda 2020 and foresees the development of ITS. Moreover, real-time data transport infrastructure for the transport system, including other infrastructure investments (e.g. for collecting and transmitting road information based on sensors), remains to be developed.

3 Understanding big data policies in the private sector

While public authorities play a major role in planning, building and managing transport infrastructure, transport activities are usually carried out by private companies and individuals. Examining the big data policies in private companies helps us to understand the potential and the roles that big data could play in the economy, in the provision of innovative products and services or in process development.

Policies in the public sector and private sector operate differently. What we are interested in identifying are the policies that enable the big data business models to be implemented. These will range from adoption of standards to actual terms and contracts held with external parties, to documents that strictly dictate the internal processes of the company. The definition is left non-exhaustive on purpose, in order to avoid restricting the exploratory analysis.

Hence, we will focus on describing the actions that a company takes, and the legal and organizational conditions that enable the actions. This will be based on several big data business models for transport activities. The following sections describe selected use cases based on the case studies to be studied in the Work Package 3 of the LeMO project.

3.1 Railway operators

The network and operations management in railway transportation systems, or railway track systems (RTS), focus on **maintenance, operations and safety**. Maintenance ensures that the rail track does not fail during the operation of trains on the particular track to avoid delays, poor level of service, and even accidents or derailment. Operations ensures that the rail traffic runs smoothly and on schedule. The data required for both may overlap, such as traffic data as input to the maintenance plans, but they are analysed in different ways befitting the aim (e.g. monitoring real-time disruptions or suggesting time for maintenance). Ensuring safety during operations is about the continuous monitoring of the railway to avoid any unsafe conditions. We will focus our discussion here on only the use of big data in maintenance, operations and digitalisation of the RTS.

Maintenance of RTS can be categorized according to *corrective, preventive and condition-based*. Many different types of data are needed or can be used to develop a track maintenance plan (see *Figure 1*) that is dynamic, cost-effective and accurate. However, these are largely internally produced and supplied data: static (grade, curvature), own historical records (grinding history, ballast history, service failure, etc.) and from rail-service operators (traffic data). Collaboration to create the datasets and implement them are non-controversial, as the rail-service operators benefit from a good track maintenance plan. Also data, such as traffic data, are already produced to support the operations of rail-service operators. Hence, further investment to produce the data is not needed. The availability of data supports the use of big data techniques in maintenance planning.

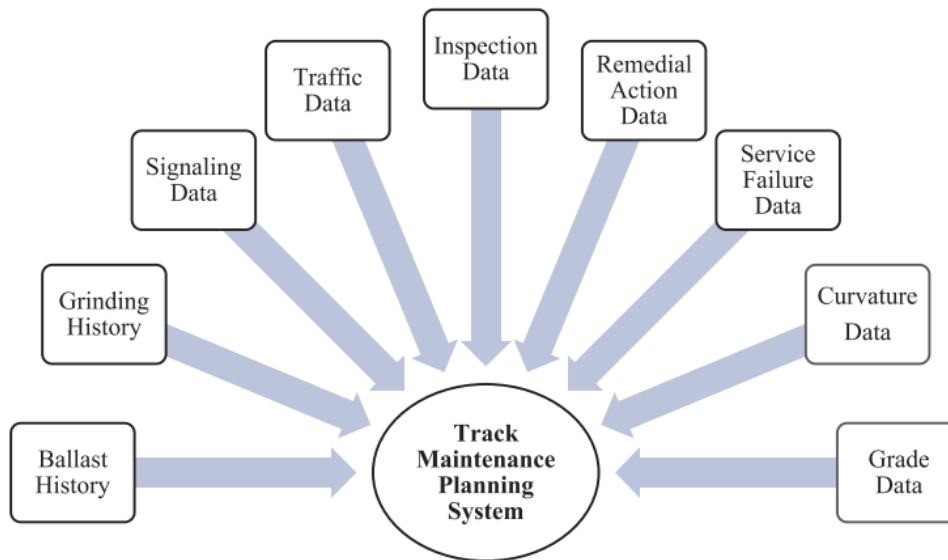


Figure 1 Data required for track maintenance planning systems (Source: (Ghofrani, He, Goverde, & Liu, 2018))

For railway operations, the main aims are to provide train users (whether passenger or freight) with sufficient information to manage their own activities, deal with disruptions, and make real time travel plans. From the operations side, identification of bottlenecks and service disruptions in real-time are important as a feedback to planning departments (i.e. network managers or transport service providers). Further, to make real-time adjustments in response to the disruptions, it is useful to also take into consideration the traffic or passenger demand. The sources of data here are therefore more complicated and diversity. Table 8 provides a summary of useful sources of data for the control of operations in passenger rail.

From Table 8, we see that most data are already readily available. For instance, rail-service providers have in their interest to share types (1) to (6) with the railway operator. In some cases, sharing the data is also mandatory in exchange for their permission to use the railway infrastructure. On the other hand, obtaining data types (7) to (9) may require agreements with other partners, such as ticketing agencies, smart card data operators, and payment system providers. As some data may be considered personal, privacy regulations may prohibit the sharing of ticketing data to any other party. Data might be available in aggregate form, which might still be useful for estimation and forecasts of transport demand and passenger traffic. However, if disaggregate data is still needed, anonymization of the data could facilitate the sharing of data. Nevertheless, at an aggregated level, such data is useful, especially if it can be linked with other data sources (e.g. weather, events, other traffic mode disruptions).

Data type (10), such as Twitter data, might also serve here to monitor passenger satisfaction (Pender, Currie, Delbosc, & Shiwakoti, 2013). Twitter also provides two-way communication, which can enable railway operators and service operators to broadcast or directly message the users they collect information from via the social media Application Programming Interfaces

(APIs¹⁴). However, despite the an international survey of railway agencies (sample size: 86) indicating that only one–tenth did not use social media during unplanned disruptions, a major concern is still the lack of staff resources to process the currently un-automated process (Pender et al., 2013). The automation of analysis could be done using big data techniques.

Table 8 Big data sources for RTS operations (Adapted from (Ghofrani et al., 2018))

Big data types	Typical contents
1) Train describer data	Track occupation and release times, train description steps, signal states (stop/go), switch states (left/ right)
2) Traffic control delay data	Delays at stations or other timetable points
3) GPS data	Train positions
4) Train event recorder data	Train positions and speeds, traction, brake applications
5) Traffic control incident registration data	Begin and end time of disruptions, failing elements
6) Timetable data	Arrival and departure times, train routes, stops
7) Ticket sales data	Tickets available
8) Automatic Fare Collection data (smart card data)	Passenger check-in and check-out times
9) Website data	Timetables, recommended travels and prices, train delays, disruption locations and times, online ticket sales
10) Social media data	Sentiment analysis

Digitalisation of the railways is one of the steps that the industry is taking to obtain greater control and efficiencies of the railway infrastructure. For instance, the Rail Delivery Group, who represents the rail industry in the UK, is supporting and pushing for the implementation of the Digital Railways schemes. The Digital Railways scheme includes the implementation of:

- European Train Control System (ETCS): this allows trains to run closer together and to travel at their best speeds while maintaining safe braking distances.
- Connected Driver Advisory Systems (CDAS) and Automatic Train Operation (ATO): these provide decision support to drivers in the cab so that they have the information they need at the right time to boost performance and safety.

¹⁴ Application Programming Interface (API) provides an interface for external systems to receive request and send responses, for instance for data requests.

- Traffic Management (TM): maximises performance as trains flow across the network, maximising the throughput that existing track can support and adapting in real-time as network conditions change to aid rapid recovery.
- Telecoms and data, through Fixed Telecommunications Network and Global System for Mobile Communications – Railway, will underpin and connect all these systems.

The aims of the digital railway programme are presented in the Figure 2. The benefits are compelling, as it could improve the management and control of the trains, improve the operation and maintenance of the physical infrastructure, and support customer service by opening the data to other parties.

The digital railway programme will...

Manage and control trains better

We can **safely and cost effectively** cater for increasing demand for rail travel and freight paths, automating control of trains to smooth the flow and **improve performance**.



It also enables the industry to...

Manage our physical infrastructure better

The system provides data that can be used to operate and maintain these assets in a **predict and prevent** way, aligning with the future operating model for how the industry will work together to maintain and operate the railway.



Provide better information

The open data generated by a digital railway will allow other parties to provide **better, more targeted information** about delays. It will provide greater level of real-time detail about the locations of trains between stations.



Figure 2 Description of the digital railway programme of the Rail Delivery Group¹⁵

A crucial element of the success of the digitalisation efforts depends on the ability of the Rail Delivery Group to mobilize its various members (which are Network Rail, the nationalised owner of the rail infrastructure; HS2; the train operating companies that provide passenger services; and the freight operating companies) in this direction. Membership in the Rail Delivery Group is mandatory for all actors in the rail industry. Development of the standards and procedures are also undertaken in collaboration with its members. However, it also obliges the implementation of the newer systems in keeping with technological improvements or government policy.

In terms of confidentiality of its data, the Rail Delivery Group is an unincorporated association owned by its members, hence data sharing is primarily between its own “members”. Furthermore, they create and share data, which are mainly “exhaust data” created from their

¹⁵ <http://digitalrailway.co.uk/our-role/industry-programme>, Accessed 18.05.2018

own internal process controls. That said the members are also obliged to share the data and adopt the practices set by the representative group.

In summary, for big data applications for the work of railway operators, the diverse types of data are increasing because of digitalisation of the railways operations in accordance with the movement of the industry. According to the literature reviewed, the most of the data, which hold value for the railway operators are readily available to them, with the exception of data collected from passengers on the sales, ticketing or social media. Since railway operators hold a strong (or more accurately a dominant) position in the rail services ecosystem, the sharing of sales and ticketing data could be obliged from the rail service providers. Access to social media information has also not been found to be difficult, although the process of gaining insights based on the social media data using big data analysis is still not widely implemented.

3.2 Open data in the airport operator context

Airports are already embedded in a mature and complex ecosystem (see Figure 3). Supporting the collaboration between industry partners and competitors are the standards and best practices adopted and promoted by the industry associations. Many are in response to regulations to improve the services of the critical and strategic airport infrastructure, which is critical for promoting tourism, facilitating business. However, they are also results of security concerns.

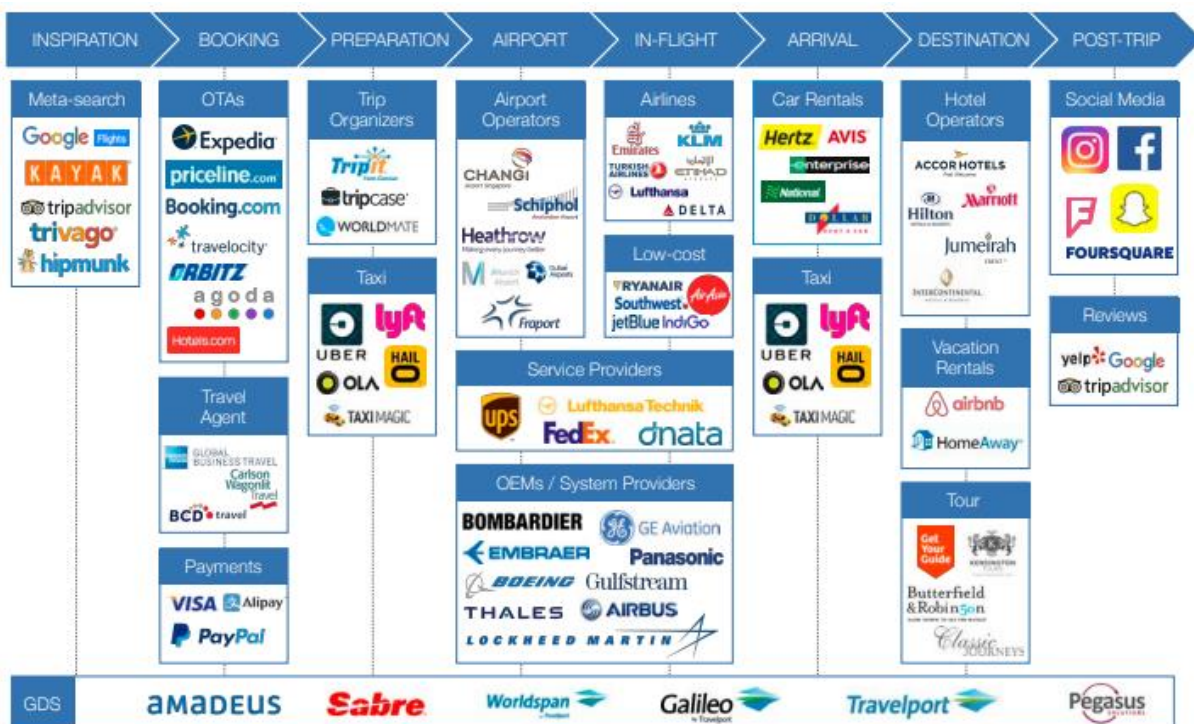


Figure 3 Illustration of the aviation, travel and tourism ecosystem (Source: (World Economic Forum, 2017))

Though the actors in the airport industry are obviously very diverse, industry associations such as International Air Transport Association (IATA) and Airports Council International (ACI), or Eurocontrol¹⁶ play a major role in driving their members in the same direction.

One such effort is the digitalisation of the airport roadmap and strategy. Digitalisation of the various control systems in the airport creates a variety of data, some that are useful for others to employ for analysis or developing other services. Some airports have made the available in an open way, as opposed to individual contracts. We discuss some of the policies relevant to open data in airports here. First, we describe the context of how some of the data is produced.

Marks & Rietsema (2014) summarize the main airport management information systems. The main aim is of these information systems is “to facilitate the airport and airline operations required to process aircraft, passengers, and air cargo: the ticketing of air travellers, ground movement of aircraft and vehicles, flight procedures of aircraft within airport airspace, and scheduling and managing of boarding and gate equipment, and weather updates.” These are composed of several distinct systems, which interact with each other significantly. Some examples are the Gate Management System, Aircraft Fuelling System, Air Traffic Control System, Weather Monitoring System, Airfield Lighting System, and Automatic Vehicle Identification System. The policy of ACI is that **centralized management of these systems** should be implemented wherever possible (ACI, 2009). A centralized management of these systems provides a single repository for many datasets, which might be easily “linked” to one another.

Most data produced in these systems would not have commercial value as open data. Nevertheless, in some cases they might. These data can be provided using open or commercial APIs to developers to develop new apps and improve the quality of the airports’ existing services - for customers, other businesses, employees, or other businesses to provide services to other customers (ACI, 2017). There is also economic and social pressure (or incentive) to digitalize the systems and enable new value to be co-created in the aviation industry. *Figure 4* summarizes the key drivers according to shareholder, customers, environment and society.

Though industry associations might show certain preferences, the airport must decide for their own **(open) sharing data policy**. ACI (2017) provides a set of considerations that can guide the strategy, which must then translate into the airports’ policy:

- The types of data that they are willing to share. This can be from the airport operations themselves, or even from third-party apps, which use their services in exchange for more data.
- The extent of the open data, which also depends on the type of data.
- The price attached to the data might also be important criteria to consider. If we think about the open data model, the data itself is considered to be “free”. However, there is a cost to access the information because of the infrastructure needed to support the external access. These could be negotiated or positioned strategically. The typical models are e.g., “charge everyone or selected organizations, or have a freemium model

¹⁶ <https://www.eurocontrol.int/network-manager>, Accessed 18.05.2018

with a service level agreement/support available for a fee, or offer pricing based on a monthly fee or on used API-calls.”

- The format for the data could be either proprietary or open data-exchange standards.

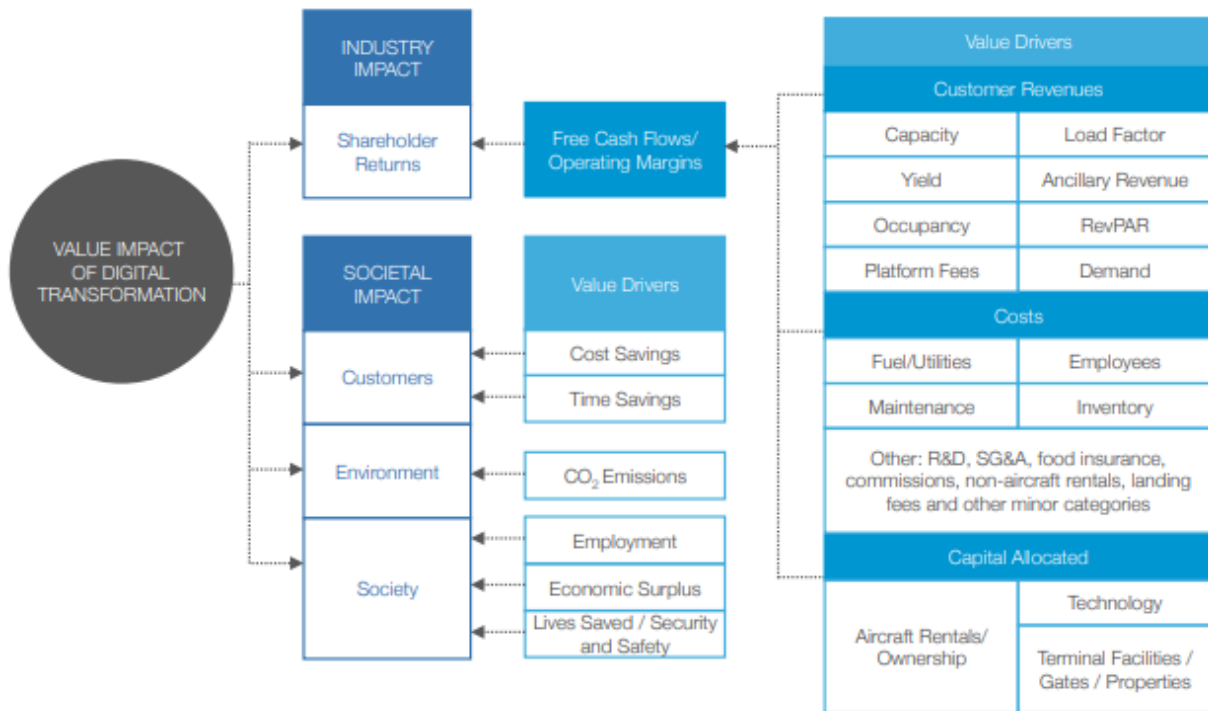


Figure 4 Value of digital transformation in aviation industry (World Economic Forum, 2017)

The Airport Community Recommended Information Services (ACRIS) working group under the ACI has developed an exchange standard called the **Seamless Travel Data Model** with the following benefits (Rogal & Stiffel, 2017):

- No bilateral negotiation between partners to define data model and interfaces;
- Common understanding of the data content;
- Simple familiarization with defined use cases;
- Same interface between different partners for B2B as well as B2C scenarios;
- Fast implementation by reusing implementation files (i.e. JSON definition files);
- High quality of pre-existing and proven interfaces; and
- Data exchange with many different partners via one common interface.

What this goes to show is that technical standards are currently being developed to further the industries foray into data sharing, as a **data supplier**. In fact, the role of open data is part of a larger roadmap towards digital airports.

Figure 5 summarizes how these technologies are to be further implemented in the airports, with the supporting infrastructure and policies to be developed internally.

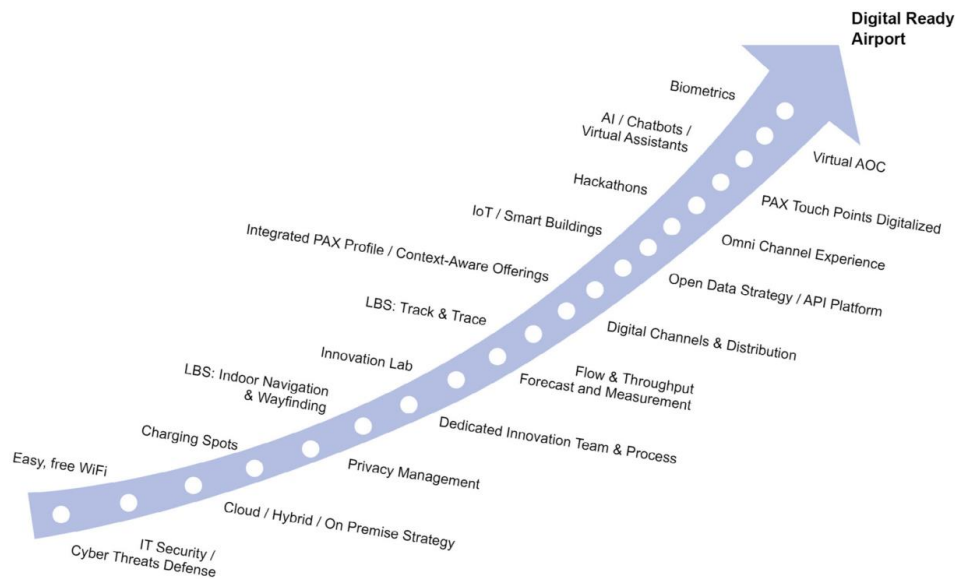


Figure 5 Enabling technologies for becoming a digital airport (ACI, 2017)

Schiphol Airport is an airport that is **actively supporting Open APIs**¹⁷. Schiphol offers several models for granting access to the information. This depends on the types of users and the purpose of using the APIs. Screening the types of data users is reasonable to ensure security and privacy concerns. Open APIs¹⁸ are “an open description format for API services that is vendor neutral, portable and open.” The types of APIs and a brief description, and types of data are presented in Table 9.

The types of users that the Open APIs are targeted at:

- Airlines & Handlers
- Airport operations
- Travel & Transport
- Contractors
- IT professionals
- Students

Many of the users are for improvements to the operations of their own contractors and employees. However, Schiphol also targets (road) transport providers to and from the airport, where they can improve the coordination of their services or even offer full-package solutions. In addition, IT professionals are given access to support the development of new business models, which might enhance passenger and freight transport services. Finally, an important target group are students who are allowed to experiment with the data APIs and learn more about the industry.

¹⁷ <https://www.schiphol.nl/en/developer-center>, Accessed 18.05.2018

¹⁸ <https://www.openapis.org/about>, Accessed 18.05.2018

Table 9 Types of Open APIs offered by Schiphol Airport

API type	Data and information	Brief description
Flight API	A wealth of information on current and scheduled flights to and from the airport.	<p>The API provides detailed flight information about:</p> <ul style="list-style-type: none"> • Destinations • Airlines that operate flights to and from Schiphol • Aircraft types operated by the airlines • Flight data elements • Status of flights
Operational Flight API	Additional operational information on current and scheduled flights to and from the airport.	<p>Operational flight information and data is key to successfully operate an airport. It can add value to the travellers' journey and the processes of airlines, ground handling agents and other stakeholders.</p> <p>The Operational Flight API makes this operational flight data available in a controlled and standardised way, to all parties involved.</p>
Wait Times API	Predicted wait times for various locations where queues might arise.	The Wait Times API provides estimated wait times, based on the current number of people in line and the processing capacity at that time. These predicted wait times can be displayed anywhere, e.g. on screens at security, in the app and on the website. The aim is to manage the travellers' expectations, give them more control over their traveller's journey and thus reducing travel stress.
Wayfinding API	Relevant information on routes and POIs at the airport, for travellers, airlines, floor managers and other staff.	<p>The Wayfinding API consists of three individual geospatial services:</p> <ul style="list-style-type: none"> • POIs – Point of Interest at Schiphol, including shops, sanitary facilities, restaurants, stairs, elevators and much more. • Routes – Wayfinding functionality for all possible routes at Schiphol. • Basemap – An airport reference layer on which users can plot POIs and routes between them.
Boarding Pass Service	Provides airlines direct information about the number of passengers who successfully scanned their boarding pass at the security filter.	The Boarding Pass Service helps airlines reduce the delay of flights, due to absent passengers. Knowing if passengers are already at Schiphol or not, airlines may decide to depart or wait for a passenger. The boarding pass scan identifies (the number of) passengers who scanned their boarding pass at security and provides this information to the airlines, who can identify these passengers in their own systems with a sequence number. The Boarding Pass Service is available via the boarding pass scan API.
Beacon Registry API	Provides details about the many crowd control beacons hanging all over the terminal buildings.	<p>The Beacon Registry API helps travellers to flow through the terminal more easily. For this, Schiphol installed over 2000 beacons in the terminal. Combined with other API's it can be used for location based services such as indoor wayfinding or proximity services.</p> <p>The Beacon Registry API offers the following data:</p> <ul style="list-style-type: none"> • Latitude and longitude per beacon. • UID per beacon. • Area name where the beacon is located. • Landside or airside locator. • Floor level of beacon (Z-index).

Of note also is an event held to encourage innovation for exploiting mobility data: the Dutch Mobility Hackathon 2017¹⁹ held on 17 and 18th November 2017 in cooperation with Royal Dutch Touring Club and the railway company of Netherlands Nederlandse Spoorwegen. The Hackthon (as with most similar events) spark a lot of interest from professionals and amateurs alike to think about new ways to make use of the big data streams. Schiphol's policy to support and attract attention to their data is supportive of open data initiatives and is a best practice for other airports to follow suit.

In summary, airport operations has led the way in digitalisation of their operations to improve operations and to support collaboration between the diverse actors. The step of providing Open Data through Open APIs is done in the hope for a greater variety of services that can help enhance the attractiveness of the services in the airport, increase efficiencies, and support new business model. While the airport industry may acknowledge the benefits, little is known on the extent that airports around the world have implemented the strategies and roadmaps.

3.3 Real-time road traffic management

The road counterpart to the railway network operator is the real-time road traffic management systems. These are actually a major part of the work that ITS do²⁰. In general, the aims of ITS and how they are measured are²¹:

- **Safety** is measured through changes in crash rates or other surrogate measures such as vehicle speeds, traffic conflicts, or traffic law violations.
- **Mobility improvements** are measured in travel time or delay savings, as well as travel time savings, and on-time performance. Travel time reliability is emerging as a new measure of travel dependability.
- **Efficiency** is typically represented through increases in capacity or level of service within existing road networks or transit systems.
- **Productivity improvements** can be documented in cost savings to transportation providers, travellers, or shippers.
- **Energy and Environment** benefits are typically documented through fuel savings and reduced pollutant emissions.
- **Customer Satisfaction** findings document the perception of deployed ITS by the traveling public, usually in the form of survey results.

ITSs are managed by road authorities, and so does not fall under the private sector responsibility. As a public agency, they have access to a lot of real-time information captured by state-owned sensors. In addition, they collect other types of information from third-party

¹⁹ <https://www.schiphol.nl/en/developer-center/page/dutch-mobility-hackathon>, Accessed 18.05.2018

²⁰ [https://www.itsbenefits.its.dot.gov/its/benecost.nsf/images/Reports/\\$File/Taxonomy.pdf](https://www.itsbenefits.its.dot.gov/its/benecost.nsf/images/Reports/$File/Taxonomy.pdf), Accessed 18.05.2018

²¹ <https://www.itsknowledgeresources.its.dot.gov/its/bcllupdate/>, Accessed 18.05.2018

aggregators or suppliers. Table 10 provides an overview of the available data sources, which may be used.

Table 10 ITS data sources and data collection technology (Khan, Rahman, Apon, & Chowdhury, 2017)

Data sources	Data collection technology	Data type	User
Roadway data	Loop detector	Volume, speed, classification occupancy, presence	Public agency
	Vision-based technology (CCTV camera)	Volume, speed, classification, occupancy, presence	Public agency
Vehicle-based data	Floating car data (with GPS and cellular network)	Vehicle position, travel time, speed, lateral and longitudinal acceleration/deceleration, obstacle detection	Public and private agencies
	Connected vehicles	Vehicle position, travel time, speed, lateral and longitudinal acceleration/deceleration, obstacle detection	Public and private agencies
Traveller-based data	Twitter, Waze	Real-time alerts, incident detection	Public and private agencies
Wide-area data	Photogrammetry	Traffic monitoring, incident management, transportation planning and design	Public agency

Roadway data are the primary sources of information usually found in the traffic management systems. Loop detectors for instance are commonly used to detect vehicle movement on particular lanes/links. Vision-based technology need to be coupled with image processing, which increasingly are based on machine learning technology. These can then be used to understand more about the types of vehicles, and even record vehicle license plates. Wide area data using satellite imaging for instance (or drones) can be used to collect long term data.

The public agencies may also **work together with private sector or public transport agencies** to collect relevant information on traffic. Floating car data can be provided by private vehicles through third-party navigation services or from fleets of public agencies (e.g. municipality services or public transport vehicles) or private companies (e.g. taxi companies, freight carriers, car-sharing platforms).

A few examples of private companies providing floating car data for public usage are:

Uber Movement²² collects data from users/drivers using the Uber Apps in cities and shares the data. Regarding privacy, “all data is anonymized and aggregated to ensure no personally identifiable information or user behaviour can be surfaced through the Movement tool. All data shared through Movement adheres to Uber’s privacy policy, and at no point will Movement

²² <https://movement.uber.com>, Accessed 18.05.2018

provide a means for partners to access individual driver or rider details in any way, shape or form.”²³

Waze’s Connected Citizens Program²⁴ collects and shares location data of its users to support “cities, departments of transportation and first responders” in the planning and operations. For privacy, “Waze will never share individual driving history or non-public user information with any partner”. The data is shared “for the partner to retrieve through a localized XML or JavaScript Object Notation (JSON) feed that is updated every two minutes.” The types of data include “filtered data for the specific, monitored area including system-generated traffic jams and user-reported traffic incidents (including jams, accidents, hazards, construction, potholes, roadkill, stopped vehicles, objects on road, and missing signs).”

The benefits according to Waze of the partnerships are (WAZE, 2018):

- **Situational Awareness:** Partners receive real-time incident information faster than other reporting methods and accurately pinpoints where incidents occur, creating faster response and clearing times and potentially saving lives
- **Two-Way Driver Communication:** Partners leverage Waze as a two-way communication channel: Partners use Waze to inform drivers of major traffic events and drivers communicate back real-time road insights through the app
- **Infrastructure Planning:** Insights into locations with frequent congestion or hazards yields smarter urban planning
- **Bridging Connections With Other Partners:** Waze gathers partners via in-person summits and an online forum to discuss case studies and exchange ideas to further impact communities globally
- **Streamlining Data Inputs:** Partners can utilize data standards designed by Waze for closure and incident reporting to reduce data fragmentation and promote transport and government data aggregation.

For both platform services (Uber and Waze), **private operators are not obliged to share the data with public agencies**. They must be also be sufficiently anonymized before sharing to protect their customers privacy. However, by partnering with public agencies, they might be able to increase the legitimacy of their services to the public. Also, private operators usually expect something in return from public agencies, such as good quality data, such as maps or traffic information. Public agencies’ open data policies help in this regard to encourage also private operators to share what they have.

These platform services bypass the services that in-built navigation systems or “connected vehicles” provide, as both services are delivered via the smartphone app. This allows it to be used by all sorts of vehicles, regardless of make or model. A similar business model could be expected for future connected vehicle manufacturers to “open” their data to public agencies.

²³ <https://movement.uber.com>, Accessed 18.05.2018

²⁴ <https://www.waze.com/ccp>, Accessed 18.05.2018

An example of a road traffic management system is the **Regional Integrated Transportation Information System (RITIS)** created by the i95 Corridor Coalition²⁵, which incorporates a wide variety of data streams including that of connected vehicles. The data flow is depicted in Figure 6. RITIS demonstrates one part of the traffic management system: the fusion and visualization of the traffic state. Based on understanding the traffic state, control of the road utilization can be exerted via traffic signals, road signs, and other traffic management tools.

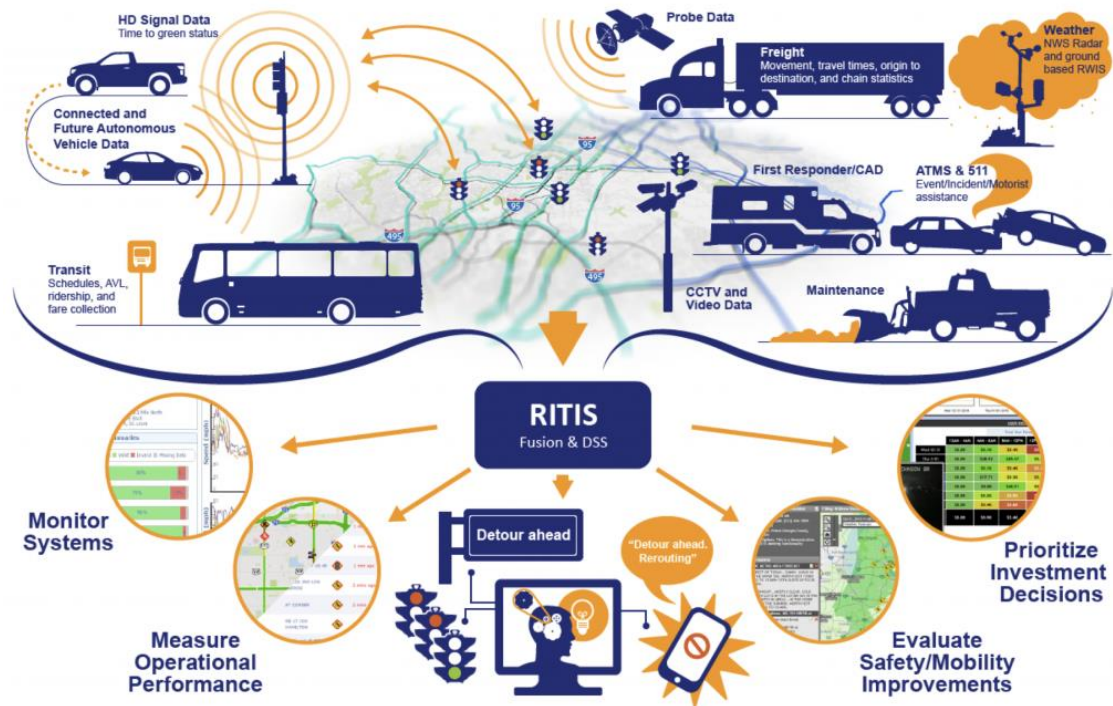


Figure 6 Overview of function of RITIS

There are three main RITIS components including: 1) real-time data feeds, 2) real-time situational awareness tools, and 3) archived data analysis tools. These are described briefly in on their website and summarized in Table 11.

The RITIS therefore works very much with big data, in **helping planning agencies to access the data**. It works as an intermediary between agencies and state borders, and third-party data suppliers. It is publicly funded and developed and operated by a public employees, the Catt Laboratory²⁶. Generating insight and helping road authorities to monitor (or as they call it being “situationally aware”) is an important role here.

²⁵ <http://i95coalition.org/>, Accessed 18.05.2018

²⁶ <http://www.cattlab.umd.edu/>

Table 11 Description of RITIS components²⁷

Component	Aim and description	Example applications
1) real-time data feeds	provide direct access to real-time incident, event, detector, probe, weather, transit, and other data sources including ITS devices.	<ul style="list-style-type: none"> - developers that need access to real-time information for dynamic mobility applications. - control data elements to be provided in the data feed or maintain secure and secluded from other agencies or the public.
2) real-time situational awareness tools	view all of the real-time RITIS data in a browser using a dynamic set of tools for situational awareness.	<ul style="list-style-type: none"> - interact with live events, incidents, weather, sensors, radio scanners, response vehicles, and other data sources and devices in maps, lists, and other graphics. - Users can apply a rich set of filters, access contact information, and even set up alerts.
3) archived data analysis tools	Query and view data from archived data. All data within RITIS is archived indefinitely, for download or export.	<ul style="list-style-type: none"> - to identify accident hot-spots, analyse queue lengths and traffic congestion/bottlenecks at specific areas, perform after-action reviews, and evaluate the effectiveness of transportation operations strategies.

In summary, the use of big data in real-time road traffic management is a major focus of the industry providing the road management systems. The ability to integrate data from a variety of government operated sources, as well as from private sources is important to these systems. Private companies that hold these data, like floating car data, are willing to share with public agencies for a variety of reasons, such as in exchange for data or to reinforce their legitimacy as a service provider.

3.4 Big data in supply chain management

The section focuses on big data in supply chain management (SCM). Throughout the history of SCM, several conceptual approaches have been developed in step with the improvement of technology and the rising requirements of SCM as a whole. A commonly used approach in the retail sector is the “Efficient Consumer Response” and “Quick Response” approaches to partnerships along the supply chain (Zairi, 1998), which aim at leveraging the whole supply chain for better performance (i.e. better reaction sales demand at a lower cost). The difficulties in the approach lie primarily in the difficulty of predicting the necessary stocks, especially that required for safety buffers to avoid stock-outs, and for suppliers to manage the operations with “increased flexibility and better predictability” (Zairi, 1998).

The variety of data available and which might be useful include “the upstream source, which is the suppliers’ side, through the intermediate stream source, which is the manufacturers’ and consolidation points or warehousing side, and finally the downstream side, which is the logistics and distribution and/or retail side” (Addo-Tenkorang & Helo, 2016). The variety can be seen in Figure 7, which presents a simple taxonomy based on the 3 (first) Vs of Big Data.

²⁷ <http://i95coalition.org/projects/regional-integrated-transportation-information-system-ritis/>

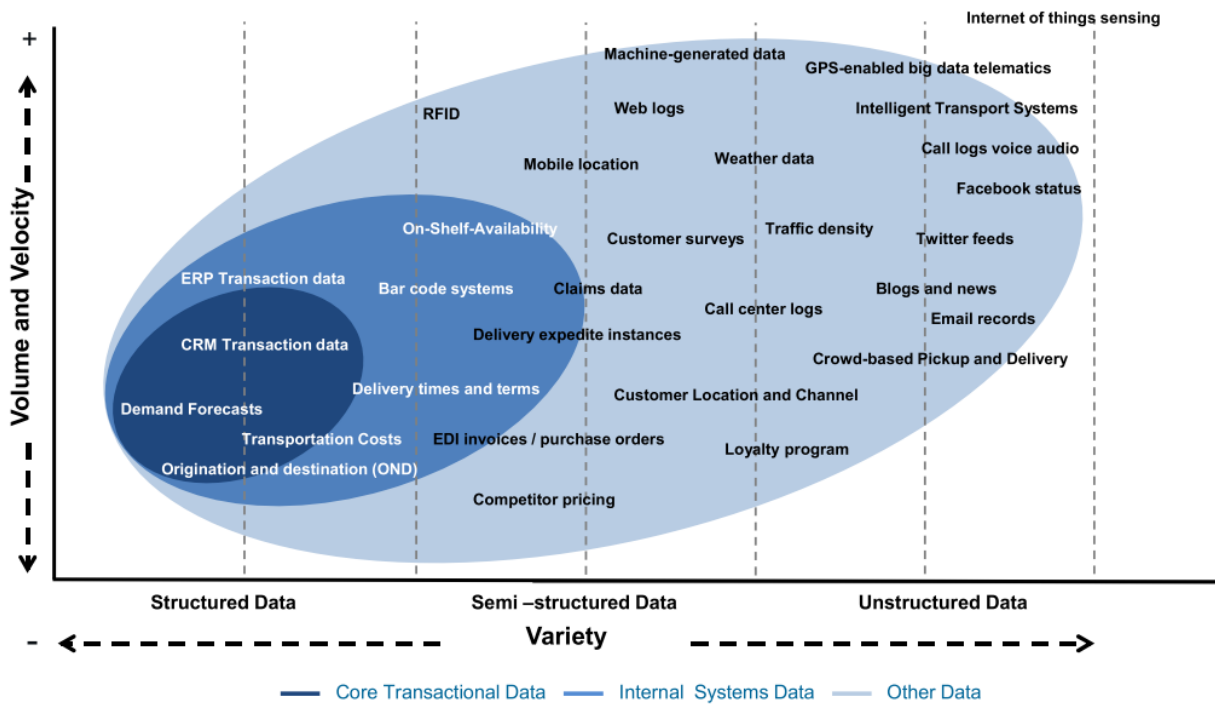


Figure 7 SCM data according to volume, velocity and variety (Rozados & Tjahjono, 2014)

Some of the **applications of big data in supply chain** are presented in Table 12.

Table 12 Applications of big data in supply chain operations (KPMG, 2017)

Supply chain operations	Descriptions and examples
Demand forecasting	More accurate estimation of demand by accessing data of sales, market trends, competitors' data, and relevant local and global economic factors.
Warehousing	Real-time conducting of big data analysis within the warehouse Enterprise Resource Planning (ERP) system and identifying inventory levels, delivery mismatches and incoming deliveries.
Inventory planning	Full transparency at the Stock-Keeping Unit (SKU) level and fully automated replenishment systems combined with demand forecast data that eliminate under/overstocking and optimize inventory ranging.
Distributing	Real-time optimization of complex webs of distribution hubs, plants and warehouses based on the material flow data.
Scheduling	Increased visibility of inventory levels, demand and manufacturing capacity; hence more accurate production and distribution scheduling.
Delivery	Monitoring of delivery routes, traffic data, weather in real-time and rerouting if necessary for capacity and asset sharing.

The application of big data in this area, which can be used to estimate the demand using downstream data, holds great potential. Some of the key themes, here are **timely delivery of retail data and forecasted sales data, without compromising privacy or business confidentiality**. However, as Whiteoak points out: “Despite an increasing general willingness to work together there remain many attitudes, prejudices, corporate cultures to be changed and hidden agendas to be exposed if we are to explore the full range of possibilities” (Whiteoak, 1999).

In short, the industry still needs to sort out their own *policy* in terms of data integration with close upstream and downstream partners. Part of this boils down to **the issue of trust between the partners**. Sharing data is both a sign of trust and a means to obtain trust, as Sahay notes “further sharing of vital information leads to building of trust by itself, since there is a shared sense of vulnerability, of the use/misuse of the information by the other party for opportunistic behaviour. The fact that such information is being supplied makes the partners feel more secure in their relationship” (Sahay, 2008).

Supply chains of the future are also expected to be more demand driven and electronically monitored and controlled. Besides providing visibility of the performance and status of the supply chain, digitalized supply chains, more importantly support “the collecting and analysing information and data on processes, energy and source materials used during manufacture and on how the product was used and performed”, which helps the partners design and implement integrated value creation processes (Swink, Davis, & Graybill, n.d.).

Figure 8 shows conceptually how a digitalized supply chain can **support new services and functions based on the new data created**. Point-of-sales (POS) data supports the pull principles in manufacturing and supply chain, as well as the concept of vendor managed inventory. Without close to real-time information on sales, such functions cannot be easily implemented. Based on the Internet of Things concept, Radio-Frequency Identification (RFID) can be placed on goods or load units, providing additional data that improves inventory management, as well as provide real-time information on the condition of products (e.g. temperature-controlled products). As a manufacturer, coordination with your supplier using their data will help to control, plan and mitigate supplier risks. Data on your internal processes will also help to manage operations and optimize maintenance plans. Finally, Global Positioning System (GPS) data from logistics vehicles will help in optimizing delivery plans.

However, this may bring certain dangers regarding **sharing confidential information between partners**. As Todeva and Knoke explain “alliance participants risk potential opportunism from their partner’s unrestricted access to proprietary secrets and patented processes” (Todeva & Knoke, 2005). Furthermore, it is unclear what the role of transparency in the partner’s processes and the negotiation of prices and terms of the partnership is. Although manufacturing companies are moving forward with digitalisation, the “integrated value creation” may still be difficult to achieve unless sufficient intermediary processes or services can help to protect each companies’ own interests.

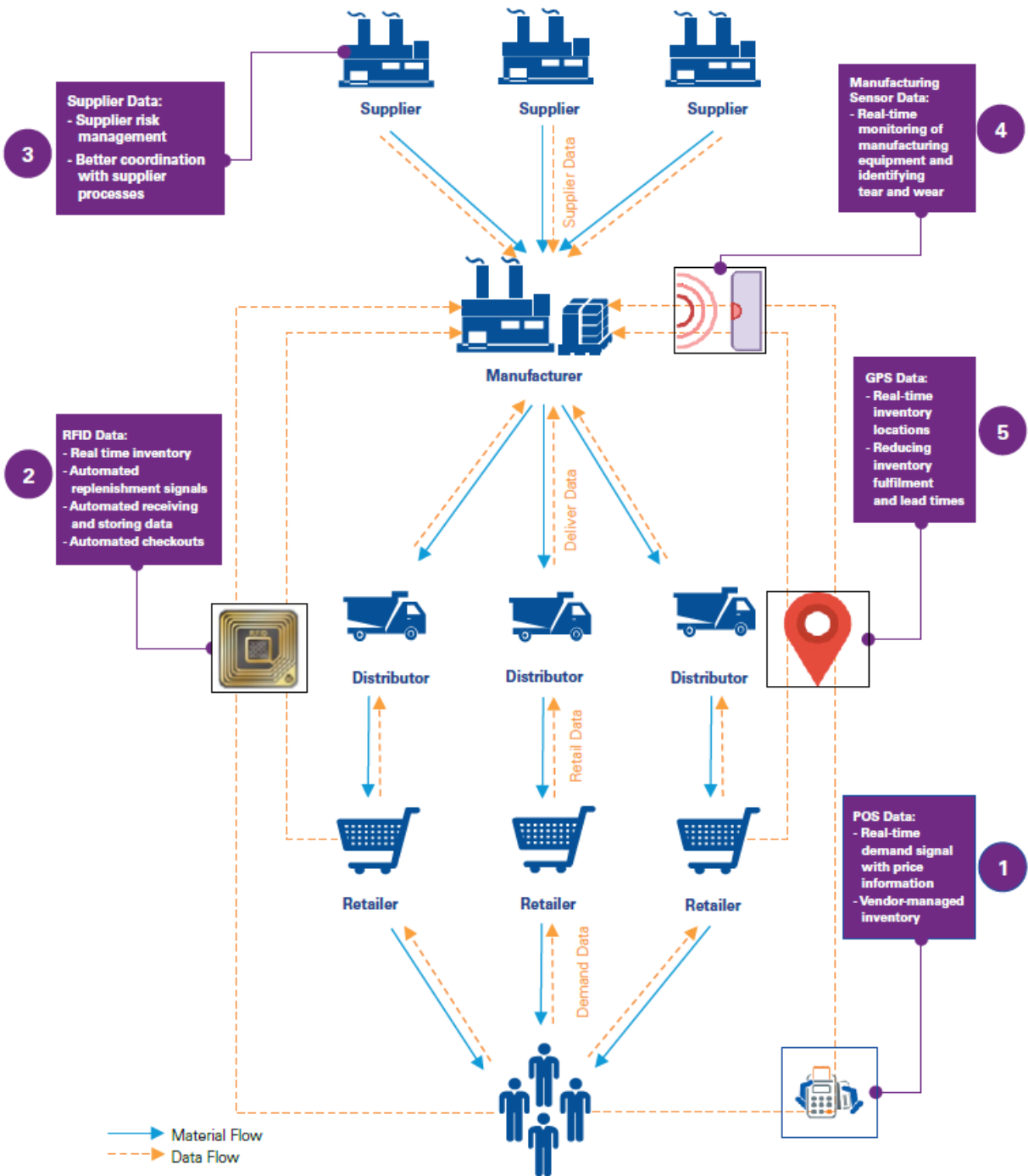


Figure 8 Illustration of data in the supply chains of the future (KPMG, 2017)

The industry needs, but does not yet have, **best practices to guide data sharing practices**. Learning from similar applications in other sectors, the following are recommended best practices for proper governance of Data-as-a-Service²⁸ (DaaS) (Sarkar, 2015a):

- Obtain executive sponsorship for the DaaS program.
- Identify the key stakeholders and data stewards for various divisions and subject areas in the organization.
- Improve reliability and trustworthiness of the published data services.
- Standardize core data assets across the organization to reduce maintenance and operational efforts.
- Align the interests of business/IT stakeholders with DaaS objectives, the mission, and goals.
- Define a common vocabulary for all key business terms and the hierarchy among them.
- Protect sensitive data from unauthorized access.

These DaaS best practices apply as well to the data sharing practices between B2B partners. However, it is uncertain how many companies have applied such practices, and whether the investment in the organizational restructuring is justified. Hence, this approach still raises a risk of an unbeneficial investment.

In summary, big data applications in SCM, particularly in digitalized supply chains are possible. Gaining transparency and visibility along the supply chain is important for manufacturers. Improving the management of risk and operations are also potential applications. Yet, to date the efforts are not yet wide-spread. Individual companies have their own constraints to deal with in digitalisation of their processes. Policies to support data sharing are not entrenched although best practices from other applications can be transferred.

3.5 Managing port operations

The management of ports, whether for sea shipping or inland waterways is a complex undertaking because of the high level operational efficiency and capability needed and that it is a hub that connects several major transport modes, and various governmental and private stakeholders.

There are two main tasks that ports need to handle.

1. Physical traffic management of vessels and cargo
2. Ensuring regulatory compliance for vessels and cargo

As an infrastructure operator, vessel operators expect **high levels of efficiency and throughput in the port operations** and the facilitation of the loading or offloading of cargo. It also has to facilitate intermodal transfers of cargo, which require stacking and retrieval of containers,

²⁸ DaaS is a “data-distribution and publication framework that can be used by organizations to engage their customers in real time” (Sarkar, 2015b)

transport to the intermodal facilities and vehicles. A number of ports have automated these operations: of gates, yards and quay cranes (Martín-Soberón, Monfort, Sapiña, Monterde, & Calduch, 2014))

Next to the traffic management, ports are often gateways into countries, which means that **regulatory requirements** also need to be observed for the incoming vessels and cargo. Facilitating trade and improving port operations (including pre- and post-processes) is the concept of the “Single Window”. A Single Window is defined as “a facility that allows parties involved in trade and transport to lodge standardised information and documents with a single entry point to fulfil all import, export, and transit-related regulatory requirements.” (European Port Community Systems Association EEIG, 2011). These require e.g. communication on all customs related issues.

There are various systems in place to support the capture and control of data flow required for ports. For instance, **RIS** “collect, process, and disseminate fairway, vessel movement, vessel traffic and transport information” to support navigation on all types of inland waterways (RISING, 2011). RIS data support (RISING, 2011):

- Voyage planning of Inland Waterway Transport (IWT) operators providing data on water level, water depth, maximum height/bridges, berth availability, lock occupation (actual and predictions/forecasts) used for routing, stowage planning, etc.
- Intermodal transport planning by enabling IWT operators to pass actual information on their schedules to the intermodal planners / freight integrators
- Voyage monitoring for IWT operators, freight integrators, inland port operators, sea port operators providing status information, e.g. vessel positions, passing waypoints, missing administrative reports, predictions of problems in continuation of the voyage
- Provision of Estimated Time of Arrival updates including reliability time windows
- Lock planning and accompanying information flows affecting IWT-based transport and hence logistics chains as well
- Fleet management of Inland Navigation equipment, like barges/lighters, by identifying their current position and status of operation
- Inland and sea port operations, e.g. berth management as well as collecting port and canal dues, Electronic Chart Display and Information System-based information for skippers, Single Window information service concept as being developed in other European projects
- Intermodal transport execution by facilitation relevant information about transport progress.

Another development, which integrates even further, especially in terms of supporting the Single Window concept is the **Port Community Systems (PCS)**. Such a system is (European Port Community Systems Association EEIG, 2011) “is a **neutral and open electronic platform** enabling intelligent and secure **exchange of information between public and private stakeholders**” and it “optimises, manages and automates port and logistics efficient processes

through a single submission of data and connecting transport and logistics chains.” A schematic of how data flows change are presented in *Figure 9*.

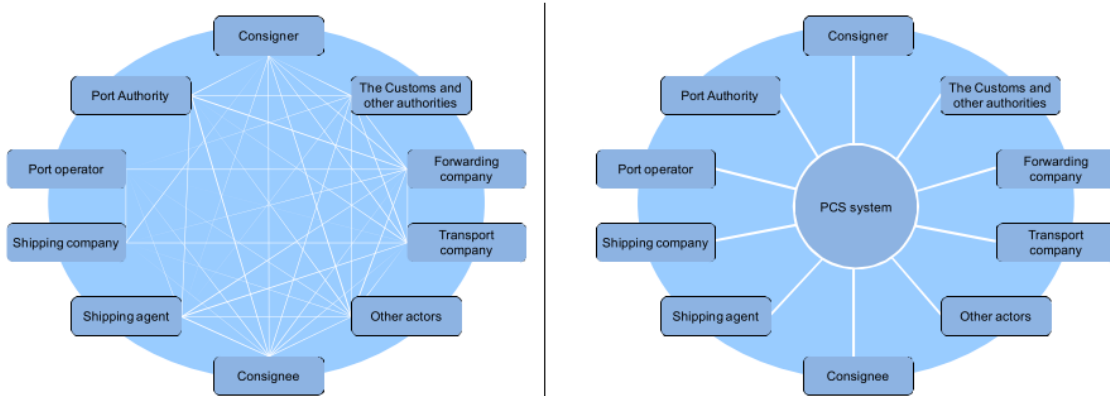


Figure 9 Traditional communication in a port versus communication with a port community system (Posti, Yliopiston, Koulutus-Ja, & Julkaisuja, 2012)

The functions of a PCS are multi-layered: organised in both functions of the transport, logistics, and regulatory domains, as well as the various application modules are presented in *Figure 10*.

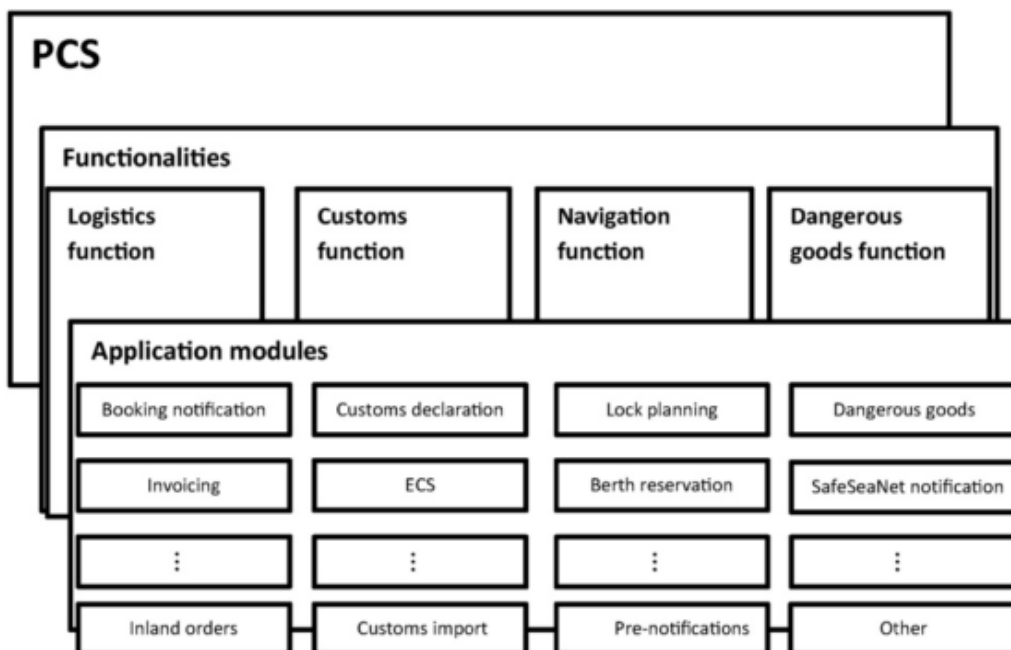


Figure 10 Multi-layer structure of PCS functions and modules (Carlan, Sys, & Vanellander, 2016)

As the PCS provides a variety of services, the benefits are also diverse (see Table 13). This will of course be different in quality and strength of the impact to each stakeholder. The benefits can attract the different actors to join the PCS, which then increases the amount of data being shared around. However, **there is no obligation for ports to evolve into a PCS or adopt any other regular exchange of information.**

Table 13 Benefits and cost for the PCS (Carlan et al., 2016)

Digital economy benefits	Community attendance benefits	Cost to PCS operator (administrator)	Cost to PCS user (supply chain actor)
Reduced access cost to information	Access to real-time information	Platform development costs (including hardware/software acquisition/development cost, staff cost, training, etc.)	Connection cost
Reduced cost of communication	Higher value-added services	Operating cost (staff cost, maintenance cost, storage and data management costs)	Hardware/software acquisition/development cost Training cost
Less illegal transactions	Compliance with community standards and regulations		Transaction fees (licence fees)
Decreased rate of errors			
Reduced rate of data inconsistency			
Fast access to information			
Efficient use of resources			

While the primary role of the PCS systems are to ensure that the data streams and the various services are well managed, the **data collected could be used for prediction of transport demand** (ALICE, n.d.). Here, it is important to ensure the quality of the data that is collected and the underlying data infrastructure. Still even a bit of data is useful for optimizing in container intermodal transport (Zuidwijk & Veenstra, 2015).

From a PCS perspective there is much concern about **cybersecurity risks and about the ownership of data** (International Port Community Systems Association, 2015). While some of the data held by the PCS is public, some are private and sensitive and pose business or security risks. Since some data is “linked” with others from various stakeholders, the security and access policies might also extend beyond the *responsibility of the PCS*. The access to the data must be clarified to (and understood by) the actual owners or contributors of the data. In terms of security, as the PCS is responsible for many functions, it must be guarded strongly from any intrusion and any errors.

From a cybersecurity standpoint the following are important elements in the strategy to develop and emphasize (International Port Community Systems Association, 2015):

- **Authorisation:** The right people should have the right access to the right data: people are often the weakest link and there should be strong policies on who is authorised to use the PCS and at what level. This should be decided and controlled on a ‘need to know’ basis. Access must be restricted – not everyone needs to know or access everything.
- **Vigilance and Preparedness:** There are individuals and organisations focusing entirely on trying to break into security systems or data – for financial gain or fraud purposes, as part of a terrorist threat, or simply to outwit the system. A PCS’s resources, and those of the wider logistics chain, must match this or at least limit their chance of success.

- **Monitoring and Traceability:** PCSs must constantly monitor their systems to eliminate any weaknesses or vulnerabilities. Processes should be in place to identify any breaches and pick up on any unusual ways of working or strange behaviour, such as large amounts of data being removed.
- **Mitigation and Business Continuity:** PCSs must be seen as Trusted Third Parties. Reliability of service, security backup and contingency plans are critical.

In summary, in port operations the development of PCS will require the use of big data techniques to manage the data, as well as to develop actionable insights in real-time management and planning. PCS are third parties – not owned or operated according to the legislation, therefore trust in the industry is important. Ensuring that the system is safe for members to share data with is paramount, hence the measures of cybersecurity need to be given high priority. However, it could be that in some cases that persuasion is not necessary, as Port Authorities might oblige the participation of the private sector actors in the PCS.

3.6 Connected and automated vehicles

The partner to smart cities and infrastructure are connected and automated vehicles. Both are two distinct types of technology concepts but are related in that automotive research are pushing the agenda forward together and in parallel. To a certain extent AVs require connected vehicle technologies to support its tasks. We describe the two technologies before moving touching on the policies that support these technologies from the automotive manufacturers' point of view.

Connected Vehicle (CV) systems build upon the development of dedicated short-range communications (DSRC) to enable coordination between vehicles and other actors and systems in road infrastructure. There are several categories of CV systems depending on the communication partner or subject: V2V (vehicle to vehicle); V2I (vehicle to infrastructure), I2V (infrastructure to vehicle); V2P (vehicle to pedestrian); and V2X (vehicle-to-anything). The functions they support are presented in Table 14.

Automated Vehicles (AVs) on the other hand are about automation of previously human functions in road vehicles. SAE International defines 5 levels of automation. Level 1 and 2 have been implemented in some systems, whereas Level 3 to Level 5 are still in the experimental stages. In Level 5, the vehicle can perform all driving tasks without the driver.

Table 14 Functions of connected vehicle systems (summarized from Shladover, 2017)

<p>V2V</p>	<ul style="list-style-type: none"> - cooperative collision warnings and hazard alerts, as tested in the Safety Pilot Model Deployment; - cooperative collision mitigation or avoidance, incorporating active braking; - cooperative adaptive cruise control, with tighter vehicle-following control than conventional adaptive cruise control and enhanced traffic flow stability; - close-formation automated platooning, enabling aerodynamic drafting and lane capacity increases; - automated manoeuvre negotiation at merging locations or intersections; - transit bus connection protection.
<p>I2V</p>	<ul style="list-style-type: none"> - providing traffic signal status information in real time for in-vehicle display, signal violation warning, or green wave speed advisories to drivers; - providing traffic and weather condition information and real-time routing advisories to drivers; - fleet management functions of vehicle routing and scheduling; - access control to closed facilities; - variable speed limits and advisories provided directly to drivers or their vehicles (I2V cooperative adaptive cruise control); - end of queue warnings; - active support for lane guidance.
<p>V2I</p>	<ul style="list-style-type: none"> - vehicle probe data applications providing detailed traffic information (speed, volume, travel time, queue length, and stops) or road surface condition information (pavement roughness or slippery conditions) - mayday and concierge services (such as OnStar); - electronic toll collection and parking payments; - traffic signal priority requests; - vehicle status information for fleet management (especially for transit and trucking fleets)
<p>V2P</p>	<p>Safety features for any vulnerable road user, including pedal cyclists, who may be carrying a nomadic device that can communicate with nearby vehicles.</p>
<p>V2X</p>	<p>At the most general level, analogous to the current interest in the ‘internet of things’ in which virtually every device could be connected to any other device.</p>

SAE level	Name	Narrative Definition	Execution of Steering and Acceleration/Deceleration	Monitoring of Driving Environment	Fallback Performance of Dynamic Driving Task	System Capability (Driving Modes)
Human driver monitors the driving environment						
0	No Automation	the full-time performance by the <i>human driver</i> of all aspects of the <i>dynamic driving task</i> , even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
1	Driver Assistance	the <i>driving mode</i> -specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the <i>human driver</i> performs all remaining aspects of the <i>dynamic driving task</i>	Human driver and system	Human driver	Human driver	Some driving modes
2	Partial Automation	the <i>driving mode</i> -specific execution by one or more driver assistance systems of both steering and acceleration/deceleration using information about the driving environment and with the expectation that the <i>human driver</i> perform all remaining aspects of the <i>dynamic driving task</i>	System	Human driver	Human driver	Some driving modes
Automated driving system ("system") monitors the driving environment						
3	Conditional Automation	the <i>driving mode</i> -specific performance by an <i>automated driving system</i> of all aspects of the dynamic driving task with the expectation that the <i>human driver</i> will respond appropriately to a request to intervene	System	System	Human driver	Some driving modes
4	High Automation	the <i>driving mode</i> -specific performance by an automated driving system of all aspects of the <i>dynamic driving task</i> , even if a <i>human driver</i> does not respond appropriately to a request to intervene	System	System	System	Some driving modes
5	Full Automation	the full-time performance by an <i>automated driving system</i> of all aspects of the <i>dynamic driving task</i> under all roadway and environmental conditions that can be managed by a <i>human driver</i>	System	System	System	All driving modes

Figure 11 Description of different levels of automation in vehicles²⁹

Taking over from the human driver, **the AV must now depend on in-vehicle sensors and the data communicated via CV systems**. Big data techniques have a large role to play in processing the amount of information very quickly with very low latency.

Automotive manufacturers work together with industry associations (such as Society for Automotive Engineers, Institute of Electrical and Electronics Engineers), with **the makers of systems**, such as Siemens, and **with road authorities** to develop and test the technologies.

Some of the benefits of connected and AVs include³⁰:

- **Crash Elimination:** Crash-free driving and improved vehicle safety, a vehicle can monitor the environment continuously, making up for lapses in driver attention.
- **Reduced Need for New Infrastructure:** By managing traffic flow, self-driving can reduce the need for building new infrastructure and reduce maintenance costs
- **Travel Time Dependability:** V2V, V2C, and V2I can substantially reduce uncertainty in travel times via real-time, predictive assessment of travel times on all routes
- **Productivity Improvements:** A reduction in driving tasks will allow travellers to use travel time more productively

²⁹ <https://connectedautomateddriving.eu/technologies>, Accessed 18.05.2018

³⁰ http://autocaat.org/Technologies/Automated_and_Connected_Vehicles, Accessed 18.05.2018

- **Improved Energy Efficiency:** Reduced energy consumption in at least three ways: more efficient driving; lighter, more fuel-efficient vehicles; and efficient infrastructure
- **New Models for Vehicle Ownership:** Self-driving vehicles could lead to a major redefinition of vehicle ownership and expand opportunities for vehicle sharing
- **New Business Models and Scenarios:** Convergence of technologies may realign industries such that companies need to compete and collaborate at the same time

An example of how the concept can develop is **Traffic Management 2.0³¹ (TM2.0)** ERTICO Platform, an initiative which “originated in 2011 from TomTom and Swarco and was formally established on 17 June 2014 during the ITS Europe Congress in Helsinki.” It is currently chaired by AustriaTech and Dynniq with more than 38 members from all ITS sectors focusing on new solutions for advanced active traffic management.

Its aims include:

- Agreement on common interfaces, principles and business models which can facilitate the exchange of data and information between the road vehicles and the Traffic Management and Control Centres
- Improving the total value chain for consistent traffic management and mobility services, while avoiding conflicting guidance information on the road and in the vehicles.

The main actors that need to be considered in this section (Vlemmings et al., n.d.) are:

- **Authorities:** administrations operating the road-side systems; implementing the TM plans; collecting traffic data and providing information to drivers by means of road-side systems.
- **Road Infrastructure Owners:** cities, regions, national road authorities; concessionaires; responsible for TM goals and/or plans (including safety, efficiency and comfort aspects).
- **Road-side / Content Service Providers:** data broker and/or aggregator of different data sources.
- **In-car Service Providers / Telecoms Providers:** OEMs (Original Equipment Manufacturers), telematics and navigation service providers, mobility mobile apps; managing interaction with consumers inside the vehicle.
- **Service Consumers:** professional and private drivers.

³¹ <http://tm20.org>, Accessed 18.05.2018

The interactions of the various actors and systems are illustrated in Figure 12.

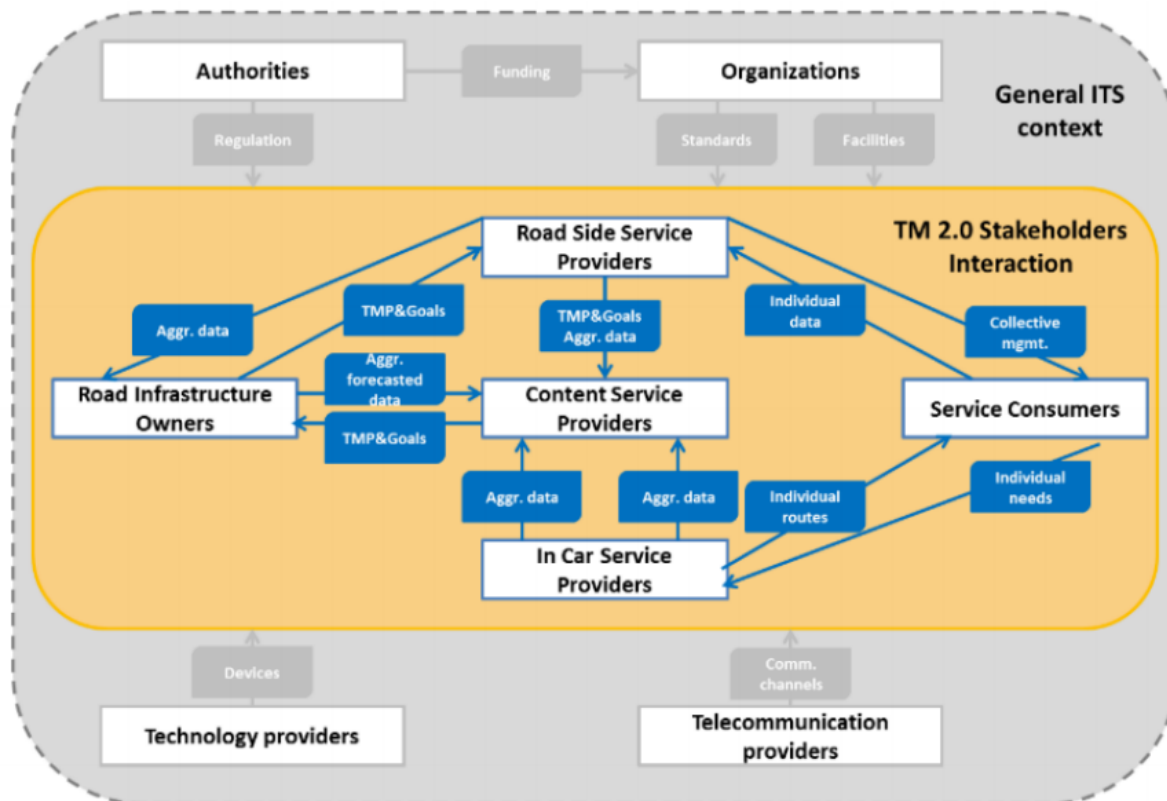


Figure 12 Traffic management 2.0 stakeholder overview and interactions (Vlemmings et al., n.d.)

It is an industry-driven initiative that looks at developing the infrastructure that support connected vehicles. One can simply state that their aim is to ensure the **optimised use of road infrastructure**.

On the other hand, the **European Automobile Manufacturers Association (ACEA)** represents the 15 Europe-based car, van, truck and bus manufacturers (i.e. BMW Group, DAF Trucks, Daimler, Fiat Chrysler Automobiles, Ford of Europe, Hyundai Motor Europe, Iveco, Jaguar Land Rover, Opel Group, PSA Group, Renault Group, Toyota Motor Europe, Volkswagen Group, Volvo Cars, and Volvo Group), which are also looking at the aspects of CVs and AVs but from a different standpoint. Their main loyalty lies with their customers, who purchase and use their vehicles.

While they recognise the benefits connectivity can bring to society, **they understand their responsibility** of not only the pleasure and convenience of driving, but also the safety of their passengers, to the extent that they can influence it.

The basic principles (ACEA, 2016a) for providing the vehicle data to any third-party including the governmental agencies are:

1. **Customer choice:** Vehicle users can obtain services from the vehicle manufacturer, his network of authorised repairers, independent aftermarket operators or any other service provider that has concluded a B2B agreement with the vehicle manufacturer.
2. **Fair competition:** Repair and maintenance information that is made available to the vehicle manufacturer's network of authorised repairers will be made available to independent aftermarket operators on non-discriminatory conditions (type, amount and quality of data, delivery time and price) in accordance with EU law. Other service providers will have access to a defined dataset to offer their services in accordance with the B2B agreement concluded with the vehicle manufacturer.
3. **Privacy and data protection:** In accordance with EU and national data protection and privacy laws, personal data of vehicle users will be made available to service providers only with the consent of the vehicle user except where a legal requirement or a contract exists. Service providers shall use this data only for the purpose(s) for which the vehicle user gave his or her consent.
4. **Safety, security and liability:** With the exception of regulated access to data for emissions control, diagnosis, repair and maintenance, data access must occur only through off-board means since direct third-party access to vehicular electronic systems would jeopardise safety, (cyber)security and vehicle integrity. Having regard to the vehicle manufacturer's obligations under product liability law, the responsibility for ensuring secure end-to-end communication between the vehicle and the off-board facility must remain exclusively with him.
5. **Interoperability:** The means of access and the interface(s) must be standardised to ensure interoperability. The ISO standard 20078 is being developed specifically for this purpose. It provides for web service access to the 'extended vehicle' as defined in ISO standard 20077-1. The extended vehicle consists of a physical road vehicle with external software and hardware extensions that are developed, implemented and managed by the vehicle manufacturer.
6. **Return on investment:** Service providers who use vehicle data for commercial purposes shall compensate vehicle manufacturers for all costs incurred, for example in generating the data and in developing, operating and maintaining the access facility and, where appropriate, for the market value of the data.

As data suppliers, who invest many resources into the research and development, business development and manufacturing, as well as undertaking much of the liability associated with the vehicles, they have a big stake in **protecting their strategic importance in the data value chain**. As their Strategy Paper outlines "In the absence of an adequate legal framework, a very small number of companies based outside the European Union could rapidly acquire the same dominant position in the area of in-vehicle services as they already have in the field of data processing, search engines, online services or smartphones. Should this occur, vehicle manufacturers risk being left with stranded investments, a loss of company know-how, commercial secrets and industrial property rights. The consequences for the competitiveness

of the auto industry, service providers and for job and value creation in Europe would be significant.” (ACEA, 2016b)

The primary actions to enable a **“fair” data ecosystem** would be

- to provide their customers with a high level of personal data protection.
- to design their vehicles and services so that where possible customers can choose whether to share personal data on a per-use contractual basis, with the freedom to deactivate geolocation functionality, unless required by contractual or legal obligations.
- to place contractual safeguards when data processing of personal data is outsourced.
- vehicle manufacturers should be paid usage fees for the development, operation and maintenance costs of accessing data.

There are also **cybersecurity concerns**. They are “fundamentally willing to share selected vehicle data with third parties provided this occurs in a way that meets strict requirements for road and product safety, as well as data security, and does not undermine their liability.” Their position is also that “third-party applications that interact with the vehicle should only be developed and approved in cooperation with the vehicle manufacturer to eliminate security, data protection and product liability risks.”

The **standards for extended vehicles**, i.e. a road vehicle with extensions for open yet protected access interfaces for selected software and hardware. The standards associated with it are ISO 20077³². Access to some data can then be safely and in a controlled way be supplied to third-party data users (such as commercial services, public agencies or traffic management systems). The “extended” part ensures that safety and security critical systems are not opened, and therefore the vulnerability to cybersecurity threats are kept minimal. The types of data accessible are also differentiated according to “intellectual property of the manufacturer, safety-critical applications, telematics and infotainment applications” where the data user and usage purposes are clearly defined and can be provided under different contractual terms.

In summary, the range of big data applications in the field of automated and connected vehicles are many, but they also depend on a wide-range of private actors. These private actors have their own reasons for not sharing the data, which range from privacy concerns, loyalty to customers, protecting proprietary technologies to the recognition of data as a new revenue source and competitive edge. While the industry is moving towards sharing some types of data, they generally will retain ownership of the data, where possible and where it allows them leverage in the market for profit. Future road management systems might require the use of the data held by the vehicles (though belonging to the private vehicle users).

³² <https://www.iso.org/standard/66975.html>, Accessed 18.05.2018

4 Conclusion

The work presented in the report highlights the state of the policies undertaken and promoted in the public and private sector to further their agendas in relation to big data in the transport sector.

In the different transport sectors, policies and initiatives have been developed in the EU to support the access, (re-) use, linking and sharing of data. Preceding and in light of these developments, the private sector has also moved ahead to incorporate the use of big data techniques into their own business models as process or product innovations.

The development of a policy roadmap to foster the growth of big data in transport will require an understanding of how existing policies affect the economic, political, social and legal environment for government and private agencies in the transport sector. This work will be carried out Work Package 2.

References

- ACEA. (2016a). ACEA Position Paper: Access to vehicle data for third-party services. Retrieved from http://www.acea.be/uploads/publications/ACEA_Position_Paper_Access_to_vehicle_data_for_third-party_services.pdf, Accessed 18.05.2018
- ACEA. (2016b). ACEA Strategy Paper on Connectivity. Retrieved from http://www.acea.be/uploads/publications/ACEA_Strategy_Paper_on_Connectivity.pdf, Accessed 18.05.2018
- ACI. (2009). Airport automation and e-business. In *Policy and Recommended Practices Handbook (Seventh Ed, pp. 1–17)*.
- ACI. (2017). Airport Digital Transformation BEST PRACTICE. ACI World.
- Addo-Tenkorang, R., & Helo, P. T. (2016). Big data applications in operations/supply-chain management: A literature review. *Computers & Industrial Engineering*, 101, 528–543. <https://doi.org/10.1016/j.cie.2016.09.023>
- ALICE. (n.d.). Roadmap “Corridors, Hubs & Synchronomodality.” Retrieved from <https://www.etp-logistics.eu/wp-content/uploads/2015/08/W26mayo-kopie.pdf>, Accessed 18.05.2018
- Carlan, V., Sys, C., & Vanelslender, T. (2016). How port community systems can contribute to port competitiveness: Developing a cost-benefit framework. *Research in Transportation Business and Management*, 19, 51–64. <https://doi.org/10.1016/j.rtbm.2016.03.009>
- European Port Community Systems Association EEIG. (2011). The role of Port Community Systems in the development of the Single Window. Retrieved from http://tfig.unece.org/pdf_files/A9R149C.pdf, Accessed 18.05.2018
- Ghofrani, F., He, Q., Goverde, R. M. P., & Liu, X. (2018). Recent applications of big data analytics in railway transportation systems: A survey. <https://doi.org/10.1016/j.trc.2018.03.010>
- International Port Community Systems Association. (2015). Cybersecurity: In the Maritime and Logistics Supply Chain - Workshop summary, conclusions and recommendations. Retrieved from <http://ipcsa.international/armoury/resources/ipcsa-cybersecurity-workshop-final-2.pdf>, Accessed 18.05.2018
- Khan, S. M., Rahman, M., Apon, A., & Chowdhury, M. (2017). Characteristics of Intelligent Transportation Systems and Its Relationship With Data Analytics. *Data Analytics for Intelligent Transportation Systems*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-809715-1.00001-8>
- KPMG. (2017). How big data is shaping the supply chains of tomorrow. Retrieved from <https://assets.kpmg.com/content/dam/kpmg/au/pdf/2017/big-data-analytics-supply-chain-performance.pdf>, Accessed 18.05.2018

- Marks, A., & Rietsema, K. (2014). Airport Information Systems—Airside Management Information Systems. *Intelligent Information Management*, 6, 149–156. <https://doi.org/10.4236/iim.2014.63016>
- Martín-Soberón, A. M., Monfort, A., Sapiña, R., Monterde, N., & Calduch, D. (2014). ScienceDirect Automation in port container terminals. *Procedia -Social and Behavioral Sciences*, 160, 195–204. <https://doi.org/10.1016/j.sbspro.2014.12.131>
- Pender, B., Currie, G., Delbosc, A., & Shiwakoti, N. (2013). Social Media Utilisation during Unplanned Passenger Rail Disruption – What’s not to “Like”? *Australasian Transport Research Forum*, 2. Retrieved from http://atrf.info/papers/2013/2013_pender_currie_delbosc_shiw.pdf
- Posti, A., Yliopiston, T., Koulutus-Ja, M., & Julkaisuja, T. (2012). E-PORT - Improving the efficiency of Finnish ports with intelligent systems - Final report of the Mobile Port project. Retrieved from https://www.utupub.fi/bitstream/handle/10024/79905/A58_E-port.pdf
- RISING. (2011). RIS Services for Improving the Integration of Inland Waterway Transports into Intermodal Chains. Retrieved from http://www.rising.eu/c/document_library/get_file?p_l_id=11324&folderId=123235&name=DLFE-2607.pdf
- Rogal, M., & Stiffel, S. (2017). ACI - AIRPORT COMMUNITY RECOMMENDED INFORMATION SERVICES WG SEAMLESS TRAVEL- BUSINESS REQUIREMENT DOCUMENT.
- Rozados, I. V., & Tjahjono, B. (2014). Big Data Analytics in Supply Chain Management: Trends and Related Research. <https://doi.org/10.13140/RG.2.1.4935.2563>
- Sahay, B. S. (2008). Industrial Management & Data Systems Understanding trust in supply chain relationships. *Industrial Management & Data Systems* *Https International Journal Business Process Management Journal*, 103(4), 553–563. Retrieved from <https://doi.org/10.1108/02635570310497602>
- Sarkar, P. (2015a). Ensuring Organizational Success. In *Data as a Service:A Framework for Providing Reusable Enterprise Data Services*.
- Sarkar, P. (2015b). Overview of Fundamental Concepts. In *Data as a Service:A Framework for Providing Reusable Enterprise Data Services*.
- Shladover, S. E. (2017). Connected and automated vehicle systems: Introduction and overview. *Journal of Intelligent Transportation Systems*, 1–11. <https://doi.org/10.1080/15472450.2017.1336053>
- Swink, D., Davis, J., & Graybill, R. (n.d.). BEST PRACTICES IN CYBER SUPPLY CHAIN RISK MANAGEMENT: Smart Manufacturing - The Future of Manufacturing and Value Chain Competitiveness. Retrieved from https://www.nist.gov/sites/default/files/documents/itl/csd/NIST_USRP-SmartManu-Cyber-SCRM-Case-Study.pdf, Accessed 18.05.2018

- Todeva, E., & Knoke, D. (2005). Strategic alliances and models of collaboration. *Management Decision*, 43(1), 123–148. <https://doi.org/10.1108/00251740510572533>
- Vlemmings, T., Vroom, O., Tzanidaki, J., Vreeswijk, J., Hofman, P., Spoelstra, J., & Rodrigues, N. (n.d.). Contractual Agreements in Interactive Traffic Management – looking for the optimal cooperation of stakeholders within the TM 2.0 concept. Retrieved from http://2r1c5r3mxgzc49mg1ey897em.wpengine.netdna-cdn.com/wp-content/uploads/sites/8/2017/06/TM2.0-TF10_Contractual-Agreements_Final-Report.pdf, Accessed 18.05.2018
- WAZE. (2018). Connected Citizens Factsheet. Retrieved from <https://web-assets.waze.com/partners/ccp/WAZE-CCP-Factsheet.pdf>, Accessed 18.05.2018
- Whiteoak, P. (1999). The realities of quick response in the grocery sector A supplier viewpoint. *International Journal of Physical Distribution & Logistics Management*, 29(7/8), 508–520. <https://doi.org/10.1108/09600039910292019>
- World Economic Forum. (2017). Digital Transformation Initiative: Aviation, Travel and Tourism Industry - In collaboration with Accenture. Retrieved from <http://reports.weforum.org/digital-transformation/wp-content/blogs.dir/94/mp/files/pages/files/wef-dti-aviation-travel-and-tourism-white-paper.pdf>
- Zairi, M. (1998). Best practice in supply chain management: the experience of the retail sector. *European Journal of Innovation Management*, 1(2), 59–66. <https://doi.org/10.1108/14601069810217239>
- Zuidwijk, R. A., & Veenstra, A. W. (2015). The Value of Information in Container Transport. *Transportation Science*, 49(3). <https://doi.org/10.1287/trsc.2014.0518>